

**Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.**

Preprint

ISSN 2198-5855

**An adaptive multi level Monte-Carlo method with
stochastic bounds for quantities of interest with
uncertain data**

Martin Eigel, Christian Merdon, Johannes Neumann

submitted: December 20, 2014

Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: martin.eigel@wias-berlin.de
christian.merdon@wias-berlin.de
johannes.neumann@wias-berlin.de

No. 2060
Berlin 2015



2010 *Mathematics Subject Classification.* 35R60, 47B80, 60H35, 65C20, 65N12, 65N22, 65J10.

Key words and phrases. partial differential equations with random coefficients, uncertainty quantification, multilevel monte carlo, adaptive methods.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: <http://www.wias-berlin.de/>

ABSTRACT. The focus of this work is the introduction of some computable a posteriori error control to the now widely popular Multilevel Monte Carlo sampling for PDE with stochastic data. We are especially interested in applications where some quantity of interest should be estimated accurately. This is a typical question in the geosciences, when e.g. groundwater flow with rather rough stochastic fields for the conductive permeability is examined. Based on a spatial discretisation by the finite element method, a goal functional is defined which encodes the quantity of interest. The devised goal-oriented a posteriori error estimator enables to determine guaranteed path-wise a posteriori error bounds for this quantity. An adaptive algorithm is proposed which employs the computed error estimates and adaptive meshes in order to control the approximation error and the stochastic error in probability likewise. Moreover, it also allows for the adaptive refinement of the mesh hierarchy used in the Multilevel Monte Carlo simulation which we use for a problem-dependent construction of discretisation levels.

Numerical experiments illustrate the performance of the presented method for a posteriori error control in Monte Carlo and Multilevel Monte Carlo Methods with respect to localised goals. Moreover, the computational efficiency of the multilevel and classical Monte Carlo approaches are compared. It is illustrated that adaptively refined meshes can yield significant benefits when combined with Monte Carlo methods. In particular, with a problem-adapted mesh hierarchy, the efficiency gains of the Multilevel Monte Carlo method in terms of computational cost can also be exploited in the context of error control.

1. INTRODUCTION

Simulations of PDE with stochastic data have become an indispensable tool in many application areas. We are especially concerned with groundwater flow models with limited knowledge about the geological media. A common approach for these problems is to assume some random field for the conductive permeability. The parameters for such a field can e.g. be based on a set of actual sample measures of the real structure of the ground. Since the feasible number of measurements is limited, this modelling is subject to a large amount of uncertainty which propagates to any computed solution based on this data. Typically, the media is assumed to be rough due to small correlation lengths with large contrast, rendering it challenging to simulate efficiently. The usual stochastic models are based on lognormal conductivity fields. Its logarithm is a (stationary) Gaussian field. For other applications, fields which are expanded in a finite set of uniform random variables are common. Then, the task is to compute some stochastic solution field and related statistics or a (stochastic) quantity of interest depending on the solution. Such a dependent goal functional is often localised to a certain region of interest, i.e. a subdomain of the physical domain $D \subset \mathbb{R}^d$ ($d = 1, 2, 3$). In particular, one commonly is interested in the expected value or the variance of the solution in some subdomain.

There are different conceptual approaches for this kind of uncertainty quantification for PDE with stochastic data. The properties of the stochastic data considered here, namely the dependence on a large number of random variables (which means many stochastic dimensions) due to the low regularity of the fields, are rather problematic for numerical methods which rely on the stochastic discretisation in orthogonal polynomials, the so called (generalised) polynomial chaos. Typically, the usage of stochastic Galerkin [25] and stochastic collocation methods [4] is most beneficial if some higher regularity of the

solution can be expected. Otherwise, the number of stochastic dimensions may cause the algebraic system to become prohibitively large. Recent reviews of numerical methods and their analytical properties can be found in [31, 19].

For this reason, for their apparent simplicity and straightforward parallelisation, Monte Carlo (MC) methods enjoy a great popularity with this type of problem. However, in the classical form, in order to achieve sufficient accuracy, the well-known slow convergence necessitates the generation of a huge number of samples where for each sample path the solution of a PDE with a realisation of the stochastic data has to be calculated. The common variance reduction techniques are only able to reduce the constant factor of the error but can not improve the convergence rate. Under certain assumptions on the solution, this can be achieved by using systematically generated sets of correlated sample points which is done in Quasi Monte Carlo (QMC) methods [23, 18].

Another approach (which can also be combined with QMC sampling) is to employ a nested set of discrete spaces. With this, it is possible to exploit the fact that the solution of the problem PDE on the coarsest space is relatively cheap and, since only components of increasing frequency but with decreasing importance (so-called corrections) are added subsequently on finer spaces, a rather small number of solutions is required on the finest space. A crucial observation then is that most of the uncertainty is already captured by the coarse space solutions. This multilevel idea is quite common in the theory of algebraic solvers for PDE and can be transferred beneficially to the problem at hand. The resulting Multilevel Monte Carlo (MLMC) method has the advantage to scale proportionally with the complexity of a single PDE solution on the finest level, at least asymptotically. It thus can be highly superior to standard MC sampling in terms of computational efficiency.

While the complexity of MLMC and the possible convergence rates in dependence on the regularity of the stochastic conductivity have been studied extensively in [7, 13, 12, 32], to our knowledge, a problem dependent spatial adaptivity of the discretisation based on some a posteriori error estimation has not been considered yet. However, an adaptive approach yields the potential to substantially reduce the complexity of MLMC calculations by focusing the computational effort on the quantity of interest, neglecting areas of the domain which do not contribute in a relevant way to this quantity.

Other recent ideas for adaptive MLMC methods can for instance be found in [20, 3] and [15].

Our approach is based on the derivation of a guaranteed goal-oriented a posteriori error estimator which leads to exact error bounds in the deterministic setting, i.e. the error of the solution for a single realisation of the stochastic data. The evaluation of this type of error estimator requires the solution of some adjoint problem to determine the spatial dependence with regard to the quantity of interest. Our specific construction only contains explicit constants which enables the favourable property that the computable error bounds are guaranteed path-wise. Since the overall MLMC error decomposes into a deterministic part (controlled by the spatial error estimator) and a stochastic part which depends on the variance of the data and the number of computed samples, we also suggest how to control the stochastic error of the goal and the error estimator in probability.

The practical application of the results is shown with two numerical examples. They specifically demonstrate the efficiency of the derived a posteriori error estimator. Moreover, we illustrate a significant computational efficiency gain when using a localised goal functional. In this common situation, the influence of the solution on the quantity of interest is restricted to its vicinity. Thus, the derived adaptive algorithm generates meshes which are predominantly refined in a small part of the domain and additionally in areas where the approximation quality is low due to reduced regularity of the solution. This problem-dependent approach can lead to substantial computational savings in comparison to uniformly refined MLMC methods which can make the adaptive MLMC approach highly efficient.

The outline of the paper is as follows. In Section 2, the model problem and data requirements are introduced. Moreover, the variational formulation and the discretisation with the first order primal and mixed finite element method (FEM) is described. In Section 3, we recall the classical Monte Carlo (MC) method and the more efficient Multilevel Monte Carlo (MLMC) algorithm. Section 4 describes the notion of error estimates for goal functionals. Guaranteed a posteriori lower and upper goal error bounds for the path-wise deterministic problem based on the solution of some dual problem is derived. Moreover, an approach to control the stochastic error with a prescribed probability is presented. In Section 5, an algorithm for the problem-dependent construction of an adapted hierarchy of meshes for MLMC is formulated and tested computationally. It is based on the previously derived a posteriori error estimation of the deterministic error of the goal which is then used for mesh refinement of the different discretisation levels. Benchmark examples illustrate the efficiency of the proposed adaptive MLMC method. In particular, we demonstrate numerically that the computed goal-oriented error bounds are guaranteed and efficient. Moreover, we compare the computational complexity of adaptivity for the classical and multi level MC.

The depicted experiments provide a clear justification for the proposed novel adaptive MLMC algorithm which can lead to a significant reduction of computational complexity by problem-dependent mesh hierarchies with guaranteed overall error control (in probability).

2. PROBLEM SETTING

Stochastic tools which enable to quantify uncertainty in subsurface flow simulations have become increasingly popular. In general, the structural composition of the media in the computational domain $D \subset \mathbb{R}^d$ ($d = 1, 2, 3$) is only known at a finite set of measurement points. However, a complete spatial field is required as model input for computational calculations which thus usually are subject to a significant degree of uncertainties. Recent progress in modelling and simulation techniques allows for an incorporation of probabilistic representations of model data.

The Darcy equation is the model equation for steady-state groundwater flow which is used in this work. We assume some simply connected polygonal Lipschitz domain D with given Dirichlet data u_D on the closed Dirichlet boundary $\Gamma_D \subseteq \partial D$ of positive surface measure, Neumann data u_N on the Neumann boundary $\Gamma_N := \partial D \setminus \Gamma_D$, source function

f and permeability coefficient κ . We then seek the solution u such that

$$\begin{aligned} -\operatorname{div}(\kappa \nabla u) &= f && \text{in } D, \\ u &= u_D && \text{on } \Gamma_D, \\ \nabla u \cdot \mathbf{n} &= u_N && \text{on } \Gamma_N. \end{aligned}$$

Assuming sufficiently smooth data with bounded coefficient

$$0 < \kappa_{\min} \leq \kappa \leq \kappa_{\max} < \infty,$$

ellipticity of the operator results in existence and uniqueness of the weak solution u .

In the stochastic setting we assume a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and the problem data consists of random fields κ, f, u_N and u_D . Thus, the solution u also is a random field. More details will be discussed in Section 2.1. For this paper only the coefficient κ is a random field and all other data is supposed to be deterministic. An extensions to consider other data as random fields should be possible but would make the analysis and implementation more complicated. In order to determine a random solution of the problem, we employ the Monte-Carlo approach which gained some popularity due to recent advances with regard to variance reduction techniques. The foundation of this classical method is the sampling of realisations of the equation at hand, i.e., a set of realisations of the data fields is drawn with which the equation is solved. Then, statistical moments of the solution are evaluated based on the sample set of solutions.

In the application of groundwater flow, lognormal random fields for the permeability tensor are popular since they exhibit properties observed in the real world such as high local oscillations on large amplitudes. The logarithm of such a random field is Gaussian. A Gaussian random field is completely defined by its first two moments, i.e., its mean $\bar{\kappa}$ and some covariance function $C : D^2 \rightarrow \mathbb{R}$.

2.1. Model problem and variational formulation. For the model problem

$$(2.1) \quad \begin{aligned} -\operatorname{div}(\kappa(\omega, x) \nabla u(\omega, x)) &= f(\omega, x) && \text{in } D \\ u(\omega, x) &= u_D && \text{on } \Gamma_D, \\ \nabla u(\omega, x) \cdot \mathbf{n} &= u_N && \text{on } \Gamma_N. \end{aligned}$$

we seek the solution u which solves (2.1) P -a.s., i.e., for almost all $\omega \in \Omega$. The differential operators are to be understood with respect to $x \in D$. For a fixed $\omega \in \Omega$, the variational form of problem (2.1) with solution $u \in H_D^1(D) := \{v \in H^1(D) \mid v|_{\partial D} = u_D\}$ reads

$$(2.2) \quad b_\omega(u(\omega, \cdot), v) = L_\omega(v) \quad \text{for all } v \in V$$

where $V := H_D^1(D) := \{v \in H^1(D) \mid v|_{\partial D} = 0 \text{ along } \Gamma_D\}$ and

$$\begin{aligned} b_\omega(u(\omega, \cdot), v) &:= \int_D \kappa(\omega, x) \nabla u(x) \cdot \nabla v(x) \, dx, \\ L_\omega(v) &:= \int_D f(\omega, x) v(x) \, dx + \int_{\Gamma_N} u_N(\omega, s) v(s) \, ds. \end{aligned}$$

We assume that $f(\omega, \cdot) \in L^2(D)$, $u_D(\omega, \cdot) \in H^{1/2}(D)$, $0 < \kappa_{\min}(\omega) \leq \kappa(\omega, x) \leq \kappa_{\max}(\omega) < \infty$ for all $x \in D$ and $\kappa(\omega, \cdot) \in L^\infty(D)$. We define

$$(2.3) \quad \kappa_{\min}(\omega) := \min_{x \in D} \kappa(\omega, x) \quad \text{and} \quad \kappa_{\max} := \max_{x \in D} \kappa(\omega, x).$$

Then, (2.2) has a unique solution due to the (extended) Lax-Milgram Lemma, see e.g. [17, 12].

2.2. FEM discretisation and error. We assume a regular triangulation \mathcal{T} of the domain D which consists of triangles $T \in \mathcal{T}$, faces \mathcal{E} and nodes \mathcal{N} . Any two triangles are either disjoint or share a common face or up to two nodes. The set of faces along the Dirichlet boundary is denoted by $\mathcal{E}(\Gamma_D) := \{E \in \mathcal{E} \mid E \subseteq \Gamma_D\}$ and the set of faces along the Neumann boundary is denoted by $\mathcal{E}(\Gamma_N) := \{E \in \mathcal{E} \mid E \subseteq \overline{\Gamma_N}\}$ and we assume that every boundary face is in exactly one of these two sets. Unit normal vectors are denoted by \mathbf{n} and the diameter of an element $T \in \mathcal{T}$ is defined by $h_T := \text{diam}(T)$. For the discretisation of (2.2), we employ a conforming finite element space of piecewise polynomial order $k \geq 1$, i.e.,

$$V_h := \left\{ v \in H_D^1(D) \mid \forall T \in \mathcal{T} \ v|_T \in P_k(T) \right\} \subset H_D^1(D)$$

where $P_k(T)$ is the space of polynomials of total degree k on T . With the nodal interpolation $u_{D,h}(\omega, \cdot)$ of the Dirichlet data $u_D(\omega, \cdot) \in H^1(D) \cap H^2(\mathcal{E}(\Gamma_D))$, the discrete weak formulation of problem (2.1) reads: Find $u_h(\omega, \cdot) \in u_{D,h} + V_h$ such that

$$(2.4) \quad b_\omega(u_h(\omega, \cdot), v_h(\omega, \cdot)) = L_\omega(v_h(\omega, \cdot)) \quad \text{for all } v_h \in V_h.$$

For the following sections, we define the L^p and energy (semi-)norms by

$$\|v\|_{L^p(D)} := \left(\int_D |v|^p \, dx \right)^{1/p},$$

$$\|v\|_\omega := b_\omega(v, v)^{1/2} = \left(\int_D \kappa(\omega, \cdot) \nabla v \cdot \nabla v \, dx \right)^{1/2}.$$

The introduction of the stress $p := \kappa \nabla u$ in equation (2.1) formally leads to the mixed problem

$$p = \kappa \nabla u \quad \& \quad -\text{div } p = f$$

with appropriate boundary conditions. Note that Dirichlet boundary conditions transform to Neumann conditions and vice versa in the mixed formulation. The definitions

$$a_\omega((u, p), (\mu, \varrho)) := \int_D u \, \text{div}(\varrho) \, dx + \int_D \kappa^{-1} p \cdot \varrho \, dx + \int_D \text{div}(p) \mu \, dx,$$

$$L_\omega((\mu, \varrho)) := \int_{\Gamma_D} u_D \cdot \varrho \cdot \mathbf{n} \, ds - \int_D f \mu \, dx$$

give rise to the weak mixed formulation of the problem: Find $(u, p) \in V \times W := L^2(D) \times H(\text{div}, D)$ such that $p \cdot \mathbf{n} = u_N$ along Γ_N and

$$(2.5) \quad a_\omega((u, p), (\mu, \varrho)) = L_\omega((\mu, \varrho)) \quad \text{for all } (\mu, \varrho) \in V \times W \quad \mathcal{P}\text{-a.s.}$$

In a discrete setting, we seek $(u_{\text{RT}}, p_{\text{RT}}) \in V_h \times W_h$. Care has to be taken to choose admissible spaces in accordance to the classic theory for mixed problems, c.f. [9]. Here, we confine to a low-order discretisation with Raviart-Thomas elements defined on $T \in \mathcal{T}$ by

$$\text{RT}_0(T) := \left\{ v : T \rightarrow \mathbb{R}^d \mid v(x) = \alpha + \beta x, \alpha \in \mathbb{R}^d, \beta \in \mathbb{R} \right\}.$$

The space of broken piecewise RT_0 finite element functions is defined by

$$\text{RT}_{-1}(\mathcal{T}) := \left\{ v \in L^2(D)^d \mid \forall T \in \mathcal{T} : v|_T \in \text{RT}_0(T) \right\}.$$

With this the lowest order Raviart-Thomas space V_h is given by

$$\text{RT}_0(\mathcal{T}) := \{ v \in \text{RT}_{-1}(\mathcal{T}) \mid \forall E \in \mathcal{E} : v \cdot \mathbf{n}_E \text{ is continuous across } E \}.$$

An admissible choice for the discrete mixed space then is

$$\text{RT}_0(\mathcal{T}) \times P_0(\mathcal{T}) =: V_h \times W_h \subset V \times W$$

where in particular $\text{div } V_h = W_h$ holds. The Neumann boundary data enters through a Fortin interpolation, such that u_{RT} satisfies $\int_E u_{\text{RT}} \cdot \mathbf{n}_E ds = \int_E u_N ds$ for all $E \in \mathcal{E}(\Gamma_N)$.

In this paper we focus on the computation of some quantity of interest defined by the linear goal functional $Q(v) := \int_D gv dx$ with $g \in L^2(D)$. This introduces the dual problem of (2.2): Find $z \in V$ such that

$$(2.6) \quad b_\omega(v, z) = Q(v) \quad \text{for all } v \in H_0^1(D).$$

3. MULTILEVEL MONTE CARLO

This section recalls some well-known facts for the classical Monte Carlo (MC) and the multilevel Monte Carlo (MLMC) methods to introduce our notation and setting for the stochastic estimates, see also [7, 17, 32, 6] for further details.

3.1. Monte Carlo. One usually is interested in the evaluation of the expected value of some goal functional $Q(u)$ which here depends on the solution u of the model problem (2.1). Since the exact solution u is in general not available, an approximate solution u_h leads to the approximation $Q(u_h)$. Now, in order to compute an estimate of the quantity

$$\mathbf{E}[Q] := \int_\Omega \int_D g(u) dx d\mathcal{P}$$

with respect to some goal function g , the standard MC estimator can be employed, i.e., for a given set of samples $\omega^{(1)}, \dots, \omega^{(N)} \in \Omega$ and the respective sample solutions $u^{(1)}, \dots, u^{(N)}$ with $u^{(i)} := u(\omega_i)$ (and $u_h^{(i)} := u_h(\omega_i)$),

$$\mathbf{E}_N^{\text{MC}}[Q(u)] := N^{-1} \sum_{i=1}^N Q(u^{(i)}).$$

This estimator is known to be convergent and unbiased. In particular, it holds

$$\mathbf{E}[\mathbf{E}_N^{\text{MC}}[Q(u_h)]] = \mathbf{E}[Q(u_h)] \quad \text{and} \quad \mathbf{Var}[\mathbf{E}_N^{\text{MC}}[Q(u_h)]] = N^{-1} \mathbf{Var}[Q(u_h)]$$

where the variance for some random variable $X : \Omega \rightarrow \mathbb{R}$ is defined as $\mathbf{Var}[X] := \mathbf{E}[(X - \mathbf{E}[X])^2]$. The expected root mean square error (RMSE) of this estimator can be decomposed into two parts, namely

$$(3.1) \quad \begin{aligned} \text{err}(\mathbf{E}_N^{\text{MC}}[Q(u_h)])^2 &:= \mathbf{E} \left[\left(\mathbf{E}_N^{\text{MC}}[Q(u_h)] - \mathbf{E}[Q(u)] \right)^2 \right] \\ &= N^{-1} \mathbf{Var}[Q(u_h)] + (\mathbf{E}[Q(u_h)] - Q(u))^2. \end{aligned}$$

The first term on the right-hand side corresponds to the stochastic error, the second term is the expected value of the FEM discretisation error. Note that on the one hand

the stochastic error can be controlled by the amount of samples drawn in the course of the MC method. On the other hand, the FEM discretisation error depends on the resolution of the underlying mesh with regard to the regularity of the pathwise solutions.

To achieve $\text{err}(\mathbf{E}_N^{\text{MC}}[Q(u_h)]) \leq \varepsilon$ with $\varepsilon > 0$, both terms can, e.g., be required to be smaller than $\varepsilon/2$. Let $\mathbf{Var}[Q(u_h)]$ be a constant which is independent of the mesh size h . We then obtain $N \gtrsim \varepsilon^{-2}$ and $h \lesssim \varepsilon^{1/\beta}$. Here the (deterministic) convergence rate β in the sense of $|\mathbf{E}[Q(u_h - u)]| \lesssim h^\beta$ depends on the regularity of the solution u , see [12, 32] for a detailed analysis in this context.

Often the variance of the quantity has to be computed alongside the expected value. In the MC Method, this can be easily done with the estimator

$$\mathbf{Var}_N^{\text{MC}}[Q(u_h)] := \mathbf{E}_N^{\text{MC}} \left[\left(Q(u_h) - \mathbf{E}_N^{\text{MC}}[Q(u_h)] \right)^2 \right]$$

using the same samples as the estimator for the expectation. This comes at almost no additional cost.

3.2. Multilevel Monte Carlo. For the MLMC algorithm, we assume a set of (increasingly finer) triangulations \mathcal{T}_ℓ of the domain D for levels $\ell = 0, \dots, L$ with the coarsest mesh \mathcal{T}_0 . Accordingly, the level mesh sizes are denoted by h_ℓ and the discrete solutions by u_ℓ . To improve the efficiency of the classical MC algorithm, the key idea of the MLMC estimator is to circumvent the costly evaluation of $\mathbf{E}[Q(u_\ell)]$ exclusively on the finest level $\ell = L$. Instead, estimate corrections $\mathbf{E}[Y_\ell]$ between the levels are utilised where

$$Y_0 := Q(u_0) \quad \text{and} \quad Y_\ell := Q(u_\ell) - Q(u_{\ell-1}).$$

Note that, by linearity of the expectation,

$$\mathbf{E}[Q(u_L)] = \mathbf{E}[Q(u_0)] + \sum_{\ell=1}^L \mathbf{E}[Q(u_\ell) - Q(u_{\ell-1})] = \sum_{\ell=0}^L \mathbf{E}[Y_\ell].$$

For the unbiased estimator $\mathbf{E}_N^{\text{MC}}[Y_\ell]$ of $\mathbf{E}[Y_\ell]$ computed with N_ℓ samples on level ℓ , the ML estimator for $Q(u_h)$ reads

$$\mathbf{E}_N^{\text{ML}}[Q(u_L)] := \sum_{\ell=0}^L \mathbf{E}_{N_\ell}^{\text{MC}}[Y_\ell] = \sum_{\ell=0}^L \left(\frac{1}{N_\ell} \sum_{i=1}^{N_\ell} Y_\ell(\omega^{(i)}) \right)$$

with

$$\mathbf{E}_N^{\text{MC}}[Y_0] := \mathbf{E}_N^{\text{MC}}[Q(u_0)] \quad \text{and} \quad \mathbf{E}_N^{\text{MC}}[Y_\ell] := \frac{1}{N_\ell} \sum_{\ell=1}^L (Q(u_\ell) - Q(u_{\ell-1})) \quad \text{for } \ell \geq 1.$$

It is crucial that the same sample $\omega^{(i)}$ is used on the two consecutive levels in $Y_\ell(\omega^{(i)}) = Q(u_\ell(\omega^{(i)})) - Q(u_{\ell-1}(\omega^{(i)}))$. Since all expectations $\mathbf{E}[Y_\ell]$ are estimated independently, it again holds

$$\mathbf{E}[\mathbf{E}_N^{\text{ML}}[Q(u_h)]] = \mathbf{E}[Q(u_h)] \quad \text{and} \quad \mathbf{Var}[\mathbf{E}_N^{\text{ML}}[Q(u_h)]] = \sum_{\ell=0}^L N_\ell^{-1} \mathbf{Var}[Y_\ell].$$

The mean square error can be expanded as before, namely

$$(3.2) \quad \begin{aligned} \text{err}(\mathbf{E}_N^{\text{ML}}[Q(u_L)])^2 &:= \mathbf{E} \left[\left(\mathbf{E}_N^{\text{ML}}[Q(u_L)] - \mathbf{E}[Q(u)] \right)^2 \right] \\ &= \sum_{\ell=0}^L N_\ell^{-1} \mathbf{Var}[Y_\ell] + (\mathbf{E}[Q(u_L)] - Q(u))^2. \end{aligned}$$

To bound err by ε , one can choose $h_L \lesssim \varepsilon^{1/\beta}$ for the second term on the right-hand side as above. Note that for the first term (which again should be bounded by $\varepsilon/2$) it holds $\mathbf{Var}[Y_\ell] \rightarrow 0$ and $N_\ell \rightarrow 1$ for $h \rightarrow 0$ and $\ell > 0$ (or $\ell \rightarrow \infty$) if $Q(u_h) \rightarrow Q(u)$ in mean square.

The coarsest admissible h_0 depends on the regularity of the solution u and thus on the regularity of the covariance function of the coefficient and its correlation length λ .

To determine (optimal) values for the number of levels L and the sample numbers $\{N_\ell\}_{\ell=0}^L$, we assume a monotone error decay of the form

$$|\mathbf{E}[Q(u_L)] - Q(u)| \simeq h^\beta \quad \text{for } h \leq h^*.$$

Then, $\mathbf{E}_N^{\text{MC}}[Y_\ell] \simeq h^\beta$ for N_ℓ sufficiently large. Denote by $C_\ell := C(Y_\ell^{(i)})$ the computational cost of a single sample Y_ℓ evaluation on level ℓ . The total MLMC cost is then given by

$$(3.3) \quad C(\mathbf{E}_N^{\text{ML}}[Q(u_L)]) = \sum_{\ell=0}^L N_\ell C_\ell.$$

To minimise the variance of the MLMC estimator for a fixed computational cost, we set

$$(3.4) \quad N_\ell \simeq \sqrt{\mathbf{Var}[Y_\ell] / C_\ell}$$

with the constant of proportionality chosen such that the total variance is $\varepsilon^2/2$. The total cost on level ℓ is then proportional to $\sqrt{\mathbf{Var}[Y_\ell] C_\ell}$ and thus

$$C(\mathbf{E}_N^{\text{ML}}[Q(u_L)]) \lesssim \sum_{\ell=0}^L \sqrt{\mathbf{Var}[Y_\ell] C_\ell}.$$

In case that the data of the PDE is sufficiently smooth, the convergence for the error of the expectation of the estimator can be shown, see [7] for detailed assumptions and the proof, and [12] for generalisations of the cited result.

Lemma 3.1. *For the error of the expectation of the MLMC-FE approximation $u_L \in L^2(\Omega; u_D + H_D^1(D))$, it holds*

$$\left\| \mathbf{E}[u] - \mathbf{E}_N^{\text{ML}}[u_L] \right\|_{L^2(\Omega; u_D + H_D^1(D))} \leq C \left(h_L + \sum_{\ell=1}^L h_\ell N_\ell^{-1/2} \right) \left(\|f\|_{L^2(D)} + \|u_N\|_{H^{1/2}(\Gamma_N)} \right).$$

In contrast to the MC context, the variance cannot be computed alongside the expectation when using the MLMC method. Nevertheless, with $W_0 = Q(u_0)^2$ and $W_\ell = Q(u_\ell)^2 -$

$Q(u_{\ell-1})^2$ it holds the decomposition of the variance

$$\begin{aligned}\mathbf{Var}[Q(u_L)] &= \mathbf{E}[Q(u_L)^2] - \mathbf{E}[Q(u_L)]^2 \\ &= \sum_{\ell=0}^L \mathbf{E}[W_\ell] - \left(\sum_{\ell=0}^L \mathbf{E}[Y_\ell] \right)^2 \\ &= \sum_{\ell=0}^L \mathbf{E}[W_\ell] - \sum_{\ell=0}^L \mathbf{E}[Y_\ell]^2 - 2 \sum_{\substack{\ell, k=0 \\ \ell < k}}^L \mathbf{E}[Y_\ell] \mathbf{E}[Y_k].\end{aligned}$$

In consequence, define the multilevel estimator for the variance as

$$\mathbf{Var}_N^{\text{ML}}[Q(u_L)] := \sum_{\ell=0}^L \mathbf{E}_{N_\ell}^{\text{MC}}[W_\ell] - \sum_{\ell=0}^L \mathbf{E}_{N_\ell}^{\text{MC}}[Y_\ell]^2 - 2 \sum_{\substack{\ell, k=0 \\ \ell < k}}^L \mathbf{E}_{N_\ell}^{\text{MC}}[Y_\ell] \mathbf{E}_{N_\ell}^{\text{MC}}[Y_k].$$

Yet again, the additional computational cost is negligible. All terms are already known except for the W_ℓ ($\ell = 0, \dots, L$) which can be cheaply sampled alongside the Y_ℓ . The additional memory needed for storage of these values is minimal. Similar decompositions are possible for higher order central moments and require additional terms as it holds

$$\mathbf{E}[(Q(u_L) - \mathbf{E}[Q(u_L)])^n] = \sum_{k=0}^n \binom{n}{k} \mathbf{E}[Q(u_L)^{n-k}] \mathbf{E}[Q(u_L)]^k.$$

4. ESTIMATION OF GOAL FUNCTIONALS

The error in the computational approximation of the quantity of interest is given by $Q(u) - Q(u_h)$. We subsequently assume a fixed $\omega \in \Omega$ to derive guaranteed upper and lower bounds η^\pm for the deterministic error of this sample. In Section 4.2 these deterministic bounds are used to compute probabilistic bounds. The following naming convention is used for the bounds with $\alpha \neq 0$

$$u^\pm := \alpha u \pm \frac{z}{\alpha}, \quad p^\pm = \alpha p \pm \frac{q}{\alpha},$$

where u and z are the solutions of Problem (2.4) and Problem (2.6), respectively, whereas p and q are the corresponding mixed solutions. Due to the linearity and symmetry of the solution operator, these are the solutions for the primal problem with the right-hand side $f^\pm := \alpha f + g/\alpha$. The same naming convention is used for their discrete counterparts.

The oscillations for a function f on some mesh \mathcal{T} are defined by $\text{osc}(f, \mathcal{T})^2 := \sum_{T \in \mathcal{T}} \text{osc}(f, T)^2$ where

$$\text{osc}(f, T) := h_T \|f - f_T\|_{L^2(T)} \quad \text{with} \quad f_T := |T|^{-1} \int_T f \, dx.$$

Analogously, the oscillations for a function u_N along the Neumann boundary Γ_N are defined by $\text{osc}(u_N, \mathcal{E}(\Gamma_N))^2 := \sum_{E \in \mathcal{E}(\Gamma_N)} \text{osc}(u_N, E)^2$ where

$$\text{osc}(u_N, E) := h_E^{1/2} \|u_N - (u_N)_E\|_{L^2(E)} \quad \text{with} \quad (u_N)_E := |E|^{-1} \int_E u_N \, dx.$$

The functions $f_{\mathcal{T}}$ and $(u_N)_{\mathcal{E}(\Gamma_N)}$ are defined piecewise for each $T \in \mathcal{T}$ and $E \in \mathcal{E}(\Gamma_N)$ by $f_{\mathcal{T}}|_T := f_T$ and $(u_N)_{\mathcal{E}(\Gamma_N)}|_E := (u_N)_E$, respectively.

4.1. Guaranteed error bounds in the deterministic case. The derivation of guaranteed deterministic bounds follows the spirit of [11] and [26]. The residual $\text{Res} \in (H_0^1(D))^*$ for some right-hand side f and the associated discrete solution u_h with discrete stress $p_h := \kappa \nabla u_h$ reads, for all $v \in V$,

$$(4.1) \quad \text{Res}(v) := b_\omega(u - u_h, v) = \int_D f v \, dx + \int_{\Gamma_N} u_N v \, ds - \int_D p_h \cdot \nabla v \, dx.$$

The naming conventions from above also apply for Res by substitution of u, u_h and f , i.e., $\text{Res}^\pm(v) := b_\omega(u^\pm - u_h^\pm, v)$. The dual norm of a residual Res is given by

$$\|\text{Res}\|_* = \sup_{v \in V} \frac{|\text{Res}(v)|}{\|v\|_\omega}.$$

Remark 4.1. For homogeneous Dirichlet data $u_D = 0$, Hilbert space theory shows the identity $\|\text{Res}\|_* = \|u - u_h\|_\omega$. In case of inhomogeneous Dirichlet boundary data $u_D \in H^1(D) \cap H^2(\mathcal{E}(\Gamma_D))$ along the Dirichlet boundary faces $\mathcal{E}(\Gamma_D)$, it holds

$$\|u - u_h\|_\omega^2 := \|\text{Res}\|_*^2 + \inf_{\substack{w \in H^1(D) \\ w = u - u_h \text{ along } \Gamma_D}} \|w\|_\omega^2.$$

The special choice of w from [5] leads to the higher-order upper bound

$$0 \leq \inf_{\substack{w \in H^1(D) \\ w = u - u_h \text{ along } \Gamma_D}} \|w\|_\omega \leq C \left\| h^{3/2} \kappa_{\max, T} \partial^2(u_D - u_h) / \partial s^2 \right\|_{L^2(\Gamma_D)} =: \text{h.o.t.}(u_D).$$

The constant C only depends on the shape of the triangles but not on their size and is set to 1 in the numerical experiments below. For triangulations into right isosceles triangles [26] claims $C \leq 0.4980$.

The point of departure for the derivation of upper and lower bounds for the quantity of interest Q is the parallelogram identity

$$(4.2) \quad \begin{aligned} Q(u - u_h) &= \frac{1}{4} \|u^+ - u_h^+\|_*^2 - \frac{1}{4} \|u^- - u_h^-\|_*^2 \\ &= \frac{1}{4} \|\text{Res}^+\|_*^2 + \inf_{\substack{w^+ \in H^1(D) \\ w^+ = u^+ - u_h^+ \text{ along } \Gamma_D}} \|w^+\|_\omega^2 \\ &\quad - \frac{1}{4} \|\text{Res}^-\|_*^2 - \inf_{\substack{w^- \in H^1(D) \\ w^- = u^- - u_h^- \text{ along } \Gamma_D}} \|w^-\|_\omega^2. \end{aligned}$$

This follows from basic algebra, the Galerkin orthogonality of the discrete solutions, and the preceding definitions, also see [2, 30, 11]. All following theorems apply to any of these residuals $\text{Res} \in \{\text{Res}^+, \text{Res}^-\}$. In the sequel $C_P(T) := \sup\{h_T^{-1} \|v - v_T\|_{L^2(T)} \mid v \in H^1(T), \|v\|_\omega = 1\}$ denotes the Poincaré constant on an element $T \in \mathcal{T}$. The piecewise constant $C_P(\mathcal{T})$ is defined by $C_P(\mathcal{T})|_T := C_P(T)$.

Remark 4.2. Since T is convex, [29, 8] show $C_P(T) := 1/\pi$. In 2D, [24] proved the better constant $C_P(T) := 1/j_{1,1}$ where $j_{1,1} \approx 3.8317$ is the first positive root of the first Bessel function $J_1(x) := \frac{1}{\pi} \int_0^\pi \cos(\varphi - x \sin \varphi) \, d\varphi$.

Theorem 4.3. For every r with $\operatorname{div} r + f_{\mathcal{T}} = 0$ and $r \cdot \mathbf{n} = (u_N)_{\mathcal{E}(\Gamma_N)}$ along Γ_N , the quantity

$$\begin{aligned} \gamma(r)^2 := & \sum_{T \in \mathcal{T}} \left(\left\| \kappa^{-1/2}(p_h - r) \right\|_{L^2(T)} + C_P(T) h_T \kappa_{\min, T}^{-1/2} \|f - f_T\|_{L^2(T)} \right. \\ & \left. + C(E) \kappa_{\min, T}^{-1/2} \sum_{E \in \mathcal{E}(\Gamma_N) \cap \mathcal{E}(T)} \|u_N - (u_N)_E\|_{L^2(E)} \right)^2 \end{aligned}$$

satisfies

$$\|\operatorname{Res}\|_* \leq \gamma(r).$$

The constants $C_P(T)$ and $C(E)$ depend only on the shape of the elements in \mathcal{T} and $\mathcal{E}(\Gamma_N)$ but not on their size.

Proof. For any $r \in H(\operatorname{div}, D)$ with $\operatorname{div} r + f_{\mathcal{T}} = 0$ and $r \cdot \mathbf{n} = (u_N)_{\mathcal{E}(\Gamma_N)}$ along Γ_N , an integration by parts yields, for any $v \in H_0^1(D)$,

$$\operatorname{Res}(v) = \int_D (f + \operatorname{div} r) v \, dx + \int_{\Gamma_N} (u_N - r \cdot \mathbf{n}) v \, ds - \int_D (p_h - r) \cdot \nabla v \, dx.$$

The last integral is bounded by

$$- \int_D (p_h - r) \cdot \nabla v \, dx \leq \left\| \kappa^{-1/2}(p_h - r) \right\|_{L^2(D)} \|v\|_{\omega}.$$

The orthogonality $\int_D (f + \operatorname{div} r) v_T \, dx = 0$ and elementwise Poincaré and Cauchy inequalities bound the first integral by

$$\begin{aligned} \int_D (f + \operatorname{div} r) v \, dx &= \sum_{T \in \mathcal{T}} \int_T (f - f_T)(v - v_T) \, dx \\ &= \sum_{T \in \mathcal{T}} \int_T \kappa_{\min, T}^{-1/2} (f - f_T) \kappa_{\min, T}^{1/2} (v - v_T) \, dx \\ &\leq \sum_{T \in \mathcal{T}} \left\| \kappa_{\min, T}^{-1/2} (f - f_T) \right\|_{L^2(T)} \kappa_{\min, T}^{1/2} \|v - v_T\|_{L^2(T)} \\ (4.3) \quad &\leq \sum_{T \in \mathcal{T}} \left\| \kappa_{\min, T}^{-1/2} (f - f_T) \right\|_{L^2(T)} C_P(\mathcal{T}) h_T \left\| \kappa^{1/2} \nabla v \right\|_{L^2(T)}. \end{aligned}$$

The estimation of the second integral is similar and starts with the orthogonality $\int_{\Gamma_N} (u_N - r \cdot \mathbf{n}) v_E \, ds = 0$ and a piecewise Cauchy inequality

$$\begin{aligned} \int_{\Gamma_N} (u_N - r \cdot \mathbf{n}) v \, ds &= \sum_{E \in \mathcal{E}(\Gamma_N)} \int_E (u_N - (u_N)_E)(v - v_E) \, ds \\ &\leq \sum_{E \in \mathcal{E}(\Gamma_N)} \left\| \kappa_{\min, T_E}^{-1/2} h_E^{1/2} (u_N - (u_N)_E) \right\|_{L^2(E)} h_E^{-1/2} \left\| \kappa_{\min, T_E}^{1/2} (v - v_E) \right\|_{L^2(E)}. \end{aligned}$$

A piecewise trace inequality for every $E \in \mathcal{E}(\Gamma_N)$ and its neighbouring element $T_E \in \mathcal{T}$, s.t. $E \subset \partial T$ shows

$$h_E^{-1/2} \left\| \kappa_{\min, T_E}^{1/2} (v - v_E) \right\|_{L^2(E)} \leq C(E) \left\| \kappa^{1/2} \nabla (v - v_E) \right\|_{L^2(T_E)}.$$

The constant $C(E)$ is independent of h_E and depends only on the shape of T_E , see e.g. [26] for details and an explicit upper bound of $C(E)$. A Cauchy inequality in $\mathbb{R}^{|\mathcal{T}|}$ concludes the proof. \square

Remark 4.4 (Lower Bounds). Lower bounds for $\|\text{Res}\|_*$ are given by any test function $v \in V$ by

$$(4.4) \quad \|\text{Res}\|_* \geq \xi(v) := |\text{Res}(v)| / \|v\|.$$

Here we employ a design similar to [1, 10]. The original design incorporates the Crouzeix-Raviart solution \hat{u}_{CR} of the corresponding Poisson problem with right-hand side $f_{\mathcal{T}}$ and computes the piecewise quadratic function

$$(4.5) \quad v^0 := \hat{u}_{CR} - \kappa_{\mathcal{T}}^{-1} f_{\mathcal{T}} \psi / 2 \in P_2(\mathcal{T}),$$

with $\psi(x) := |x - \text{mid}(T)|^2 / 2 - \frac{1}{|T|} \int_T |y - \text{mid}(T)|^2 dy$ for $x \in T \in \mathcal{T}$. According to [1], $\kappa_{\mathcal{T}} \nabla v^0 = \hat{p}_{RT}$ equals the solution \hat{p}_{RT} of the mixed problem with κ replaced by $\kappa_{\mathcal{T}}$. Furthermore, it holds for all $T \in \mathcal{T}$

$$v^0|_T := \underset{v \in P_2(T)}{\text{argmin}} \left\{ \|\kappa_{\mathcal{T}} \nabla v - \hat{p}_{RT}\|_{L^2(T)} \mid \int_T v dx / |T| = \hat{p}_{RT}(\text{mid}(T)) \right\}.$$

This means that there are two ways to compute v^0 which involve the computation of either \hat{u}_{CR} or \hat{p}_{RT} (which is different from p_{RT} since κ is not piecewise constant in our applications). We use (4.5) and compute some nearby $v_A \in P_2(\mathcal{T}) \cap C_0(D)$ which ensures $v_A = u_h$ along Γ_D . Eventually, the test function $v = v_A - u_h \in V$ in (4.4) leads to the lower bound

$$(4.6) \quad \|\text{Res}\|_* \geq \xi(v_A - u_h) = |\text{Res}(v_A - u_h)| / \|v_A - u_h\|.$$

Proposition 4.5. For any r^{\pm} with $\text{div } r^{\pm} + f_{\mathcal{T}}^{\pm} = 0$ and $r \cdot \mathbf{n} = u_{N\mathcal{E}(\Gamma_N)}$ along Γ_N , $v^{\pm} \in V$ and $\alpha \neq 0$, it holds

$$\eta^-(\alpha, r^+, r^-, v^+) \leq Q(u - u_h) \leq \eta^+(\alpha, r^+, r^-, v^-)$$

where η^- and η^+ are defined as

$$\eta^- := \frac{1}{4} \xi^+(v^+) - \frac{1}{4} \gamma^-(r^-)^2, \quad \eta^+ := -\frac{1}{4} \xi^-(v^-) + \frac{1}{4} \gamma^+(r^+)^2.$$

Proof. Application of Theorem 4.3 and Remark 4.4 to Equation (4.2) gives the result. Here, γ^+ is the upper bound of $\|\text{Res}^+\|_*$ and γ^- is the upper bound of $\|\text{Res}^-\|_*$. \square

In the numerical examples below, we employ $r^{\pm} = p_{RT}^{\pm}$ and $v^{\pm} = v_A^{\pm} - u_h^{\pm}$ from Remark 4.4. A proper choice of α is paramount for the efficiency of the bounds. Similar to [11], it is $\alpha = \left\| \kappa^{-1/2} (q_h - q_{RT}) \right\|_{L^2(D)}^{1/2} / \left\| \kappa^{-1/2} (p_h - p_{RT}) \right\|_{L^2(D)}^{1/2}$ optimal up to oscillations in input data.

4.2. Error bounds in probability for the stochastic case. Let $\mathcal{N}_{0,1}$ denote the standard normal distribution. With some stochastic variable X and its standard deviation σ_X the central limit theorem by Lindeberg and Lévy (see eg. [16, 21]) states the limit of the error in the Monte Carlo estimator as

$$\sigma_X^{-1} N^{1/2} (\mathbf{E}[X] - \mathbf{E}_N^{\text{MC}}[X]) \rightarrow \mathcal{N}_{0,1} \quad \text{for } N \rightarrow \infty.$$

The cumulative distribution function of $\mathcal{N}_{0,1}$ is $\Phi(x) := (2\pi)^{-1/2} \int_{-\infty}^x \exp(-t^2/2) dt$. It thus holds the limit of the following probability for any $s > 0$

$$\mathbf{p} := \mathbf{P}\left[\left(\mathbf{E}[X] - \mathbf{E}_N^{\text{MC}}[X]\right) < \sigma_X N^{-1/2} s\right] \rightarrow \Phi(s) \quad \text{for } N \rightarrow \infty.$$

For $\varepsilon_X(N) := \sigma_X N^{-1/2} s$, there exists some sequence $(\varepsilon_X^*(N, \mathbf{p}))_{N \in \mathbb{N}}$ such that $\varepsilon_X^*(N, \mathbf{p}) \rightarrow 0$ for $N \rightarrow \infty$ and

$$(4.7) \quad \varepsilon_X(N) = \sigma_X N^{-1/2} \Phi^{-1}(\mathbf{p}) + \varepsilon_X^*(N, \mathbf{p}).$$

This equation can now be used to control the stochastic error of the Monte Carlo estimator $\mathbf{E}_N^{\text{MC}}[X]$.

In the following theorem we assume that the events

$$A^\pm := \left(\mathbf{E}[\eta^\pm] - \mathbf{E}_N^{\text{MC}}[\eta^\pm]\right) < \mathbf{Var}[\eta^\pm]^{1/2} N^{-1/2} \Phi^{-1}(\mathbf{p}^\pm) + \varepsilon_{\eta^\pm}^*(N, \mathbf{p}^\pm),$$

$$B := \left(\mathbf{E}[Q(u_h)] - \mathbf{E}_N^{\text{MC}}[Q(u_h)]\right) < \mathbf{Var}[Q(u_h)]^{1/2} N^{-1/2} \Phi^{-1}(\mathbf{p}^*) + \varepsilon_{Q(u_h)}^*(N, \mathbf{p}^*)$$

hold with the prescribed probabilities $\mathbf{P}[A^\pm] \geq \mathbf{p}^\pm$ and $\mathbf{P}[B] \geq \mathbf{p}^*$ (for a sufficiently large number of samples such that $\varepsilon_{\eta^\pm}^*(N, \mathbf{p}^\pm)$ and $\varepsilon_{Q(u_h)}^*(N, \mathbf{p}^*)$ are negligible).

Corollary 4.6. *Let $\mathbf{p}, \mathbf{p}^*, \mathbf{p}^-, \mathbf{p}^+ \in (0, 1)$ be some prescribed probabilities with $\mathbf{p} = 2\mathbf{p}^* + \mathbf{p}^+ + \mathbf{p}^- - 3$. Then, the bounds*

$$\begin{aligned} \eta^\ominus(\mathbf{p}^*, \mathbf{p}^-, N) &:= \mathbf{E}_N^{\text{MC}}[\eta^-] - \mathbf{Var}[\eta^-]^{1/2} N^{-1/2} \Phi^{-1}(\mathbf{p}^-) - \varepsilon_{\eta^-}^*(N, \mathbf{p}^-) \\ &\quad - \mathbf{Var}[Q(u_h)]^{1/2} N^{-1/2} \Phi^{-1}(\mathbf{p}^*) - \varepsilon_{Q(u_h)}^*(N, \mathbf{p}^*), \end{aligned}$$

$$\begin{aligned} \eta^\oplus(\mathbf{p}^*, \mathbf{p}^+, N) &:= \mathbf{E}_N^{\text{MC}}[\eta^+] + \mathbf{Var}[\eta^+]^{1/2} N^{-1/2} \Phi^{-1}(\mathbf{p}^+) + \varepsilon_{\eta^+}^*(N, \mathbf{p}^+) \\ &\quad + \mathbf{Var}[Q(u_h)]^{1/2} N^{-1/2} \Phi^{-1}(\mathbf{p}^*) + \varepsilon_{Q(u_h)}^*(N, \mathbf{p}^*) \end{aligned}$$

satisfy

$$\mathbf{P}\left[\eta^\ominus(\mathbf{p}^*, \mathbf{p}^-, N) \leq \mathbf{E}[Q(u)] - \mathbf{E}_N^{\text{MC}}[Q(u_h)] \leq \eta^\oplus(\mathbf{p}^*, \mathbf{p}^+, N)\right] \geq \mathbf{p}.$$

Proof. Proposition 4.5 leads to the inequality

$$\begin{aligned} &\mathbf{E}[Q(u)] - \mathbf{E}_N^{\text{MC}}[Q(u_h)] \\ &\leq \mathbf{E}[\eta^+] + \mathbf{E}[Q(u_h)] - \mathbf{E}_N^{\text{MC}}[Q(u_h)] \\ &= \mathbf{E}_N^{\text{MC}}[\eta^+] + (\mathbf{E}[\eta^+] - \mathbf{E}_N^{\text{MC}}[\eta^+]) + (\mathbf{E}[Q(u_h)] - \mathbf{E}_N^{\text{MC}}[Q(u_h)]). \end{aligned}$$

The terms in brackets define stochastic errors of the Monte Carlo estimators for the stochastic variables $X = \eta^+$ and $X = Q(u_h)$ and are related to the events A^+ and B . The conditional probability for both events A^+ and B to hold can be estimated by

$$\mathbf{P}[A^+ \cap B] = \mathbf{P}[A^+] + \mathbf{P}[B] - \mathbf{P}[A^+ \cup B] \geq \mathbf{P}[A^+] + \mathbf{P}[B] - 1 \geq \mathbf{p}^* + \mathbf{p}^+ - 1.$$

This yields the probabilistic upper bound

$$\mathbf{P}\left[\mathbf{E}[Q(u)] - \mathbf{E}_N^{\text{MC}}[Q(u_h)] \leq \eta^\oplus(\mathbf{p}^*, \mathbf{p}^+, N)\right] \geq \mathbf{p}^* + \mathbf{p}^+ - 1.$$

The lower bound

$$\mathbf{P}\left[\eta^\ominus(\mathbf{p}^*, \mathbf{p}^-, N) \leq \mathbf{E}[Q(u)] - \mathbf{E}_N^{\text{MC}}[Q(u_h)]\right] \geq \mathbf{p}^* + \mathbf{p}^- - 1$$

can be derived similarly as $\mathcal{N}_{0,1}$ is even and its cumulative distribution function is odd. Another application of the formula for conditional probabilities (for the probability that both bounds hold) yields the assertion and concludes the proof. \square

In the case of the multilevel Monte Carlo approach consider the bounds on each level $\eta_\ell^- \leq Q(u - u_\ell) \leq \eta_\ell^+$ and consequently $Z_0^+ := \eta_0^+$ as well as $Z_\ell^\pm := \eta_\ell^\pm - \eta_{\ell-1}^\pm$ for $\ell = 1, \dots, L$. Furthermore, the events

$$A_\ell^\pm := (\mathbf{E}[Z_\ell^\pm] - \mathbf{E}_{N_\ell}^{\text{MC}}[Z_\ell^\pm]) < \mathbf{Var}[Z_\ell^\pm]^{1/2} N_\ell^{-1/2} \Phi^{-1}(\mathbf{p}_\ell^\pm) + \varepsilon_{Z_\ell^\pm}^*(N_\ell, \mathbf{p}_\ell^\pm),$$

$$B_\ell := (\mathbf{E}[Y_\ell] - \mathbf{E}_{N_\ell}^{\text{MC}}[Y_\ell]) < \mathbf{Var}[Y_\ell]^{1/2} N_\ell^{-1/2} \Phi^{-1}(\mathbf{p}_\ell^*) + \varepsilon_{Y_\ell}^*(N_\ell, \mathbf{p}_\ell^*)$$

on each level $\ell \in L$ are assumed to hold with the probabilities $\mathbf{P}[A_\ell^\pm] \geq \mathbf{p}_\ell^\pm$ and $\mathbf{P}[B_\ell] \geq \mathbf{p}_\ell^*$.

Corollary 4.7. *Let $\mathbf{p} \in (0, 1)$, $\mathbf{p}^*, \mathbf{p}^-, \mathbf{p}^+ \in (0, 1)^L$ be some prescribed probabilities with $\mathbf{p} = 2 \sum_{\ell=0}^L \mathbf{p}_\ell^* + \sum_{\ell=0}^L \mathbf{p}_\ell^- + \sum_{\ell=0}^L \mathbf{p}_\ell^+ - 4L - 3$. Then, the bounds*

$$\eta_{\text{ML}}^\ominus(\mathbf{p}^*, \mathbf{p}^-, \mathbf{N}) := \mathbf{E}_N^{\text{ML}}[\eta_L^-] - \sum_{\ell=0}^L \mathbf{Var}[Z_\ell^-]^{1/2} N_\ell^{-1/2} \Phi^{-1}(\mathbf{p}_\ell^-) - \sum_{\ell=0}^L \varepsilon_{Z_\ell^-}^*(N_\ell, \mathbf{p}_\ell^-)$$

$$- \sum_{\ell=0}^L \mathbf{Var}[Y_\ell]^{1/2} N_\ell^{-1/2} \Phi^{-1}(\mathbf{p}_\ell^*) - \sum_{\ell=0}^L \varepsilon_{Y_\ell}^*(N_\ell, \mathbf{p}_\ell^*),$$

$$\eta_{\text{ML}}^\oplus(\mathbf{p}^*, \mathbf{p}^+, \mathbf{N}) := \mathbf{E}_N^{\text{ML}}[\eta_L^+] + \sum_{\ell=0}^L \mathbf{Var}[Z_\ell^+]^{1/2} N_\ell^{-1/2} \Phi^{-1}(\mathbf{p}_\ell^+) + \sum_{\ell=0}^L \varepsilon_{Z_\ell^+}^*(N_\ell, \mathbf{p}_\ell^+)$$

$$+ \sum_{\ell=0}^L \mathbf{Var}[Y_\ell]^{1/2} N_\ell^{-1/2} \Phi^{-1}(\mathbf{p}_\ell^*) + \sum_{\ell=0}^L \varepsilon_{Y_\ell}^*(N_\ell, \mathbf{p}_\ell^*)$$

satisfy

$$\mathbf{P}\left[\eta_{\text{ML}}^\ominus(\mathbf{p}^*, \mathbf{p}^-, \mathbf{N}) \leq \mathbf{E}[Q(u)] - \mathbf{E}_N^{\text{ML}}[Q(u_L)] \leq \eta_{\text{ML}}^\oplus(\mathbf{p}^*, \mathbf{p}^+, \mathbf{N})\right] \geq \mathbf{p}.$$

Proof. Similarly to Corollary 4.6, Equation (4.7) and Proposition 4.5 and the notation of Y_ℓ and Z_ℓ^\pm from above show the inequality

$$\begin{aligned} & \mathbf{E}[Q(u)] - \mathbf{E}_N^{\text{ML}}[Q(u_L)] \\ &= \mathbf{E}[Q(u)] - \sum_{\ell=0}^L \mathbf{E}_{N_\ell}^{\text{MC}}[Y_\ell] \\ &= \mathbf{E}[Q(u)] - \mathbf{E}[Q(u_L)] + \sum_{\ell=0}^L (\mathbf{E}[Y_\ell] - \mathbf{E}_{N_\ell}^{\text{MC}}[Y_\ell]) \\ &\leq \mathbf{E}[\eta_L^+] + \sum_{\ell=0}^L (\mathbf{E}[Y_\ell] - \mathbf{E}_{N_\ell}^{\text{MC}}[Y_\ell]) \\ &= \sum_{\ell=0}^L \mathbf{E}_{N_\ell}^{\text{MC}}[Z_\ell^+] + \sum_{\ell=0}^L (\mathbf{E}[Z_\ell^+] - \mathbf{E}_{N_\ell}^{\text{MC}}[Z_\ell^+]) + \sum_{\ell=0}^L (\mathbf{E}[Y_\ell] - \mathbf{E}_{N_\ell}^{\text{MC}}[Y_\ell]). \end{aligned}$$

The probability of the events A_ℓ^\pm and B_ℓ from above allow to bound the probability that the upper bound $\mathbf{E}[Q(u)] - \mathbf{E}_N^{\text{ML}}[Q(u_L)] \leq \eta_{\text{ML}}^\oplus(\mathbf{p}^*, \mathbf{p}^+, \mathbf{N})$ holds by $\mathbf{P}\left[\bigcap_{\ell=0}^L (A_\ell^+ \cap B_\ell)\right]$. Commulative application of the formula for conditional probabilities that was already used in Corollary 4.6 gives

$$\begin{aligned} \mathbf{P}\left[\mathbf{E}[Q(u)] - \mathbf{E}_N^{\text{ML}}[Q(u_L)] \leq \eta_{\text{ML}}^\oplus(\mathbf{p}^*, \mathbf{p}^+, \mathbf{N})\right] &\geq \sum_{\ell=0}^L \mathbf{P}[A_\ell^+ \cap B_\ell] - L \\ &\geq \sum_{\ell=0}^L \mathbf{P}[A_\ell^+] + \mathbf{P}[B_\ell] - 2L - 1 \\ &\geq \sum_{\ell=0}^L p_\ell^* + p_\ell^+ - 2L - 1. \end{aligned}$$

The probability for the lower bound is derived analogously and another application of the formula for conditional probabilities concludes the proof. \square

Remark 4.8. The sharpest bounds in Corollary 4.6 and 4.7 are given for an optimal choice of p^*, p^-, p^+ and $\mathbf{p}^*, \mathbf{p}^-, \mathbf{p}^+$ respectively and thus it remains to find the minimum of the bounds under the given constraints. This has to be done numerically (e.g. as in [22]), as the inverse cumulative distribution function cannot be expressed in terms of elementary functions. As the variances can differ greatly in size by orders of magnitude it might be prudent in practical applications to choose the probabilities evenly in order to prevent blow ups of errors in the approximation of the variances.

5. EXPERIMENTS

This section describes our adaptive mesh refinement algorithm and reports on two numerical experiments.

5.1. Adaptive mesh refinement. Quantities of interest with local support lead to slow convergence rates with uniformly refined meshes. Furthermore, classic FEM adaptivity might lead to suboptimal convergence or even no convergence at all in the goal quantity despite improved convergence rates for the energy norm of u_h in the primal problem [28]. In the following, we aim to adapt the approach for goal-driven adaptivity from [27] to the stochastic context based on the Cauchy inequality

$$\begin{aligned} \mathbf{E}[|Q(u - u_\ell)|] &= \mathbf{E}[|b_\omega(u - u_\ell, z - z_\ell)|] \leq \mathbf{E}[\|u - u_\ell\|_\omega \|z - z_\ell\|_\omega] \\ (5.1) \quad &\leq \left(\mathbf{E}[\|u - u_\ell\|_\omega^2] \mathbf{E}[\|z - z_\ell\|_\omega^2]\right)^{1/2}. \end{aligned}$$

The generation of mesh $\mathcal{T}_{\ell+1}$ now follows in the spirit of [27]. For each realisation $u(\omega)$ of the solution on some mesh \mathcal{T}_ℓ , classic FEM analysis gives local refinement indicators for the primal and the dual problem,

$$\begin{aligned} \eta_{\ell,\omega}^{\text{loc}}(T)^2 &:= h_T^2/\kappa_{\max,T} \|f\|_{L^2(T)}^2 + \sum_{E \in \mathcal{E}(T)} h_T/\kappa_{\max,E} \|\kappa \nabla u_\ell \cdot \mathbf{n}_E\|_{L^2(E)}^2, \\ \hat{\eta}_{\ell,\omega}^{\text{loc}}(T)^2 &:= h_T^2/\kappa_{\max,T} \|g\|_{L^2(T)}^2 + \sum_{E \in \mathcal{E}(T)} h_T/\kappa_{\max,E} \|\kappa \nabla z_\ell \cdot \mathbf{n}_E\|_{L^2(E)}^2. \end{aligned}$$

Algorithm 1: Generation of MLMC mesh hierarchy**input** : initial mesh \mathcal{T}_0 , number of samples N , bulk criterion ϑ **output** : hierarchy of meshes $(\mathcal{T}_\ell)_{\ell=0,1,\dots}$ **for** $\ell = 0, 1, \dots$ **do** solve $(u_\ell^i)_{i=0,\dots,N}$ and $(z_\ell^i)_{i=0,\dots,N}$ on \mathcal{T}_ℓ calculate $(\eta_{\ell,i}^{\text{loc}}(T))_{i=0,\dots,N; T \in \mathcal{T}}$ and $(\tilde{\eta}_{\ell,i}^{\text{loc}}(T))_{i=0,\dots,N; T \in \mathcal{T}}$ approximate mean values $(\mathbf{E}_N^{\text{MC}}[\eta_\ell^{\text{loc}}(T)])_{T \in \mathcal{T}}$ and $(\mathbf{E}_N^{\text{MC}}[\tilde{\eta}_\ell^{\text{loc}}(T)])_{T \in \mathcal{T}}$ mark $\mathcal{M}_{u,\ell}$ and $\mathcal{M}_{z,\ell}$ by (5.2) **if** $|\mathcal{M}_{u,\ell}| \leq |\mathcal{M}_{z,\ell}|$ **then** $\mathcal{T}_{\ell+1} = \text{refine}(\mathcal{T}_\ell, \mathcal{M}_{u,\ell})$ **else** $\mathcal{T}_{\ell+1} = \text{refine}(\mathcal{T}_\ell, \mathcal{M}_{z,\ell})$ **end****end**

The elementwise mean value of these indicators yields refinement indicators for the adaptive mesh refinement algorithm as follows. A bulk criterion with parameter $0 < \vartheta \leq 1$ chooses a subset $\mathcal{M}_{u,\ell} \subseteq \mathcal{T}_\ell$ of smallest cardinality such that

$$(5.2) \quad \sum_{T \in \mathcal{M}} \mathbf{E}[\eta_{\ell,\omega}^{\text{loc}}(T)]^2 \geq \vartheta \sum_{T \in \mathcal{T}} \mathbf{E}[\eta_{\ell,\omega}^{\text{loc}}(T)]^2$$

and $\mathcal{M}_{z,\ell} \subseteq \mathcal{T}_\ell$ accordingly. That is, we get an error reduction property for each ω in both the primal and the dual problem. The mean indicator represents a refinement, that propagates optimal error reduction in mean for each problem. Finally, choose the refinement, that is either primal or dual, that gives this property with the least amount of marked elements. This will result in the desired error reduction in Equation (5.1).

For the experiments below we set $\vartheta := 0.5$ and $N := 100$ in Algorithm 1. The refinement of all elements in the chosen set \mathcal{M}_ℓ leads to the output triangulation $\mathcal{T}_{\ell+1}$. Possible further refinements of a closure step guarantee a shape-regular series of triangulations. The algorithm for generating an appropriate hierarchy of meshes based on the goal-oriented error indicators is shown in Algorithm 1.

5.2. On the choice of meshes for the MLMC method. Starting with a coarse initial mesh, uniform refinement generates a sequence of meshes $(\mathcal{T}_\ell^U)_{\ell=0}^L$. The adaptive algorithm from Section 5.1, on the other side, generates an adaptively refined mesh sequence $(\mathcal{T}_k^A)_{k=0}^K$. Then, for every uniform mesh \mathcal{T}_ℓ^U , we select the coarsest adaptive mesh $\mathcal{T}_{k_\ell}^A$ that is finer than \mathcal{T}_ℓ^U , i.e.

$$k_\ell := \min \left\{ k \in \mathbb{N} \mid |V(\mathcal{T}_\ell^U)| \leq |V(\mathcal{T}_k^A)| \right\}$$

This defines a subsequence of adaptively refined meshes $(\mathcal{T}_{k_\ell}^A)_{\ell=0}^L$ with a comparable rate of growth in the number of degrees of freedom as for the uniform meshes.

In the numerical experiments a three level multilevel algorithm is used. Therefore, for each experiment a tuple of three consecutive meshes is selected from each of these sequences.

Subsequently, all possible tuples are used and identified through the number of degrees of freedom on their finest mesh, that is the target mesh on which the solution is computed.

5.3. Notation and parameters. In all experiments, we set $p = 0.98$ and choose the parameters \mathbf{p}^* , \mathbf{p}^- , \mathbf{p}^+ as suggested in Remark 4.8. Furthermore, we always calculate with $N = 10^5$ samples in the Monte Carlo method and with $N_0 = 10^5$ samples on the coarsest grid of the multi level Monte Carlo method. The computational cost C_ℓ for one sample on level ℓ is assumed to be proportional to the number of degrees of freedom ndof . In Equation (3.4) the equivalence constant is set to $N_0/\sqrt{C_0/\mathbf{Var}[Y_0]}$ and the variance on level ℓ is estimated with 100 samples. The unknown exact mean solutions were approximated by a fine grid solution on an adaptively refined mesh with 10^6 degrees of freedom and with 10^6 samples. The number of samples is assumed to be large enough to neglect all the unknown ε^* quantities in Corollaries 4.6 and 4.7.

The convergence history plots below show the goal error $e := |\mathbf{E}[Q(u)] - \mathbf{E}_N^{\text{ML}}[Q(u_h)]|$ and the quantities from Corollary 4.6. The plots also contain the sum of the confidence interval lengths in the upper and lower bounds, i.e.

$$c := \eta^\oplus - \mathbf{E}_N^{\text{ML}}[\eta^+] - (\eta^\ominus - \mathbf{E}_N^{\text{ML}}[\eta^-])$$

and their limit (for $h \rightarrow 0$)

$$c_{\text{lim}} := 2 \mathbf{Var}[Q(u)]^{1/2} N^{-1/2} \Phi^{-1}(p^*).$$

Similarly, $e_{\text{lim}} := \mathbf{Var}[Q(u)]^{1/2} N^{-1/2}$ denotes the remaining stochastic part of (3.1) and (3.2) for $h \rightarrow 0$. As the number of samples stays constant, e_{lim} does not converge to zero and therefore limits the accuracy for the determination of $\mathbf{E}[Q(u)]$. From this point of view it is clear that the convergence rate of e near e_{lim} breaks down.

5.4. First experiment. The first experiment concerns the slit domain $\Omega := (0, 1)^2 \setminus ([0.5, 1] \times \{0.5\})$ and employs the right-hand side data $f \equiv 1$, homogeneous Dirichlet data and the goal weight function

$$(5.3) \quad g = \begin{cases} C^{-1} r^{-2} \exp\left(\frac{-1}{1 - \|x_0 - x\|^2 r^{-2}}\right) & \text{if } \|x_0 - x\| \leq r, \\ 0 & \text{else} \end{cases}$$

where

$$(5.4) \quad C = \int_D r^{-2} \exp\left(\frac{-1}{1 - \|x_0 - x\|^2 r^{-2}}\right) dx.$$

For this experiment, a relatively large goal domain is chosen with radius $r = 0.3$ and center $x_0 = (0.3, 0.3)$.

| ndof on L | N_0 | N_1 | N_2 | ndof on L | N_0 | N_1 | N_2 |
|-------------|-------|-------|-------|-------------|-------|-------|-------|
| 149 | 10000 | 19817 | 5586 | 166 | 10000 | 17620 | 4257 |
| 553 | 10000 | 2699 | 566 | 601 | 10000 | 2315 | 380 |
| 2129 | 10000 | 956 | 182 | 2372 | 10000 | 647 | 220 |
| 8353 | 10000 | 331 | 194 | 9225 | 10000 | 394 | 104 |
| 33089 | 10000 | 387 | 163 | 34929 | 10000 | 211 | 100 |

TABLE 1. The number of samples per level in the MLMC algorithm with uniform (left) and adaptive (right) meshes for the simulation of Section 5.4.

The random field κ is sampled from an expansion similar to that from [14]. The coefficients of this expansion are defined with $\sigma > 1$ and $0 < A < 1/\zeta(\sigma)$ by

$$\begin{aligned}
a_m(x) &:= \alpha_m \cos(2\pi\beta_1(m)x_1) \cos(2\pi\beta_2(m)x_2), \\
\alpha_m &:= Am^{-\sigma}, \\
\beta_1(m) &:= m - k(m)(k(m) + 1)/2, \\
\beta_2(m) &:= k(m) - \beta_1(m), \\
k(m) &:= \lfloor -1/2 + \sqrt{1/4 + 2m} \rfloor,
\end{aligned}$$

where ζ denotes the Riemann zeta function. Now, consider t uniformly distributed random variables φ_m on the interval $[0, 1]$ with $t \in \mathbb{N}$ and the minimum $\alpha_{\min} = \sum_{m=0}^t \alpha_m$ of the sum of the α_m . Then, the random field κ is defined as

$$\kappa(x) := \frac{c}{\alpha_{\min}} \left(\sum_{m=0}^t a_m(x) \varphi_m + \alpha_{\min} \right) + \varepsilon.$$

The last term in the brackets as well as $\varepsilon > 0$ ensures the positivity of the random field whereas $c > 0$ scales the output. In this example the parameters are set to $A = 0.6$, $\sigma = 2$, $t = 5$, $\varepsilon = 5 \cdot 10^{-6}$ and $c = 10^{-3}$. Figure 7 depicts an example realisation.

Figure 1 shows the convergence history of the exact goal error e for a Monte Carlo simulation as defined in Section 5.3. The adaptive mesh refinement clearly leads to a faster approach of the e_{lim} line and to the optimal quadratic convergence (with respect to the average mesh width $h := \text{ndof}^{-1/2}$) while the uniform mesh refinement yields only a reduced convergence order. The upper and lower bounds are guaranteed and efficient. In the case of adaptive mesh refinement the efficiency becomes slightly worse on the finest mesh as the stochastic contribution c begins to dominate the bounds. In this case, the accuracy of the goal error and the efficiency of the bounds can be increased by a larger number of samples.

Figure 2 shows the convergence history of the quantities under investigation for a multi level Monte Carlo simulation. There is no significant difference in the efficiency of the guaranteed bounds compared to MC. Table 1 lists the number of samples for each level of the MLMC algorithm. As expected, the number of samples on the finer grids $\ell = 1$ and $\ell = 2$ decreases with $Y_\ell \rightarrow 0$ for $h \rightarrow 0$. For the adaptively generated meshes this decrease is even faster than for the uniform refined meshes.

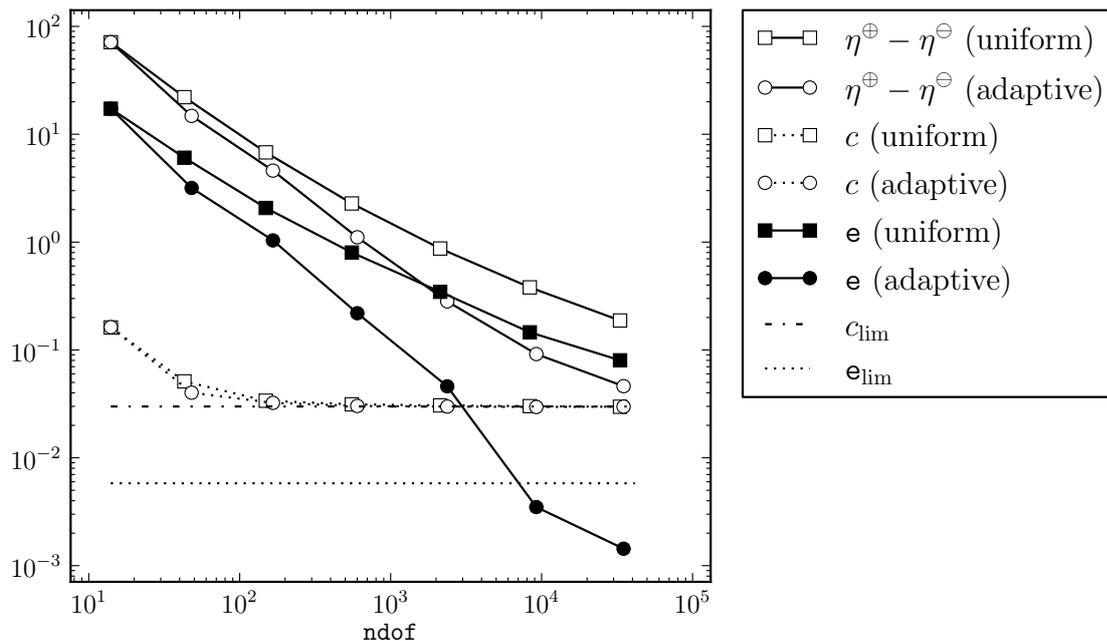


FIGURE 1. Convergence history of the exact goal error, the difference between upper bound and lower bound of the error estimator and the estimates for the stochastic error and their limits as defined in Section 5.3 for the MC simulation of Section 5.4.

Figure 3 shows the convergence history where the mentioned quantities for the MLMC and MC simulations are plotted against the computational costs defined by (3.3). The MLMC method reaches the same error level with much less effort than the MC method. In fact, the costs of MC to reach a certain accuracy are about 10 times higher than the costs for MLMC to reach the same accuracy. Moreover, the adaptive multilevel algorithm outperforms the uniform Monte Carlo method by two to three orders of magnitude. In fact, the uniform convergence rates with respect to the computational cost impede practical applications whereas the adaptive Algorithms render it practically computable. In the latter case, multilevel Monte Carlo reduces the computational cost by an additional order of magnitude.

Figure 8 depicts an adaptive mesh generated by the algorithm from Section 5.1.

5.5. Second experiment. The second experiment considers the problem from Section 5.4 with a different random field κ . Real world examples usually contain high amplitude variations at small scales. Lognormal random fields mimic this behaviour in a suitable fashion. We consider the correlation function with scale σ_{κ} and correlation length λ_{κ} ,

$$C(x, y) := \sigma_{\kappa}^2 \exp(-\|x - y\|_{L^2(D)} / \lambda_{\kappa}^2).$$

The aim is to create a lognormal random field on some triangulation \mathcal{T}_{κ} where the underlying normal random field is generated from the covariance matrix $\mathcal{C} = (C(z_i, z_j))_{i,j=1}^N$ for the vertices $z_i \in \mathcal{N}_{\kappa}$ of \mathcal{T}_{κ} . For a vector of N normally distributed random variables

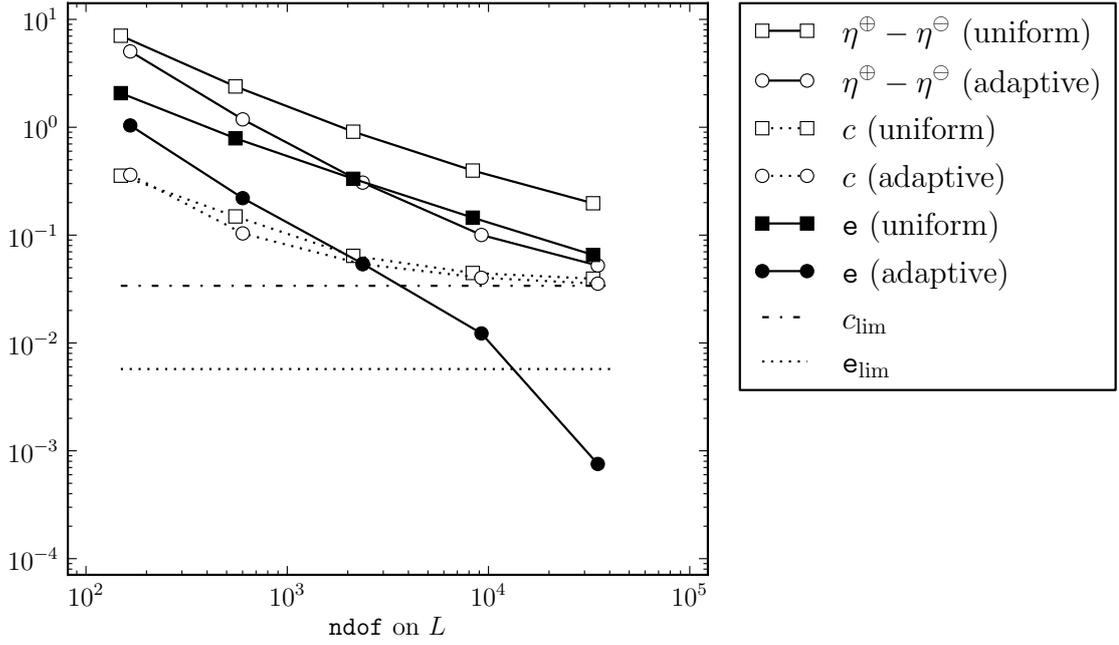


FIGURE 2. Convergence history of the exact goal error, the difference between upper bound and lower bound of the error estimator and the estimates for the stochastic error and their limits as defined in Section 5.3 for the MLMC of Section 5.4.

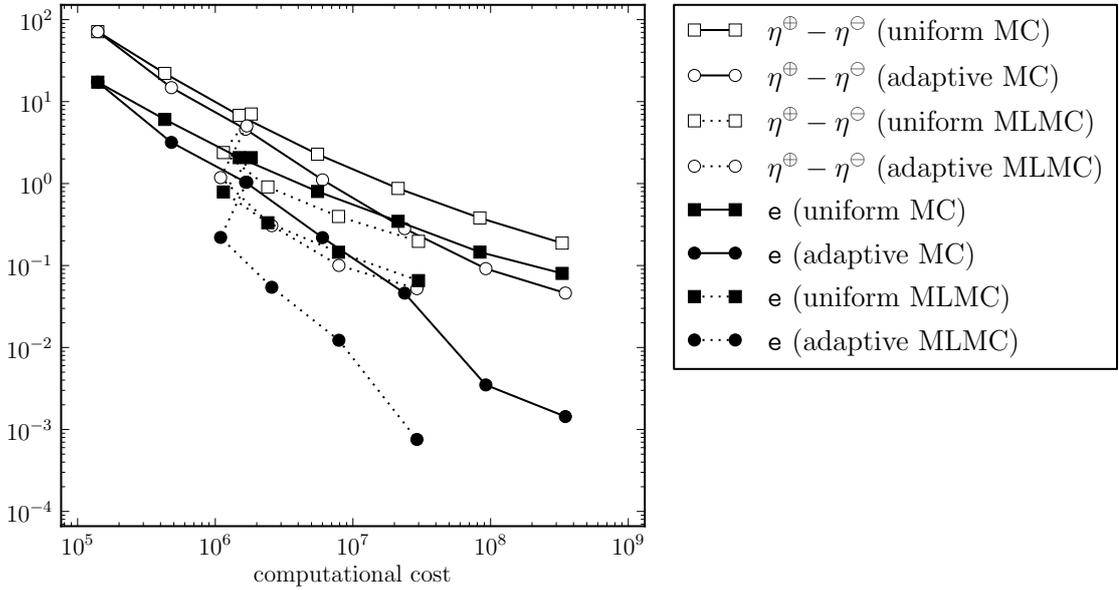


FIGURE 3. Comparison of computational costs vs. exact goal error and error bounds for MC and MLMC simulations of Section 5.4.

| ndof on L | N_0 | N_1 | N_2 | ndof on L | N_0 | N_1 | N_2 |
|-------------|-------|-------|-------|-------------|-------|-------|-------|
| 149 | 10000 | 6049 | 3004 | 153 | 10000 | 6453 | 2640 |
| 553 | 10000 | 4591 | 1529 | 627 | 10000 | 3579 | 1186 |
| 2129 | 10000 | 2502 | 842 | 2295 | 10000 | 2086 | 656 |
| 8353 | 10000 | 1431 | 389 | 9306 | 10000 | 1529 | 396 |
| 33089 | 10000 | 858 | 230 | 36258 | 10000 | 722 | 170 |

TABLE 2. The number of samples per level in the MLMC algorithm with uniform (left) and adaptive (right) meshes for the simulation of Section 5.5.

\mathbf{x} and the Cholesky decomposition $\mathbf{C} = \mathbf{L}\mathbf{L}^T$, it holds $\mathbf{Cov}[\mathbf{L}\mathbf{x}] = \mathbf{C}$. Thus, the $P_1(\mathcal{T}_\kappa)$ interpolation of $\kappa(z_1, \dots, z_N) := \bar{\kappa} + c_\kappa \exp(\mathbf{L}\mathbf{x})$ satisfies the desired properties with any function $\bar{\kappa} > 0$ on D . Here, we set $\sigma_\kappa = 1$, $\lambda_\kappa = 0.3$, $c_\kappa = 10^{-4}$ and $\bar{\kappa} = 10^{-3}$. Figure 7 depicts an example realisation.

Figure 4 shows the convergence history of the exact goal error e for a Monte Carlo simulation as defined in Section 5.3. Also, for this rough stochastic field, the error for adaptive mesh refinement shows the optimal convergence rate. Moreover, the upper and lower goal error bounds are as efficient as in the first example.

The conclusions are similar in case of a multi level Monte Carlo simulation as depicted in Figure 5. For the present random field κ , the numbers of samples N_1 and N_2 decrease not as fast as in the first example (see Table 2).

The comparison of both simulations in terms of computational costs in Figure 3 again renders the multi level Monte Carlo method superior to the standard Monte Carlo method in both the uniform and the adaptive case. With this rough field, the adaptive Monte Carlo method performs just as well as the uniform multilevel Monte Carlo method. The combination of the multilevel method with the adaptive mesh generation leads to an error reduction of one order of magnitude at one tenth of the computational cost compared to uniform Monte Carlo.

Figure 8 depicts an adaptive mesh generated by the algorithm from Section 5.1. Both meshes show refinement at the singularity in the center (mainly induced by refinement indicators for the primal problem) and in the support of the goal functional (mainly induced by refinement indicators for the dual problem).

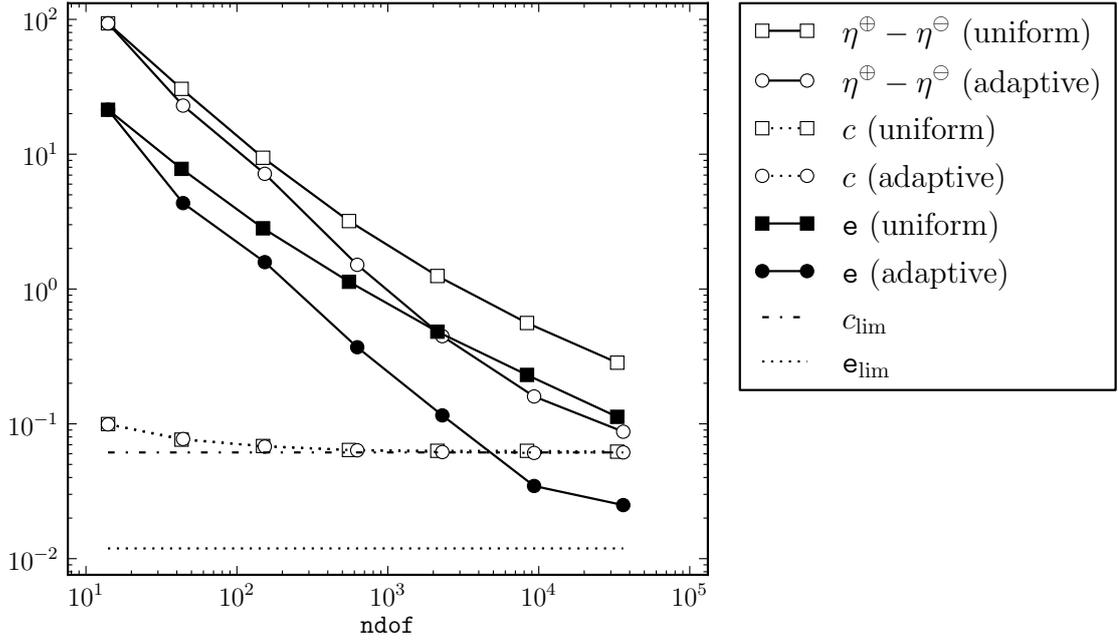


FIGURE 4. Convergence history of the exact goal error, the difference between upper bound and lower bound of the error estimator and the estimates for the stochastic error and their limits as defined in Section 5.3 for the MC simulation of Section 5.5.

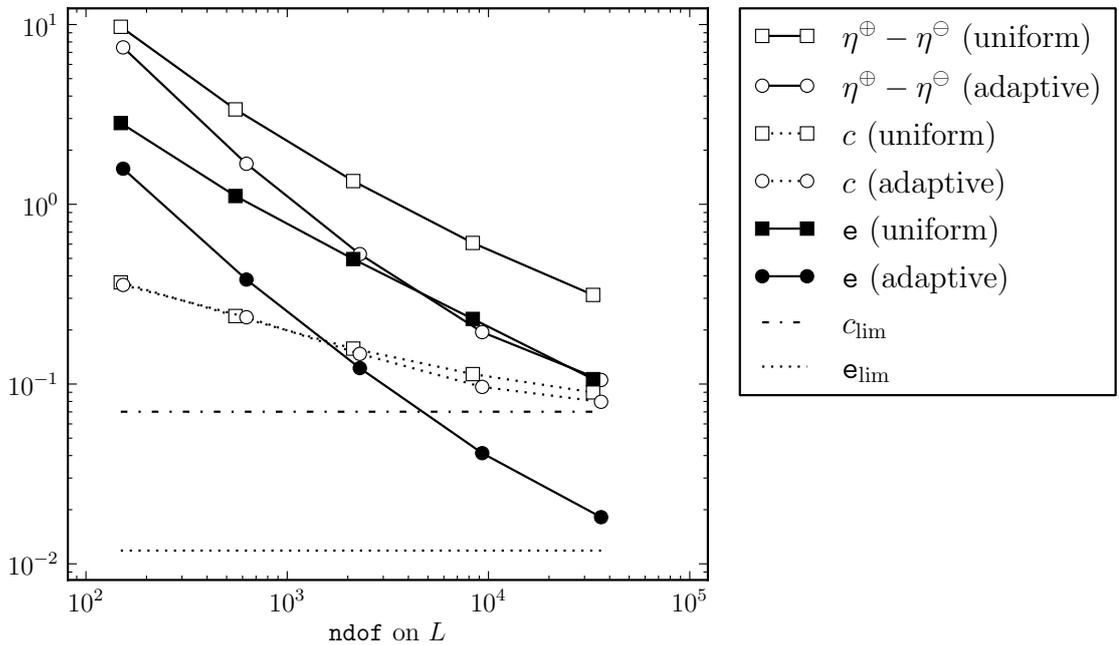


FIGURE 5. Convergence history of the exact goal error, the difference between upper bound and lower bound of the error estimator and the estimates for the stochastic error and their limits as defined in Section 5.3 for the MLMC of Section 5.5.

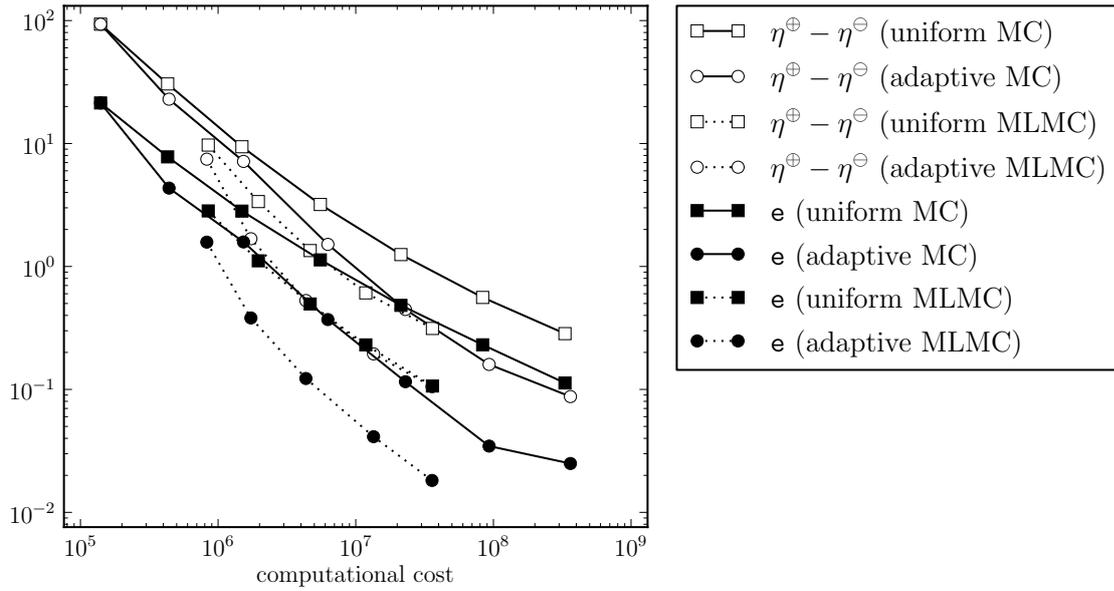


FIGURE 6. Comparison of computational costs vs. exact goal error and error bounds for MC and MLMC simulations of Section 5.5.

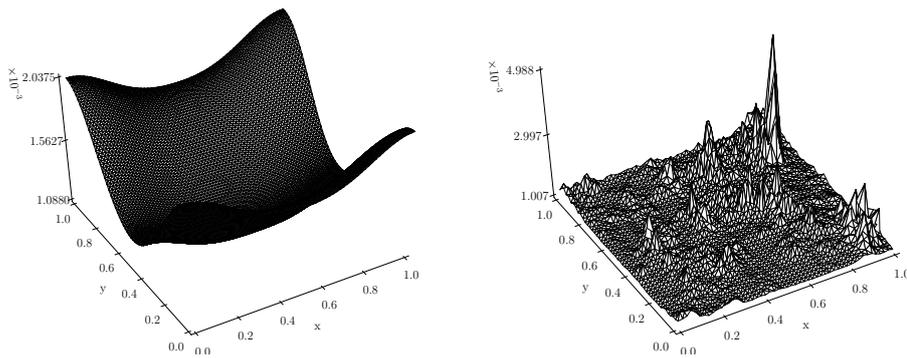


FIGURE 7. Example realisation for the random field κ from Section 5.4 (left) and Section 5.5 (right).

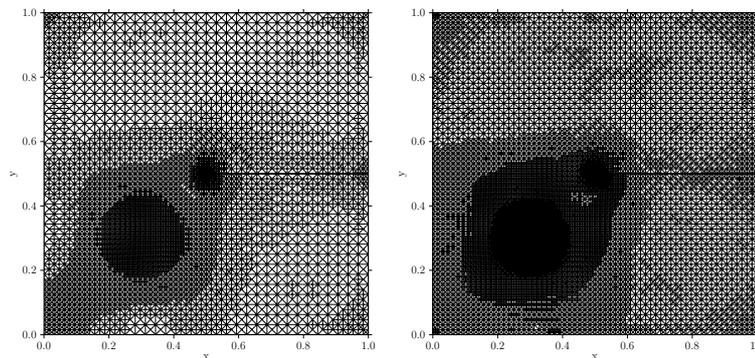


FIGURE 8. Adaptive mesh with approximately 5000 degrees of freedom from Section 5.4 (left) and Section 5.5 (right).

REFERENCES

- [1] M. AINSWORTH, *A posteriori error estimation for lowest order Raviart-Thomas mixed finite elements*, SIAM J. Sci. Comput., 30 (2007/08), pp. 189–204.
- [2] M. AINSWORTH AND J. T. ODEN, *A posteriori error estimation in finite element analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2000.
- [3] A. L. H. ALI, F. NOBILE, E. VON SCHWERIN, AND R. TEMPONE, *Optimization of mesh hierarchies in multilevel monte carlo samplers*, arXiv preprint arXiv:1403.2480, (2014).
- [4] I. BABUŠKA, F. NOBILE, AND R. TEMPONE, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM J. Numer. Anal., 45 (2007), pp. 1005–1034.
- [5] S. BARTELS, C. CARSTENSEN, AND G. DOLZMANN, *Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis*, Numer. Math., 99 (2004), pp. 1–24.
- [6] A. BARTH, A. LANG, AND C. SCHWAB, *Multilevel Monte Carlo method for parabolic stochastic partial differential equations*, BIT, 53 (2013), pp. 3–27.
- [7] A. BARTH, C. SCHWAB, AND N. ZOLLINGER, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, Numer. Math., 119 (2011), pp. 123–161.
- [8] M. BEBENDORF, *A note on the Poincaré inequality for convex domains*, Z. Anal. Anwendungen, 22 (2003), pp. 751–756.
- [9] D. BOFFI, F. BREZZI, AND M. FORTIN, *Mixed finite element methods and applications*, vol. 44 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2013.
- [10] D. BRAESS, *An a posteriori error estimate and a comparison theorem for the nonconforming P_1 element*, Calcolo, 46 (2009), pp. 149–155.
- [11] C. CARSTENSEN, C. MERDON, AND J. NEUMANN, *Aspects of guaranteed error control in CPDEs*, in Numerical Solution of Partial Differential Equations: Theory, Algorithms, and Their Applications, vol. 45 of Springer Proceedings in Mathematics & Statistics, Springer New York, 2013, pp. 103–119.
- [12] 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*, SIAM J. Numer. Anal., 51 (2013), pp. 322–352.
- [13] K. A. CLIFFE, M. B. GILES, R. SCHEICHL, AND A. L. TECKENTRUP, *Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients*, Comput. Vis. Sci., 14 (2011), pp. 3–15.
- [14] M. EIGEL, C. J. GITTELSON, C. SCHWAB, AND E. ZANDER, *Adaptive stochastic galerkin fem*, Computer Methods in Applied Mechanics and Engineering, 270 (2014), pp. 247–269.
- [15] D. ELFVerson, F. HELLMAN, AND A. MÅLQVIST, *A multilevel monte carlo method for computing failure probabilities*, arXiv preprint arXiv:1408.6856, (2014).
- [16] W. FELLER, *An introduction to probability theory and its application*, Wiley, 1971.
- [17] C. J. GITTELSON, J. KÖNNÖ, C. SCHWAB, AND R. STENBERG, *The multi-level Monte Carlo finite element method for a stochastic Brinkman problem*, Numer. Math., 125 (2013), pp. 347–386.
- [18] I. G. GRAHAM, F. Y. KUO, D. NUYENS, R. SCHEICHL, AND I. H. SLOAN, *Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications*, J. Comput. Phys., 230 (2011), pp. 3668–3694.
- [19] M. D. GUNZBURGER, C. G. WEBSTER, AND G. ZHANG, *Stochastic finite element methods for partial differential equations with random input data*, Acta Numer., 23 (2014), pp. 521–650.
- [20] A.-L. HAJI-ALI, F. NOBILE, AND R. TEMPONE, *Multi index monte carlo: when sparsity meets sampling*, arXiv preprint arXiv:1405.3757, (2014).
- [21] O. KALLENBERG, *Foundations of modern probability*, springer, 2002.
- [22] D. KRAFT, *A software package for sequential quadratic programming*, report DFVLR-FB-88-28, Deutsche Forschungs- und Versuchsanstalt fuer Luft- und Raumfahrt e.V. (DFVLR), Koeln (Germany), 1988.
- [23] 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 coefficients*, SIAM J. Numer. Anal., 50 (2012), pp. 3351–3374.

- [24] R. S. LAUGESEN AND B. A. SIUDEJA, *Minimizing Neumann fundamental tones of triangles: an optimal Poincaré inequality*, J. Differential Equations, 249 (2010), pp. 118–135.
- [25] H. G. MATTHIES AND A. KEESE, *Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 1295–1331.
- [26] C. MERDON, *Aspects of guaranteed error control in computations for partial differential equations*, PhD thesis, Humboldt-Universität zu Berlin, 2013.
- [27] M. S. MOMMER AND R. STEVENSON, *A goal-oriented adaptive finite element method with convergence rates*, SIAM Journal on Numerical Analysis, 47 (2009), pp. 861–886.
- [28] J. T. ODEN AND S. PRUDHOMME, *Goal-oriented error estimation and adaptivity for the finite element method*, Computers & mathematics with applications, 41 (2001), pp. 735–756.
- [29] L. PAYNE AND H. WEINBERGER, *An optimal poincaré inequality for convex domains*, Archive for Rational Mechanics and Analysis, 5 (1960), pp. 286–292.
- [30] S. PRUDHOMME AND J. T. ODEN, *On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors*, Comput. Methods Appl. Mech. Engrg., 176 (1999), pp. 313–331. New advances in computational methods (Cachan, 1997).
- [31] C. SCHWAB AND C. J. GITTELSON, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Numer., 20 (2011), pp. 291–467.
- [32] A. L. TECKENTRUP, R. SCHEICHL, M. B. GILES, AND E. ULLMANN, *Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients*, Numer. Math., 125 (2013), pp. 569–600.