Multilevel Monte Carlo methods for Stochastic elliptic multiscale PDES

A. Abdulle, A. Barth, C. Schwab

MULTILEVEL MONTE CARLO METHODS
FOR
STOCHASTIC ELLIPTIC MULTISCALE PDES

ASSYR ABDULLE, ANDREA BARTH, AND CHRISTOPH SCHWAB

ABSTRACT. In this paper Monte Carlo Finite Element (MC FE) approximations for elliptic homogenization problems with random coefficients which oscillate on $n \in \mathbb{N}$ a-priori known, separated length scales are considered. The convergence of multilevel MC FE (MLMC FE) discretizations is analyzed. In particular, it is considered that the multilevel FE discretization resolves the finest physical length scale, but the coarsest FE mesh does not, so that the so-called "resonance" case occurs at intermediate MLMC sampling levels. It is proved that switching to an Hierarchic Multiscale Finite Element method such as the Finite Element Heterogeneous Multiscale method (FE–HMM) to compute all MLMC FE samples on meshes which under-resolve the physical length scales implies once more optimal efficiency (in terms of accuracy versus computational work) for the numerical estimates of statistical moments with first and second order FE–HMMs. Specifically, the method proposed here allows to obtain estimates of the expectation of the random solution, with accuracy versus work that is identical to the solution of a single deterministic problem obtained by a FE–HMM, and which is, moreover, robust with respect to the physical length scales. Numerical experiments corroborate our analytical findings.

1. INTRODUCTION

Monte Carlo (MC) methods are, among others, a key tool to obtain computable estimates of statistical moments of random quantities. In the case of partial differential equations (PDEs) with random inputs, a MC method entails for each realization of the stochastic data the numerical solution of a deterministic PDE. For time dependent, parabolic problems driven by noise (see, e.g., [7, 8, 9, 33, 26]), numerous "realizations" of the PDE in space-time must be simulated.

In order to reduce the computational complexity in such simulations, multilevel Monte Carlo (MLMC) sampling has been introduced, to the authors' knowledge, by M. Giles in [28, 27] for stochastic ordinary differential equations of Itô type after earlier work by S. Heinrich on numerical quadrature (see [30]). The basic idea of MLMC is, however, by no means restricted to stochastic ordinary differential equations: its convergence analysis was generalized to nonlinear, scalar hyperbolic conservation laws with random initial data in [36] and in [37], as a computational strategy (based on the convergence analysis in [36]) also for systems of nonlinear, hyperbolic conservation laws in several space dimensions. In [10], the MLMC algorithm was introduced and analyzed for elliptic PDEs with stochastic coefficients.

Here, we investigate MLMC methods for elliptic PDEs in divergence form, where the coefficients are random with multiple scales; we assume that *the random coefficients' multiple length scales are a-priori known, separated and deterministic.* Such problems arise, among others, in the numerical simulation of subsurface flow problems (see, e.g., [38, 39] and the references therein). In particular, the coefficients are non ergodic and are not amenable to classical homogenization results based on the Birkhoff theorem.

The basic theory for MLMC Finite Element methods (MLMC FE methods) for elliptic problems with stochastic coefficients has recently been developed in [10]. There, it was assumed that the coefficients were random and nondegenerate with a single scale. In [10], the authors show optimal complexity of a combined MLMC FE sampling strategy, for simplicial first and second order Finite Elements in polygonal and polyhedral domains, under natural assumptions on the coefficient. In this context, *optimal* refers to the accuracy versus work in the estimate of (FE approximations of) moments of the unknown stochastic solution. It was shown that for example the expectation of the FE approximation of the solution can be computed in accuracy versus work which equals that of the FE solution of one deterministic, elliptic, one-scale problem with a linear complexity multilevel solver: due to its discretization level dependent sampling strategy, the work of sampling does not increase the asymptotic complexity of the algorithm, as compared to a linear-complexity, deterministic solver. Key to this efficiency of the MLMC discretizations is the simultaneous sampling of the PDE in a *hierarchy of discretizations*, typically (but not necessarily) obtained from a *multilevel* discretization of the differential equation of interest.

In [10], the authors analyzed the MLMC FE method for the numerical solution of the following elliptic model problem

$$(1.1) \qquad -\nabla \cdot (a(\omega, x)\nabla u(\omega, x)) = f(x) \qquad \text{for } x \text{ in } D,$$

with Dirichlet boundary conditions $u|_{\partial D} = 0$ and where $D \subset \mathbb{R}^d$, $2 \le d \in \mathbb{N}$ is a bounded Lipschitz domain (mixed Dirichlet and Neumann boundary conditions are equally admissible, see [10]). For the right hand side $f \in L^2(D)$ was assumed (a stochastic right hand side is also possible if it is independent of a). The random diffusion coefficient a was assumed to be a possibly correlated random field defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ taking values in $L^\infty(D)$. The coefficient is *nondegenerate*, i.e., $\mathbb{P}$-a.s. bounded away from zero as well as from above. Thus, the random solution u of Problem (1.1) not only depends on $x \in D$ but also on a stochastic parameter $\omega \in \Omega$.

For $\mathbb{P}$-a.s. each realization (i.e., for given, fixed $\omega \in \Omega$) of the random diffusion coefficient $a(\omega, \cdot) \in L^\infty(D)$, Equation (1.1) is an elliptic PDE with inhomogeneous coefficient. Such linear, second order elliptic problems can be solved efficiently by multilevel Galerkin Finite Element methods. Here, as in the previous papers [10, 36, 37], we shall be interested in particular in the computation of statistical moments of the stochastic solution by a Monte Carlo method. As it is well known, the convergence rate of the MC method for the statistical estimation of, for example, mathematical expectations is, at best, $1/2$ with respect to the number of samples (given the random variable has finite variance). This entails the necessity of generating a large number of "samples" which, in the context of the present paper, amounts to the numerical solution of one elliptic multiscale PDE per sample.

One of the aims of the multilevel Monte Carlo Finite Element method (MLMC FE method) proposed here is to decrease the cost of this computation to log-linear work and memory in terms of the number of degrees of freedom $N \in \mathbb{N}$, i.e., to the complexity of solving a single deterministic instance of Equation (1.1) by a multilevel method. As it is shown

in [10], for elliptic problems with spatially inhomogeneous random coefficients which vary on a single spatial length scale, this is indeed so, at least for low order Finite Elements in space dimensions $d = 2, 3$. Higher order Finite Element discretizations are shown to be hampered by the convergence rate $1/2$ of the MC method. In the present paper, we investigate the convergence of the MLMC FE method for the numerical solution of elliptic problems with *spatially heterogeneous random coefficients which vary on multiple, a-priori known, deterministic, separated, spatial length scales.*

As in [10], to accommodate the MLMC scheme, we introduce a nested sequence of hierarchic Finite Element spaces (FE spaces), in each of which we calculate a certain number of MC samples of the approximation of the solution. As observed initially by M. Giles in the context of stochastic ordinary differential equations in [28], to balance the errors, this leads on the one hand to a large number of samples on a very coarse grid, whereas on the finest grid only few, expensive samples are drawn. In our setting, on the finest FE mesh, all physical length scales are resolved, but solving the linear system of equations for each MC sample of the random coefficient is costly. In [10], the authors showed how the MLMC FE method can exploit this fact: the number of MC samples is related to an inverse power of the mesh width at each level. This strategy allows to achieve asymptotically the convergence of the MC method on the finest grid (for which the MC samples are chosen accordingly), but the computational costs are only a fraction of the latter. With the use of a full Multigrid solver, the computational costs of the MLMC method are proved to be log-linear in $\mathbb{R}^d$, for $d > 1$. However, in the presence of heterogeneous random coefficients which vary on multiple length scales, the coarse FE grids (which necessarily appear in a MLMC estimator) may not resolve the smallest physical length scale. In turn, the convergence rate estimates in [10] are no longer meaningful for scale under-resolved FE meshes and, further, since so-called "resonance" errors occur at intermediate MLMC sampling levels.

It has been argued that MLMC estimates of discretized ensemble averages where the finest mesh resolves the physical length scales "inherit" the scale-resolving properties of the finest mesh. The (analytical and computational) investigation of this assertion is one purpose of the present paper. The principal conclusion is that even if only the discretization at the finest mesh level used in the MLMC resolves all physical length scales of interest which appear in the exact solution, the large number of scale under-resolved FE approximations on coarse meshes does not seem to affect the asymptotic convergence of the MLMC FE method.

Further, we show in this paper that a MLMC simulation using a Hierarchic Multiscale Finite Element method (FE–HMM) allows *robust convergence rates.* The FE–HMM which we use here was introduced in [23] and analyzed in [1, 2, 24] (see [3, 4] for recent reviews). As we show in the present paper, the FE–HMM allows accurate computation of MLMC samples on a mesh that does not resolve the fine scale of the problem. To resolve the fine scale features of the solution on coarse meshes, the FE–HMM solves a "micro-problem" on a patch with a scale-independent number of oscillations of the finest scale.

The error analysis in the present paper provides bounds on the error of the MLMC FE–HMM for the stochastic, elliptic, multiscale model problem considered, which are explicit with respect to the physical length scale(s) in the stochastic, elliptic problem. The error is bounded by the sum of the errors of the spatial approximation of the FE–HMM and of the error of the MC approximation. We infer that the MC sample size should be increased as a certain function of the mesh width, in order to yield maximum accuracy at minimum computational cost. The error analysis in the present paper also indicates that the efficiency of the MLMC FE–HMM depends on the dimension of the physical domain D: it is nearly

optimal in dimension $d = 2$, and slightly suboptimal in dimension $d = 3$, whereas in dimension $d = 1$, it is suboptimal in terms of work versus accuracy. This is due to the convergence rate 1, in terms of the number of degrees of freedom in the space dimension $d = 1$, for the error in the energy norm achieved by a first order FE method, which requires an excessive number of MC samples to balance statistical and discretization error contributions.

This paper is structured as follows. In Section 2, we formulate the stochastic elliptic multiscale model problem, show its well posedness and, in particular, establish existence and regularity of statistical moments of the stochastic solution. We also recapitulate several recent results from [31] on the homogenization of the stochastic multiscale solutions. In Section 3 we derive and analyze the rate of convergence of the multilevel Monte Carlo Finite Element method. We prove in particular convergence rates of the Monte Carlo approximations of the solution for the non-discretized as well as for the "pathwise" Galerkin Finite Element approximation of the MC samples. In Section 4 we derive bounds for the multilevel Monte Carlo Heterogeneous Multiscale Finite Element method. We resort to the FE–HMM which is a Finite Element method with numerical upscaling to achieve a robust discretization of small (subgrid) length scales in the solution. In the final section, we present numerical experiments in one spatial dimension which corroborate the theoretical results in the preceding sections. We present further numerical experiments with the standard MLMC FE method analyzed in [10] for the random multiscale problems which are under consideration here. An Appendix collects for the readers' convenience a result of [31] on multiscale Karhunen–Loève expansion.

2. Elliptic stochastic multiscale problem

In this section, we present the stochastic, elliptic, multiscale problem, in particular its (pathwise) weak formulation, and its unique solvability and, more importantly, some recent results from [31] on the regularity and the asymptotics, as the periodic length scale in the stochastic coefficient (a in (1.1)) tends to zero. At this point we remark that stochastic, elliptic homogenization problems have been considered before; we mention explicitly [11, 12, 32] and the references there. However, usually only two scales were considered and an ergodic hypothesis was imposed. In this work, following the analysis in [31] neither stationarity nor ergodicity is assumed.

As in the MC convergence theory the *finiteness of second moments* is crucial, we use throughout the remainder of this paper two variational formulations of the elliptic PDE with random inputs: First, *variational stochastic formulations* where the deterministic variational principle is required to hold in mean-square sense with respect to the probability measure. Second, the (stronger) *"pathwise" variational formulation* which is, essentially, the deterministic variational formulation of the problem with the random input considered as a "parameter". Accordingly, we shall refer to this second kind of variational formulation as "parametric, deterministic" variational formulation; rather than elementary events $\omega \in \Omega$, in this formulation the parameter is given in terms of principal components of the input data, via their Karhunen–Loève expansion.

2.1. **Random multiscale diffusion problem.** In a bounded Lipschitz domain $D \subset \mathbb{R}^d$, $d \in \mathbb{N}$, (to which we shall refer as the "physical domain"), we consider random elliptic problems of the form

$$(2.1) \qquad -\nabla \cdot (A\nabla u) = f \qquad \text{for } x \text{ in } D.$$

Here, the diffusion coefficient A resp. the permeability is uncertain in a 'slow', (or macroscopic) variable $x \in D$ and exhibits microstructure in $n \geq 1$, $n \in \mathbb{N}$, 'fast' (or microscopic) length scales. We assume these $n + 1$ length scales to be *separated* and *deterministic* and *a priori known*. To specify the random coefficients' (periodic) microstructure, let Y denote the unit cube in $\mathbb{R}^d$ and let $(Y_i, i = 1, \ldots, n)$ be $n \geq 1$ copies of Y which we assume to be the ranges of the n fast (or "microscopic") variables. We remark that all our results generalize straightforwardly to the case when the Y_i, for $i = 1, \ldots, n$, are nonidentical.

To describe the random coefficients that are admissible in our analysis, we assume that we are given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where, as usual, Ω denotes a set of elementary events, $\mathcal{A} \subset 2^\Omega$ the σ-algebra of all possible events and where $\mathbb{P} : \mathcal{A} \to [0, 1]$ is a probability measure. Throughout, for $0 < p \leq \infty$ and a Banach space B, we denote by $L^p(\Omega; B)$ the Bochner space of strongly measurable mappings from Ω to B which are p-summable (resp. $\mathbb{P}$-a.s. bounded in B in case that $p = \infty$). Further, let $\varphi \in \mathcal{L}(H_1, H_2)$ denote a continuous, linear mapping from H_1 to another separable Hilbert space H_2. For a random field $X \in L^p(\Omega; H_1)$ this mapping defines a random variable $Y = \varphi X \in L^p(\Omega; H_2)$ and

$$\|\varphi X\|_{L^p(\Omega; H_2)} \leq C \|X\|_{L^p(\Omega; H_1)}.$$

Furthermore, there holds

$$\varphi \int_\Omega X \, d\mathbb{P} = \int_\Omega \varphi X \, d\mathbb{P}.$$

We refer to Chapter 1 of [19] for a synopsis of these and further results of Banach-space-valued random variables.

We assume we are given a random field A, defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, taking values in $L^\infty(D; C_\#(Y_1 \times \ldots \times Y_n; \mathbb{S}^d))$, such that

$$(2.2) \qquad A \in L^\infty(\Omega; L^\infty(D; C_\#(Y_1 \times \ldots \times Y_n; \mathbb{S}^d))).$$

With $\mathbb{S}^d$ we denote the space of all real, symmetric $d \times d$ matrices, endowed with the spectral norm. Here and in the following, the notation $\#$ indicates that the functions admit Y_i-periodic extensions to all of $\mathbb{R}^d$ with respect to each of the variables $y_i \in Y_i$, for $i = 1, \ldots, n$, which locally, i.e., on compact subsets of $\mathbb{R}^d$, belong to the same function spaces on the Y_i. We denote by $\mathbf{Y} = Y_1 \times \ldots \times Y_n$ and by $\boldsymbol{y} = (y_1, \ldots, y_n) \in \mathbf{Y}$. Therefore, we write $C_\#(\mathbf{Y}; \mathbb{S}^d)$ in place of $C_\#(Y_1 \times \ldots \times Y_n; \mathbb{S}^d)$. Further, we denote by $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

To ensure well-posedness of the random elliptic problem, given in Equation (2.1), we impose boundedness and ellipticity on the matrix A $\mathbb{P}$-a.s..

Assumption 2.1. The random matrix A satisfies property (2.2) and is, moreover, symmetric coercive, i.e., there exist positive constants α and β such that for $\mathbb{P}$-a.e. $\omega \in \Omega$, for every $x \in D$ and for every $\boldsymbol{y} \in \mathbf{Y}$ it holds, for all $\xi \in \mathbb{R}^d$

$$(2.3) \qquad \alpha |\xi|^2 \leq \xi^\top A(\omega; x, \boldsymbol{y}) \xi \leq \beta |\xi|^2.$$

Next, we parameterize the coefficient A, before we prove well-posedness of Equation (2.1).

2.2. **Parametrization of the stochastic, multiscale coefficient.** We assume that the stochastic coefficient A is characterized by a sequence $\boldsymbol{z} = (z_k, k \in \mathbb{N})$ of real-valued, centered random variables, defined on $(\Omega, \mathcal{A}, \mathbb{P})$, by

$$(2.4) \qquad A(\omega; x, \boldsymbol{y}) = \mu(x, \boldsymbol{y}) + \sum_{k=1}^{\infty} z_k(\omega) \Psi_k(x, \boldsymbol{y}), \quad (\omega, x, \boldsymbol{y}) \in \Omega \times D \times \mathbf{Y},$$

where for $k \in \mathbb{N}$, $\Psi_k \in L^{\infty}(D \times \mathbf{Y}; \mathbb{S}^d)$ and $\mu(x, \boldsymbol{y}) = \mathbb{E}[A(\cdot; x, \boldsymbol{y})]$ denotes the mean of A. One example for such an expansion is the multiscale Karhunen–Loève expansion.

Whenever it holds that $A \in L^2(\Omega; L^{\infty}(D \times \mathbf{Y}; \mathbb{S}^d))$, then a Karhunen–Loève expansion of A exists and converges in $L^2(\Omega; L^{\infty}(D \times \mathbf{Y}; \mathbb{S}^d))$; see Appendix A for details. Additional conditions need to be imposed, however, for pointwise convergence ($\mathbb{P}$-a.s.) to hold.

Without any assumption on the normalization of the z_k or Ψ_k, the parametric representation, given in Equation (2.4), is not unique. We therefore assume that the sequence $(z_k, k \in \mathbb{N})$ consists of independent, identically distributed random variables with uniform distribution[1], i.e., $z_k \sim \mathcal{U}([-1, 1])$. We further denote by $\boldsymbol{z}(\omega)$ the vector $(z_1(\omega), z_2(\omega), \dots) \in U := [-1, 1]^{\mathbb{N}}$ of a realization. In Equation (2.4), for a sequence $\beta = (\beta_k, k \in \mathbb{N}) \in \ell^1$, we assume that the parametric, deterministic matrix functions $(\Psi_k, k \in \mathbb{N})$ satisfy for all $\xi \in \mathbb{R}^d$, $x \in D$, and $\boldsymbol{y} \in \mathbf{Y}$

$$(2.5) \qquad |\xi^{\top} \Psi_k(x, \boldsymbol{y}) \xi| \leq \beta_k |\xi|^2,$$

which implies that the series in Equation (2.4) converges unconditionally, for $\mathbb{P}$-a.e. $\boldsymbol{z}(\omega) \in U$. Further, we assume that there exist positive constants α_0 and β_0 such that the mean of A in Equation (2.4), i.e., the matrix function $\mu \in L^{\infty}(D \times \mathbf{Y}; \mathbb{S}^d)$, satisfies for all $\xi \in \mathbb{R}^d$, $x \in D$, and $\boldsymbol{y} \in \mathbf{Y}$

$$(2.6) \qquad \alpha_0 |\xi|^2 \leq \xi^{\top} \mu(x, \boldsymbol{y}) \xi \leq \beta_0 |\xi|^2.$$

To ensure that the stochastic coefficient A in Equation (2.4) is well defined and coercive, we assume that in Equation (2.4) the (centered) random field $A - \mu$, is bounded by the (deterministic) mean field μ in the following sense.

Assumption 2.2. Let Equation (2.5) and Equation (2.6) be satisfied and assume that in Equation (2.4), the random variables z_k, for $k \in \mathbb{N}$, are independent, identically, and uniformly distributed on $[-1, 1]$. Moreover, with (the ellipticity constant) α_0 in Equation (2.6), the matrix functions μ and Ψ_k satisfy for some $\kappa > 0$

$$\sum_{k \in \mathbb{N}} \beta_k \leq \frac{\kappa}{1 + \kappa} \alpha_0.$$

We remark that Assumption 2.2 implies Assumption 2.1, since we may choose (see [31] for details)

$$\alpha = \alpha_0 - \frac{\kappa}{1 + \kappa} \alpha_0 = \frac{1}{1 + \kappa} \alpha_0 \quad \text{and} \quad \beta = \alpha_0 + \frac{\kappa}{1 + \kappa} \alpha_0.$$

With Equation (2.5) and Assumption 2.2, we have the estimate

$$(2.7) \qquad \sum_{k \in \mathbb{N}} \|\Psi_k\|_{L^{\infty}(D \times \mathbf{Y}; \mathbb{S}^d)} \leq 2 \frac{\kappa}{1 + \kappa} \alpha_0.$$

---

[1] The assumptions can be relaxed, at the expense of additional complications in the notation, while allowing for results which are completely analogous to those in the present paper.

Further follows from Assumption 2.2 that for $\mathbb{P}$-a.e. $\boldsymbol{z}(\omega) = (z_k(\omega), k \in \mathbb{N}) \in U$ the series

$$(2.8) \qquad A(\boldsymbol{z}(\omega); x, \boldsymbol{y}) := \mu(x, \boldsymbol{y}) + \sum_{k=1}^{\infty} z_k(\omega) \Psi_k(x, \boldsymbol{y}),$$

for $(x, \boldsymbol{y}) \in D \times \mathbf{Y}$, converges in $L^{\infty}(D \times \mathbf{Y}; \mathbb{S}^d)$.

A further assumption on the diffusion coefficient is *scale separation*. This means that for a (nondimensional) scale parameter $\varepsilon > 0$, there are $n \in \mathbb{N}$ deterministic, positive functions $(\varepsilon_i, i = 1, \ldots, n)$ which depend continuously and monotonically on ε, with $1 > \varepsilon_1(\varepsilon) \geq \ldots \geq \varepsilon_n(\varepsilon) > 0$, and which describe the n microscopic length scales on which the stochastic diffusion coefficient depends. Without loss of generality, we set $\varepsilon_1(\varepsilon) = \varepsilon$ (or $\varepsilon_1 \equiv \varepsilon$). We say that the random coefficient is *scale separated* if, for all $i = 1, \ldots, n-1$, there holds

$$(2.9) \qquad \lim_{\varepsilon \to 0} \frac{\varepsilon_{i+1}(\varepsilon)}{\varepsilon_i(\varepsilon)} = 0.$$

For the (ε independent!) random coefficient A defined in Equation (2.8), satisfying both Assumption 2.1 and 2.2, and for a given family of scale functions $(\varepsilon_i, i = 1 \ldots, n)$, satisfying Equation (2.9), we define a family of random, multiscale diffusion coefficients $A^{\varepsilon} \in L^{\infty}(\Omega; L^{\infty}(D; \mathbb{S}^d))$ by

$$(2.10) \qquad A^{\varepsilon}(\omega; x) := A(\omega; x, \frac{x}{\varepsilon_1(\varepsilon)}, \ldots, \frac{x}{\varepsilon_n(\varepsilon)}).$$

This we might also express in terms of the coefficient vector $\boldsymbol{z}(\omega) \in U$: for the parametric and deterministic coefficient $A(\boldsymbol{z}; x, \boldsymbol{y})$ as in Equation (2.8) and with the scale functions $(\varepsilon_i, i = 1 \ldots, n)$, we define the *parametric, deterministic multiscale coefficient*

$$(2.11) \qquad A^{\varepsilon}(\boldsymbol{z}; x) := A(\boldsymbol{z}; x, \frac{x}{\varepsilon_1(\varepsilon)}, \ldots, \frac{x}{\varepsilon_n(\varepsilon)}).$$

Here, $\boldsymbol{z} = (z_k(\omega), k \in \mathbb{N}) \in U$ denotes the coefficient vector of one realization for $\omega \in \Omega$ in Equation (2.8). However, we define the random coefficient in this sense independently of $\omega \in \Omega$ for every parameter $\boldsymbol{z} \in U$. In the next section we specify the probability space where A^{ε} is defined on in more detail. Due to Assumption 2.1 and Equation (2.7), for every $\varepsilon > 0$, the stochastic, elliptic, multiscale problem, defined in Equation (2.1), with the stochastic, multiscale coefficient A^{ε} defined in Equation (2.10) and Equation (2.11) admits $\mathbb{P}$-a.s. a unique solution which we denote in what follows by u^{ε}.

2.3. **Image probability space.** A tool in our analysis is a parametric, deterministic representation of the law of the stochastic multiscale solution u^{ε} of Equation (2.1) with stochastic coefficient A^{ε}. We shall use this representation in order to prove various convergence results of u^{ε} as the scale parameter ε tends to 0. Below, we analyze the dependence of this representation of u^{ε} on the parameter vector $\boldsymbol{z}$. This, in turn, also allows for the proof of sharp bounds on spectral approximations of polynomial chaos type of the parametric random solution u^{ε}. To this end, following [17, 18], we represent the law of the random multiscale solution u^{ε} in terms of countably many "random coordinates" z_k which correspond to "principal components" of the random coefficient A in the representation given in Equation (2.4). We collect the sequence $(z_k, k \in \mathbb{N})$ in a vector $\boldsymbol{z}$ and define the *parametric, deterministic multiscale coefficient* $A^{\varepsilon}(\boldsymbol{z}; x)$ as in Equation (2.11).

We define a probability measure on the parameter space U. To this end, we introduce on $U = [-1, 1]^{\mathbb{N}}$ the σ-algebra $\Theta = \mathcal{B}([-1, 1]^{\mathbb{N}})$.

On the measurable space (U, Θ) thus obtained, we define a product probability measure as image measure of $\mathbb{P}$ under the random variable $\boldsymbol{z}$ by

$$(2.12) \qquad\qquad \rho := \bigotimes_{j \in \mathbb{N}} \rho_j,$$

where ρ_j, $j \in \mathbb{N}$, is the uniform distribution on the interval $[-1, 1]$, i.e., $\rho_j(d\xi) = \frac{1}{2}d\xi$. For any set of the form $S = \prod_{j=1}^{\infty} S_j$ with $S_j \in \mathcal{B}([-1, 1])$, it holds $S \in \Theta$ and

$$\rho(S) = \prod_{j=1}^{\infty} \mathbb{P}(\{\omega : \ z_j(\omega) \in S_j\}).$$

With this structure of the probability space, the series in Equation (2.8) converges unconditionally, and almost surely with respect to the (image) measure defined in Equation (2.12), i.e., ρ-a.s.. In the following subsection we shall define the elliptic problem with parametric, deterministic, multiscale coefficient $A^{\varepsilon}(\boldsymbol{z}; x)$.

2.4. **Variational formulation and well-posedness.** With the parametric, deterministic, multiscale coefficient A^{ε} defined in Equation (2.11), for a scale parameter $0 < \varepsilon < 1$ and for $\boldsymbol{z} \in U$, we consider the parametric, deterministic, n-scale Dirichlet problem[2] for x in D by

$$(2.13) \qquad\qquad -\nabla \cdot (A^{\varepsilon}(\boldsymbol{z}; x)\nabla u^{\varepsilon}(\boldsymbol{z}; x)) = f(x),$$

with $u^{\varepsilon}|_{\partial D} = 0$, where ∂D denotes the boundary of the bounded Lipschitz domain $D \subset \mathbb{R}^d$, for $d \in \mathbb{N}$. For simplicity of exposition, we assume, in what follows, that $f \in L^2(D)$ is a given source term which is independent of ε. The significance of Equation (2.13) for Equation (2.1) with coefficient A^{ε} consists in $u^{\varepsilon}(\omega; x) = u^{\varepsilon}(\boldsymbol{z}; x)|_{\boldsymbol{z}=\boldsymbol{z}(\omega)}$.

To present the variational formulation of Equation (2.13) we introduce the Hilbert space

$$(2.14) \qquad\qquad V := H_0^1(D) = \{v \in H^1(D) : \gamma_0 v = 0 \text{ on } \partial D\}.$$

We remark that for bounded Lipschitz domains D the trace operator γ_0 is well-defined and continuous from $H^1(D)$ onto $H^{1/2}(\partial D)$ and, hence, the space V is a closed, linear subspace of $H^1(D)$. By the Poincaré inequality the expressions

$$\|v\|_V := \left(\int_D |\nabla v|^2 dx \right)^{1/2},$$

for $v \in V$ and

$$\|v\|_{L^2(U;V)} := \left(\int_U \int_D |\nabla v|^2 dx \rho(d\boldsymbol{z}) \right)^{1/2}$$

for $v \in L^2(U; V)$ are norms on V and $L^2(U; V)$, respectively. In what follows, we identify $L^2(D)$ with its dual and denote by V^* the dual of V with respect to the "pivot" space $L^2(D)$, i.e., we work in the triplet $V \subset L^2(D) \simeq L^2(D)^* \subset V^*$.

To derive the variational formulation of the stochastic elliptic boundary value problem in Equation (2.13) we fix $\boldsymbol{z} = (z_k(\omega), k \in \mathbb{N}) \in U$ for the moment. We begin our analysis by casting problem (2.13) in variational form: we multiply Equation (2.13) by a test function

---

[2]Mixed boundary conditions could equally be considered, at the expense of a more involved notation.

$v \in V$ and integrate by parts in D to obtain the *parametric, deterministic weak formulation and solution*: for every fixed $\boldsymbol{z} \in U$ and $0 < \varepsilon \leq 1$, find $u^\varepsilon(\boldsymbol{z}; \cdot) \in V$ such that for all $v \in V$

$$(2.15) \qquad b(\varepsilon, \boldsymbol{z}; u^\varepsilon(\boldsymbol{z}; \cdot), v) = \int_D f(x) v(x) \, dx,$$

where the *parametric, deterministic bilinear form* is defined, for $v, w \in V$, by

$$b(\varepsilon, \boldsymbol{z}; w, v) := \int_D A^\varepsilon(\boldsymbol{z}; x) \nabla w(x) \cdot \nabla v(x) \, dx.$$

By Assumption 2.1 and Assumption 2.2, for $\mathbb{P}$-a.e. $\boldsymbol{z}(\omega) \in U$ and $0 < \varepsilon \leq 1$, the deterministic, parametric variational problem admits a unique solution $u^\varepsilon(\boldsymbol{z}, \cdot) \in V$. Moreover, as shown in [31], for every $0 < \varepsilon \leq 1$, the mapping $\boldsymbol{z} \mapsto u^\varepsilon(\boldsymbol{z}, \cdot)$ is measurable, i.e., $u^\varepsilon \in L^0(U; V)$, where we denote the set of all measurable mappings from U to V by $L^0(U; V)$.

Integrating both sides of Equation (2.15) with respect to ρ over U, we derive at the *(integrated) weak formulation* of the parametric, deterministic, elliptic boundary value problem given in Equation (2.13): given the coefficient A^ε, for $0 < \varepsilon \leq 1$, as in Equation (2.11) with separated length scales $\varepsilon_i(\varepsilon)$, for $i = 1, \ldots, n$, satisfying Equation (2.9) and for $f \in V^*$, find $u^\varepsilon \in L^2(U; V)$ such that for all $v \in L^2(U; V)$

$$(2.16) \qquad B(\varepsilon; u^\varepsilon, v) = F(v),$$

where the (ε-dependent) bilinear form $B(\varepsilon; \cdot, \cdot) : L^2(U; V) \times L^2(U; V) \to \mathbb{R}$ is defined by

$$B(\varepsilon; w, v) := \int_U b(\varepsilon, \boldsymbol{z}; w(\boldsymbol{z}; \cdot), v(\boldsymbol{z}; \cdot)) \, \rho(d\boldsymbol{z}),$$

and the linear functional $F(\cdot) \in L^2(U; V)^* \simeq L^2(U; V^*)$ is given by

$$F(v) = \int_U \Big(\int_D f(x) v(\boldsymbol{z}; x) \, dx \Big) \rho(d\boldsymbol{z}).$$

Theorem 2.3. *For a random diffusion coefficient A^ε as defined in Equation (2.11) which satisfies Assumption 2.2 and for every $f \in V^*$ the integrated weak formulation, Equation (2.16), of the stochastic elliptic boundary value problem, given by Equation (2.13), admits a unique solution $u^\varepsilon \in L^2(U; V)$. This solution is measurable as a map $\boldsymbol{z} \mapsto u^\varepsilon(\boldsymbol{z}; \cdot) \in V$ and satisfies for ρ-a.e. $\boldsymbol{z} \in U$ the a-priori estimate*

$$(2.17) \qquad \|u^\varepsilon(\boldsymbol{z}; \cdot)\|_V \leq \frac{\|f\|_{V^*}}{\alpha}.$$

Moreover, the moments of the solution u^ε are bounded, for $1 \leq p \leq \infty$, by

$$(2.18) \qquad \|u^\varepsilon\|_{L^p(U; V)} \leq \frac{\|f\|_{V^*}}{\alpha}.$$

Proof. By Assumption 2.1, we have, for every $0 < \varepsilon \leq 1$ and every $w, v \in L^2(U; V)$

$$(2.19) \qquad |B(\varepsilon; v, w)| \leq \|A^\varepsilon\|_{L^\infty(U; L^\infty(D; \mathbb{S}^d))} \|w\|_{L^2(U; V)} \|v\|_{L^2(U; V)} \leq \beta \|w\|_{L^2(U; V)} \|v\|_{L^2(U; V)}$$

and there exists $\alpha > 0$ such that for every $0 < \varepsilon \leq 1$ holds

$$(2.20) \qquad B(\varepsilon; w, w) \geq \alpha \|w\|_{L^2(U; V)}^2.$$

Moreover, for given $f \in V^*$, we have for all $v \in L^2(U; V)$

$$(2.21) \qquad |F(v)| \leq \|f\|_{V^*} \|v\|_{L^2(U; V)}.$$

The assertion in Equation (2.17) now follows from the Lax-Milgram Lemma. The bound in Equation (2.18) is a consequence of the measurability of the mapping $\boldsymbol{z} \rightarrow u^{\varepsilon}(\boldsymbol{z}; \cdot) \in V$ and of the independence of $\boldsymbol{z}$ of the a-priori estimate in Equation (2.17). □

Remark 2.4. The variational formulation, Equation (2.16), requires in Assumption 2.1 only the definition and boundedness of the random coefficient A^{ε} ρ-almost surely. Assumption 2.1 implies $\mathbb{P}$-a.s. unique solvability and $\mathbb{P}$-a.s. boundedness of the stochastic diffusion problem, defined in Equation (2.13) for ρ-a.e. $\boldsymbol{z} \in U$; this renders the MC FE method and the MLMC FE method well-defined.

2.5. **Regularity of solutions.** Here, we collect several results on the regularity of the solutions u^{ε} of Equation (2.13) which are required for the error analysis of the MLMC FE method as well as of the MLMC FE–HMM. We start with a result on the spatial regularity of u^{ε}. To this end, we impose the following assumption.

Assumption 2.5. The random coefficient A in Assumption 2.2 satisfies

$$A \in L^{\infty}(U; W^{1,\infty}(D \times \mathbf{Y}; \mathbb{S}^d)).$$

Hence, also the coefficient A^{ε} defined in Equation (2.11) satisfies, for every $\varepsilon > 0$,

$$(2.22) \qquad A^{\varepsilon} \in L^{\infty}(U; W^{1,\infty}(D; \mathbb{S}^d)).$$

In [31, Thm. 1.12] it was shown, that for every fixed realization of $\boldsymbol{z}$ in U the family of parametric solutions $(u^{\varepsilon}(\boldsymbol{z}; \cdot), \varepsilon \in (0, 1])$ $(n + 1)$-scale converges as ε tends to 0 to the respective homogenization limit $\bar{u}(\boldsymbol{z}; \cdot) \in V$. This limit solution is measurable as a map from U to V and is, for every $\boldsymbol{z}$, the solution of the parametric, deterministic, one-scale limiting (or "homogenized") elliptic problem

$$(2.23) \qquad -\nabla \cdot (\bar{A}(\boldsymbol{z}; x) \nabla \bar{u}(\boldsymbol{z}; x)) = f(x) \quad \text{for } x \text{ in } D,$$

where $\bar{u}|_{\partial D} = 0$. Here, the effective (or homogenized) parametric coefficient $\bar{A}(\boldsymbol{z}; \cdot)$ is defined, for every $\boldsymbol{z}$ in U, in the usual fashion (see the proof of [31, Prop. 4.1] for an explicit formula of $\bar{A}(\boldsymbol{z}; x)$). The corresponding weak form is similar to Equation (2.15) with the homogenized, parametric, deterministic coefficient $\bar{A}(\boldsymbol{z}; x)$ in place of $A^{\epsilon}(\boldsymbol{z}; x)$: find $\bar{u} \in L^2(U; V)$ such that for all $v \in L^2(U; V)$

$$(2.24) \qquad \bar{B}(\bar{u}, v) = F(v).$$

Here, the bilinear form $\bar{B}$ corresponding to the random, homogenized problems is defined by

$$(2.25) \qquad \bar{B}(w, v) := \int_U \bar{b}(\boldsymbol{z}; w(\boldsymbol{z}, \cdot), v(\boldsymbol{z}, \cdot)) \, \rho(d\boldsymbol{z}),$$

where

$$\bar{b}(\boldsymbol{z}; w, v) := \int_D \bar{A}(\boldsymbol{z}; x) \nabla w \cdot \nabla v \, dx.$$

The right hand side of Equation (2.24) is defined as in Equation (2.16). Assumption 2.1 and 2.5 imply local $H^2(D)$ regularity of the homogenized solution $\bar{u}(\boldsymbol{z}; \cdot)$ (see also Remark 3.3 ahead). As in standard homogenization theory, it can be shown that $\bar{A}(\boldsymbol{z}; \cdot)$ is again symmetric and uniformly bounded and coercive, thus Equation (2.19) and Equation (2.20) hold for the bilinear form $\bar{B}(v, w)$ of the variational formulation in Equation (2.24).

Remark 2.6. Under Assumption 2.5, for $f \in L^2(D)$, the solutions $u^\varepsilon(\boldsymbol{z}; \cdot)$ of the parametric, deterministic multiscale problem (2.16) and the solution $\bar{u}(\boldsymbol{z}; \cdot)$ of the parametric, homogenized problem (2.23) belong, as functions of the spatial variable $x \in D$, to a regularity space W which is contained in $H^2_{loc}(D)$ but has to be specified in each concrete application. For example, in the case that the tensor A is isotropic and the domain D is convex, we have $W = H^2(D) \cap H^1_0(D)$. In general, however, we only have $W = H^2_\gamma(D)$ where the subscript γ indicates a weighted $L^2(D)$ bound on the second derivatives of u, with a weight function which depends on the distance to corners (and to the edges in the case $d = 3$) of the domain D (see, e.g., [29] for details). There holds, for $f \in L^2(D)$, the regularity estimate

$$(2.26) \qquad \sup_{\boldsymbol{z} \in U} \|\bar{u}(\boldsymbol{z}; \cdot)\|_{H^2_\gamma(D)} \le C \|f\|_{L^2(D)},$$

and furthermore, since the bound is independent of $\boldsymbol{z} \in U$

$$(2.27) \qquad \|\bar{u}\|_{L^2(U; H^2_\gamma(D))} \le C \|f\|_{L^2(D)}.$$

Analogously, for the solutions u^ε of the parametric, deterministic multiscale problem we have

$$\sup_{\boldsymbol{z} \in U} \|u^\varepsilon(\boldsymbol{z}; \cdot)\|_{H^2_\gamma(D)} \le \frac{C}{\varepsilon_n(\varepsilon)} \|f\|_{L^2(D)}$$

and

$$(2.28) \qquad \|u^\varepsilon\|_{L^2(U; H^2_\gamma(D))} \le \frac{C}{\varepsilon_n(\varepsilon)} \|f\|_{L^2(D)},$$

with a constant $C > 0$ independent of ε. This result follows from a direct application of the second fundamental inequality for elliptic operators (see [35, Chap. 2.6]) provided Assumption 2.5 holds and using $\|A^\varepsilon\|_{L^\infty(U; W^{1,\infty}(D; \mathbb{S}^d))} \le C/\varepsilon_n$, which follows from Equation (2.11).

3. Multilevel Monte Carlo Finite Element method

We are interested in estimating the expectations $\mathbb{E}[u^\varepsilon]$ and $\mathbb{E}[\bar{u}]$ of the random solutions $u^\varepsilon, \bar{u}$ of the variational formulations (2.16) and (2.24), respectively. Before giving specific bounds to our problems we recapitulate the general idea of MC methods.

3.1. **Monte Carlo method.** The Monte Carlo (MC) estimator of the mathematical expectation of a square integrable random variable u taking values in a separable Hilbert space V, is given by the average over M random variables u^i, for $i = 1, \ldots, M$, i.e.,

$$(3.1) \qquad E_M[u] := \frac{1}{M} \sum_{i=1}^{M} u^i.$$

Here, the u^i are M independent, identically distributed copies of the random variable u. Therefore, we have $E_M[u] \in L^2(U; V)$ and, by the law of large numbers, the sample average $E_M[u]$ converges to $\mathbb{E}[u]$ as M tends to ∞. For a Monte Carlo simulation, Equation (3.1) amounts to generating M independent draws $\hat{\boldsymbol{z}}^i \in U$ of the parameter vector $\boldsymbol{z} \in U$ which are identically, uniformly distributed and then to generate corresponding samples $u^{\varepsilon,i} := u^\varepsilon(\hat{\boldsymbol{z}}^i; \cdot) \in V$ of u^ε by solving the variational problem (2.15).

The following result is a bound on the variance of the Monte Carlo estimator E_M, assuming *availability of the solutions u^i, for $i = 1, \ldots, M$.*

Lemma 3.1. *For any $M \in \mathbb{N}$ and for $u \in L^2(U;V)$ holds*

$$(3.2) \qquad \|\mathbb{E}[u] - E_M[u]\|_{L^2(U;V)} = M^{-1/2}\,Var\,(u) \le M^{-1/2}\|u\|_{L^2(U;V)}.$$

For the proof, we refer to [10, Lemma 4.1] or to [41]. Based on Equation (3.1), we estimate the expectation of u^ε, and $\bar{u}$ by

$$(3.3) \qquad E_M[u^\varepsilon] := \frac{1}{M}\sum_{i=1}^{M} u^{\varepsilon,i} \quad \text{and} \quad E_M[\bar{u}] := \frac{1}{M}\sum_{i=1}^{M} \bar{u}^i.$$

Since $u^\varepsilon \in L^2(U;V)$ and $\bar{u} \in L^2(U;V)$ Lemma 3.1 applies to both.

Remark 3.2. Based on the a-priori estimates in Equation (2.17) and Equation (2.18), we immediately infer from Lemma 3.1 that the *convergence rate of the MC estimator is independent of the scale parameter ε.*

In the preceding argument we assumed that realizations of the solution $u^{\varepsilon,i}$ appearing in the definition of the MC estimator (3.3) are on one hand well-defined and on the other exactly available. That they are well defined is ensured by the ρ-a.s. (with respect to the parameter vector $\boldsymbol{z}$) well-posedness of the parametric, deterministic problem. However, they are in general not exactly available and therefore *approximate pathwise solutions must be used instead.* This introduces an additional discretization error in the MC estimate.

Aiming at MLMC methods, we introduce next so-called *hierarchic FE discretizations* of the stochastic elliptic problem (2.13).

3.2. Finite Element discretization.

We consider continuous, piecewise linear Finite Elements based on sequences of regular simplicial meshes of quasi–uniform triangles or tetrahedra $(\mathcal{T}_l, l \in \mathbb{N}_0)$ of the polygonal, respectively polyhedral, domain D. For any level $l \in \mathbb{N}_0$ of mesh refinement, we denote the mesh width of $\mathcal{T}_l$ by

$$h_l = \max\{h_K : K \in \mathcal{T}_l\},$$

with $h_K := \mathrm{diam}(K)$. We recall (see, e.g., [13, 14]) that the nested family $(\mathcal{T}_l, l \in \mathbb{N}_0)$ of regular, simplicial meshes is called κ-*shape regular* if and only if there exists a $\kappa < \infty$ such that

$$\kappa := \sup_l(\kappa_l) = \sup_l(\max_{K \in \mathcal{T}_l}(\frac{h_K}{\rho_K})).$$

Here ρ_K is the radius of the largest ball that can be inscribed into any $K \in \mathcal{T}_l$.

We generate the sequence $(\mathcal{T}_l, l \in \mathbb{N}_0)$ of uniformly refined, shape regular, simplicial meshes by regular subdivision of an initial, regular triangulation $\mathcal{T}_0$ of D. This results in the mesh width $h_l = 2^{-l}h_0$, where h_0 denotes the mesh width of $\mathcal{T}_0$. Since regular subdivision preserves shape-regularity, the nested family $(\mathcal{T}_l, l \in \mathbb{N}_0)$ of regular, simplicial triangulations obtained in this way is κ-shape regular, ie. $\kappa_l = \kappa_0 = \kappa$. For $l \in \mathbb{N}_0$, we define the FE spaces

$$(3.4) \qquad \mathcal{S}^1(D, \mathcal{T}_l) = \{v \in H^1(D) : v|_K \in \mathcal{P}_1,\ \forall K \in \mathcal{T}_l\},$$

where, as usual, $\mathcal{P}_1 = \mathrm{span}\{x^\alpha : |\alpha| \le 1\}$ denotes the space of polynomials of total degree not exceeding 1. We denote further by

$$\mathcal{S}_0^1(D, \mathcal{T}_l) = \{v \in H_0^1(D) : v|_K \in \mathcal{P}_1,\ \forall K \in \mathcal{T}_l\} = \mathcal{S}^1(D, \mathcal{T}) \cap H_0^1(D),$$

the corresponding Finite Element space taking into account the Dirichlet boundary condition. The spaces $\mathcal{S}_0^1(D, \mathcal{T}_l)$ are well-known to have the approximation property (see, e.g., [15])

$$(3.5) \qquad \inf_{v \in \mathcal{S}_0^1(D, \mathcal{T}_l)} \|w - v\|_{H^1(D)} \leq C h_l \|w\|_W,$$

where W is as in Remark 2.6.

Remark 3.3. In polygonal or polyhedral domains D, only *local $H^2(D)$ regularity* can be asserted due to possible corner (and edge for $d = 3$) singularities; to reduce technicalities in our exposition, we do not develop this aspect here. All FE approximation results which follow admit, however, direct generalizations which cover this situation, *with identical conclusions*, if the triangulations $\mathcal{T}_l$, $l \in \mathbb{N}_0$, are suitably refined towards corners (and edges for $d = 3$).

We introduce the family $\mathcal{V} = (V_l, l \in \mathbb{N}_0)$ of subspaces of V defined by $V_l := \mathcal{S}_0^1(D, \mathcal{T}_l) \subset H_0^1(D)$ with finite dimensions $N_l = \dim(V_l)$. Since the triangulations in the sequence $(\mathcal{T}_l, l \in \mathbb{N}_0)$ are nested, we have

$$(3.6) \qquad V_0 \subset V_1 \subset \ldots \subset V_l \subset \ldots \subset V.$$

The FE approximation is based on $L^2(U; V)$-Galerkin projections of the weak formulation in Equation (2.16) of the stochastic, elliptic boundary value problem in Equation (2.13). The discrete problem corresponding to the integrated weak formulation, Equation (2.16), reads: find $u_{l,FEM}^\varepsilon \in L^2(U; V_l)$ such that for all $v_l \in L^2(U; V_l)$

$$(3.7) \qquad B(\varepsilon; u_{l,FEM}^\varepsilon, v_l) = F(v_l).$$

In [10, Theorem 4.3], the following mean square error bound for the MC FE estimate of $\mathbb{E}[u^\varepsilon]$ was obtained for solutions u^ε which belong to $H^2(D)$ for every $0 < \varepsilon \leq 1$ and for every $\boldsymbol{z} \in U$:

$$(3.8) \qquad \|\mathbb{E}[u] - E_M[u_{l,FEM}^\varepsilon]\|_{L^2(U;V)} \leq C \left(\frac{1}{\sqrt{M}} \|u^\varepsilon\|_{L^2(U;V)} + h_l \|u^\varepsilon\|_{L^2(U;H^2(D))} \right).$$

We see from Equation (3.8) that the combined MC FE error consists of a *statistical error* and of a *discretization error*. With Equation (2.17), the statistical error is, indeed, of order $O(M^{-1/2})$. On the other hand, under the condition of $H^2(D)$ regularity, the discretization error admits, due to the fine-scale components of the pathwise solutions $u^\varepsilon(\boldsymbol{z}, \cdot)$, only an ε-dependent bound (cf. Remark 2.6).

This implies for the bound in Equation (3.8), combined with Equation (3.5) and the a-priori estimate in Theorem 2.3, the error bound

$$(3.9) \qquad \|\mathbb{E}[u^\varepsilon] - E_M[u_{l,FEM}^\varepsilon]\|_{L^2(U;V)} \leq C \left(\frac{1}{\sqrt{M}} + \left(1 \wedge \frac{h_l}{\varepsilon_n(\varepsilon)} \right) \right) \|f\|_{L^2(D)}.$$

Here, we used the (uniform with respect to ε and to l) $L^2(U; V)$-boundedness of the Finite Element projection and Assumption 2.1. For two positive numbers $a \wedge b$ denotes the minimum of a and b.

From Equation (3.9), convergence can therefore only be expected if the mesh width h_l satisfies the *scale-resolution requirement* $h_l \leq \varepsilon_n(\varepsilon)$.

Error bounds which are independent of ε can be obtained for the homogenized solution in Equation (2.24) (which, however, is in general not explicitly available since the homogenized tensor $\bar{A}(\boldsymbol{z}; x)$ is not explicitly known). Its solution $\bar{u}$ is still a random variable but

is independent of ε. The FE approximation of the random, homogenized problem is to find $\bar{u}_{l,FEM} \in L^2(U; V_l)$ such that there holds

$$(3.10) \qquad \bar{B}(\bar{u}_{l,FEM}, v_l) = F(v_l),$$

for all $v_l \in L^2(U; V_l)$, where the load functional F is given by Equation (2.16). Here, under the assumption $f \in L^2(D)$ and $\bar{u} \in L^2(U; H^2(D))$, we have the error bound

$$(3.11) \qquad \|\mathbb{E}[\bar{u}] - E_M[\bar{u}_{l,FEM}]\|_{L^2(U;V)} \leq C \left(\frac{1}{\sqrt{M}} + h_l \right) \|f\|_{L^2(D)}.$$

3.3. Multilevel Monte Carlo Finite Element method.

For the MLMC FE method we discretize the parametric, deterministic variational formulation, given by Equation (2.15), by a Galerkin projection onto *all* subspaces $(V_l, l = 1, \ldots, L)$, for $L \in \mathbb{N}_0$. To derive the MLMC FE estimator for $\mathbb{E}[v]$ for $v \in L^2(U; V)$, we denote by $v_l \in L^2(U; V_l)$ projections (such as, e.g., FE projections) of v onto V_l. With the convention $v_{-1} \equiv 0$, we use the linearity of the mathematical expectation to write for every $L \in \mathbb{N}_0$

$$(3.12) \qquad \mathbb{E}[v_L] = \mathbb{E}\Big[\sum_{l=0}^{L} (v_l - v_{l-1}) \Big] = \sum_{l=0}^{L} \mathbb{E}[v_l - v_{l-1}].$$

The MLMC estimator for $\mathbb{E}[u^\varepsilon_{L,FEM}]$ is based on replacing the expectation in each term on the right hand side of Equation (3.12) by its MC estimator E_{M_l} with M_l samples (possibly different for each level $l = 1, \ldots, L$). Thus, we obtain the *MLMC FE estimator* with $L \in \mathbb{N}_0$ levels of mesh-refinement and with level-dependent sample sizes M_l (with the convention that $u^\varepsilon_{-1,FEM} \equiv 0$), defined by

$$(3.13) \qquad E^L[u^\varepsilon_{L,FEM}] := \sum_{l=0}^{L} E_{M_l} \left[u^\varepsilon_{l,FEM} - u^\varepsilon_{l-1,FEM} \right].$$

Here, the MC estimators $(E_{M_l}, l \in \mathbb{N}_0)$ are defined as in Equation (3.1). The FE solutions in two successive terms in the telescoping sum are required on the same refinement level $l - 1$, but are sampled at M_l resp. M_{l-1} *independent* realizations of the random coefficient. In [10, Lemma 4.4], an a-priori error bound for the MLMC FE estimator for $\mathbb{E}[u]$ was proved. In our setting this yields the following error bound.

Theorem 3.4. *For $L \in \mathbb{N}_0$, under Assumption 2.5, the MLMC FE estimator, given in Equation (3.13), of the expectation $\mathbb{E}[u^\varepsilon]$ of the solution $u^\varepsilon \in L^2(U; W)$ to the random, elliptic, multiscale problem in Equation (2.13), in the polyhedral domain $D \subset \mathbb{R}^d$ admits the error bound*

$$(3.14)
\begin{aligned}
&\|\mathbb{E}[u^\varepsilon] - E^L[u^\varepsilon_{L,FEM}]\|_{L^2(U;V)} \\
&\leq C \left(\left(1 \wedge \frac{h_L}{\varepsilon n(\varepsilon)}\right) + \frac{1}{\sqrt{M_0}} + \sum_{l=1}^{L} \left(1 \wedge \frac{h_l}{\varepsilon n(\varepsilon)}\right) \frac{1}{\sqrt{M_l}} \right) \|f\|_{L^2(D)}.
\end{aligned}$$

Here, the constant C is independent of $0 < \varepsilon \leq 1$, M_l and h_l, for $0 \leq l \leq L$.

Proof. The proof follows Lemma 4.4 in [10], but we track the dependence on ε of the constants in the argument. With the convention that $u_{-1,FEM} \equiv 0$, we write

$$
\begin{aligned}
&\|\mathbb{E}[u^\varepsilon] - E^L[u^\varepsilon_{L,FEM}]\|_{L^2(U;V)} \\
&\quad = \|\mathbb{E}[u^\varepsilon] - \mathbb{E}[u^\varepsilon_{L,FEM}] + \mathbb{E}[u^\varepsilon_{L,FEM}] - \sum_{l=0}^L E_{M_l}[u^\varepsilon_{l,FEM} - u^\varepsilon_{l-1,FEM}]\|_{L^2(U;V)} \\
&\quad \leq \|\mathbb{E}[u^\varepsilon] - \mathbb{E}[u^\varepsilon_{L,FEM}]\|_{L^2(U;V)} + \|\sum_{l=0}^L \big((\mathbb{E} - E_{M_l})[u^\varepsilon_{l,FEM} - u^\varepsilon_{l-1,FEM}]\big)\|_{L^2(U;V)} \\
&\quad =: I + II.
\end{aligned}
$$

As in standard FE theory, for every $l \in \mathbb{N}_0$ and $\boldsymbol{z} \in U$, holds the following approximation property for the FE solution

$$(3.15) \qquad \|u^\varepsilon(\boldsymbol{z};\cdot) - u^\varepsilon_{l,FEM}(\boldsymbol{z};\cdot)\|_V \leq C \inf_{v_l \in V_l} \|v_l - u^\varepsilon(\boldsymbol{z};\cdot)\|_V.$$

From Assumption 2.1, we infer the mean-square stability of the Galerkin projection, i.e.

$$(3.16) \qquad \sup_{\boldsymbol{z} \in U} \|u^\varepsilon_{l,FEM}(\boldsymbol{z};\cdot)\|_V \leq C\alpha^{-1}\|f\|_{V^*}$$

where the constant $C > 0$ depends on the parameter β in Assumption 2.1, but is independent of ε and of $l \in \mathbb{N}_0$. We estimate the terms I and II separately, ignoring (to ease notation) temporarily (3.16).

Term I: By Jensen's and the Cauchy–Schwarz inequality, we get with Equation (3.15), Equation (3.5) and by Equation (2.28)

$$I \leq \|u^\varepsilon - u^\varepsilon_{L,FEM}\|_{L^2(U;V)} \leq Ch_L\|u^\varepsilon\|_{L^2(U;W)} \leq C\frac{h_L}{\varepsilon_n(\varepsilon)}\|f\|_{L^2(D)}.$$

For Term II we use the triangle inequality and Lemma 3.1 to get

$$
\begin{aligned}
II &\leq \sum_{l=0}^L \|(\mathbb{E} - E_{M_l})[u^\varepsilon_{l,FEM} - u^\varepsilon_{l-1,FEM}]\|_{L^2(U;V)} \\
&\leq \sum_{l=0}^L \frac{1}{\sqrt{M_l}}\|u^\varepsilon_{l,FEM} - u^\varepsilon_{l-1,FEM}\|_{L^2(U;V)}.
\end{aligned}
$$

For the first term of the sum ($l = 0$) we have with Theorem 2.3, the stability of the Finite Element approximation, and the convention $u^\varepsilon_{-1,FEM} \equiv 0$ the bound

$$\frac{1}{\sqrt{M_0}}\|u^\varepsilon_{0,FEM}\|_{L^2(U;V)} \leq \frac{1}{\sqrt{M_0}}\big(\frac{1}{\alpha}\|f\|_{L^2(D)}\big) \leq C\frac{1}{\sqrt{M_0}}\|f\|_{L^2(D)}.$$

All the other terms in the sum ($l = 1, \ldots, L$) we bound with the triangle inequality, Equation (3.15), and Equation (3.5) to obtain

$$
\frac{1}{\sqrt{M_l}} \|u_{l,FEM}^\varepsilon - u_{l-1,FEM}^\varepsilon\|_{L^2(U;V)} \leq \frac{1}{\sqrt{M_l}} \left(\|u^\varepsilon - u_{l,FEM}^\varepsilon\|_{L^2(U;V)} + \|u^\varepsilon - u_{l-1,FEM}^\varepsilon\|_{L^2(U;V)} \right)
$$

$$
\leq C \frac{1}{\sqrt{M_l}} (h_l + h_{l-1}) \|u^\varepsilon\|_{L^2(U;W)}
$$

$$
\leq C \frac{1}{\sqrt{M_l}} (h_l + h_{l-1}) \frac{1}{\varepsilon_n(\varepsilon)} \|f\|_{L^2(D)}
$$

$$
\leq C \frac{1}{\sqrt{M_l}} \frac{h_l}{\varepsilon_n(\varepsilon)} \|f\|_{L^2(D)}.
$$

Summing this bound for $l = 1, \ldots, L$ we find

$$
\|\mathbb{E}[u^\varepsilon] - E^L[u_{L,FEM}^\varepsilon]\|_{L^2(U;V)} \leq C \Big(\frac{h_L}{\varepsilon_n(\varepsilon)} + \frac{1}{\sqrt{M_0}} + \sum_{l=1}^{L} \frac{1}{\sqrt{M_l}} \frac{h_l}{\varepsilon_n(\varepsilon)} \Big) \|f\|_{L^2(D)}.
$$

Using Equation (3.16) we may replace in this bound the quotients $h_l/\varepsilon_n(\varepsilon)$ for $l = 0, \ldots, L$ by $1 \wedge h_l/\varepsilon_n(\varepsilon)$ which implies the assertion. $\qquad\square$

Remark 3.5. Comparing the singlelevel MC and the multilevel MC *mean-square* error bounds, i.e., Equation (3.9) and Equation (3.14), respectively, we observe that the discretization error contributions from MLMC at under-resolved FE mesh levels (where $h_l > \varepsilon_n(\varepsilon)$) still converge in mean square for large numbers of samples. This effect is not present in the singlelevel MC FE method.

We gather from Theorem 3.4 that we have to distinguish two cases. Namely, the first, where all levels are under-resolved, i.e., $h_L \geq \varepsilon_n(\varepsilon)$ and the other where only the first levels are under-resolved but at least the finest one is resolved, i.e., $h_L < \varepsilon_n(\varepsilon)$. In the first case the sample number on the coarsest level M_0 should be chosen as

$$
M_0 = \mathrm{O}\Big(\frac{(\varepsilon_n(\varepsilon))^2}{h_L^2} \Big) = \mathrm{O}((\varepsilon_n(\varepsilon))^2 2^{2L})
$$

and for all other levels $l = 1, \ldots, L$ as

$$
M_l = \mathrm{O}\Big(\frac{(\varepsilon_n(\varepsilon))^2}{h_L^2} l^{2(1+\eta)} \Big) = \mathrm{O}((\varepsilon_n(\varepsilon))^2 2^{2L} l^{2(1+\eta)}).
$$

However, we can only expect convergence of the Monte Carlo part of the error contribution in this case. In the case where we have under-resolved levels up to level $l^*(\varepsilon_n)$, for $0 \leq l^*(\varepsilon_n) < L$, meaning $h_{l^*(\varepsilon_n)} \geq \varepsilon_n(\varepsilon)$ and $h_{l^*(\varepsilon_n)+1} < \varepsilon_n(\varepsilon)$, we choose on the first level

$$
M_0 = \mathrm{O}\Big(\frac{(\varepsilon_n(\varepsilon))^2}{h_L^2} \Big) = \mathrm{O}(2^{2L} (\varepsilon_n(\varepsilon))^2)
$$

to equilibrate the statistical and the discretization error contributions. The MLMC sample numbers M_l, for $l = 1, \ldots, l^*(\varepsilon_n), \ldots, L$, should be chosen according to

$$
(3.17) \qquad M_l = \mathrm{O}\Big(\frac{h_l^2 \wedge (\varepsilon_n(\varepsilon))^2}{h_L^2} l^{2(1+\eta)} \Big) = \mathrm{O}\Big(2^{2L} l^{2(1+\eta)} \big(2^{-2l} \wedge (\varepsilon_n(\varepsilon))^2 \big) \Big),
$$

for $\nu > 0$. Comparing these sample numbers with the corresponding sample numbers in [10], our analysis indicates that it suffices to sample less on under-resolved levels. The total work

for this estimates is the number of samples multiplied by the work for one solve on each level, summed over all levels. Therefore we achieve, for $L \in \mathbb{N}$, a total work for the MLMC FE method in the case that $h_L < \varepsilon_n(\varepsilon)$ of

$$\mathcal{W}(L) = \begin{cases} O\left(N_L^2 (\log_2 N_{l^*(\varepsilon_n)})^{3+2\eta}\right) & \text{for } d = 1, \\ O\left(N_L (\log_2 N_L)^{3+2\eta}\right) & \text{for } d = 2, \\ O\left(N_L (\log_2 N_L)^{2+2\eta}\right) & \text{for } d = 3. \end{cases}$$

Here, N_L denotes the degrees of freedom of the FE method on level L. For the corresponding singlelevel Monte Carlo method (see Equation (3.9)) the total work at level L is bounded by $\mathcal{W}(L) = O((\varepsilon_n(\varepsilon))^2 h_L^{-(2+d)})$ or equivalently by

$$\mathcal{W}(L) = \begin{cases} O\left(\varepsilon_n(\varepsilon)^2 N_L^3\right) & \text{for } d = 1, \\ O\left(\varepsilon_n(\varepsilon)^2 N_L^2\right) & \text{for } d = 2, \\ O\left(\varepsilon_n(\varepsilon)^2 N_L^{5/3}\right) & \text{for } d = 3. \end{cases}$$

We infer from the asymptotic error bound in Theorem 3.4 that convergence of discretization errors of the FE method can only be expected if $h_L < \varepsilon_n(\varepsilon)$ holds. Since $\varepsilon_n(\varepsilon)$ denotes the smallest physical length scale (cf. Equation (2.9)), $h_L < \varepsilon_n(\varepsilon)$ is a *scale-resolution requirement* for the FE method. Inspecting the singlelevel and the multilevel error bounds, Equation (3.9) and Equation (3.14), respectively, we see that scale-resolution is necessary for achieving $O(h_L)$ convergence rates. Further, Equation (3.14) indicates that ensemble averaging can, in part, compensate for lack of scale resolution in the multilevel MC method. Avoiding the preasymptotic range altogether, however, will require modifications in the discretization. In this case, i.e. when $h_L \geq \varepsilon_n(\varepsilon)$, we introduce and analyze next (following [28, 10]) the multilevel Monte Carlo Finite Element Heterogeneous Multiscale method.

4. Multilevel Monte Carlo Finite Element Heterogeneous Multiscale method

From now on and throughout the remainder of the present paper, we confine our analysis and development of the MLMC FE–HMM to problems with *one microscopic spatial length scale*, i.e. to $n = 1$ in Equation (2.2). Accordingly, we write from now on ε in place of $\varepsilon_1(\varepsilon)$ for the (nondimensional) physical length scale. We first give a short description of the FE–HMM for deterministic, elliptic, multiscale problems based on [1, 24] to the extent necessary for the exposition of the MLMC FE–HMM. The FE–HMM was introduced first in [23] and analyzed for elliptic problems in [1, 24] – we refer to [3, 4] for a review. We confine the ensuing analysis to first order FE–HMM but indicate that all results below have been established in recent years for deterministic elliptic and parabolic multiscale problems for macro and micro FE spaces of arbitrary order (see [4] for details).

To simplify the present exposition, we describe the method for piecewise linear FE spaces both, in the macroscopic as well as in the microscopic variable.

4.1. **FE–HMM.** The FE–HMM is based on macro and micro FE spaces. Where the macro problem is approximated with a hierarchy of macro FE methods, based on the family of standard FE spaces $\mathcal{V} = (V_l = \mathcal{S}_0^1(D, \mathcal{T}_l), l \in \mathbb{N}_0)$ as introduced in Section 3. The term "macro FE" is used here to underline that $h_l \gg \varepsilon$ is allowed and no scale resolution is imposed on these spaces.

Micro finite element space. For each $K \in \mathcal{T}_l$, $l \in \mathbb{N}_0$, we consider a node x_K located at the barycenter of the element K and a weight $\omega_K = |K|$ (the volume of the element K). Associated to each node x_K we consider a *micro domain* $K_\delta = x_K + \delta I$, where $I = (-1/2, 1/2)^d$ and, typically, $h_l >> \delta \geq \varepsilon$.

For a coefficient A^ε, given in Equation (2.11), which is periodic in the fast variable (with period ε), we can choose $\delta = \varepsilon$ and denote $K_\delta = K_\varepsilon$. The method still applies for $\delta > \varepsilon$ – for example in a situation where the period of the problem is a priori unknown– but a resonance error, e.g. δ/ε, appears in the FE–HMM convergence rate estimates in [24, 4] if $\delta/\varepsilon \notin \mathbb{N}$. For simplicity we adopt the choice $\varepsilon = \delta$ in what follows.

On each micro domain $K_\varepsilon \subset K \in \mathcal{T}_l$, $l \in N_0$, we consider a sequence $(\widetilde{\mathcal{T}}_l, l \in \mathbb{N}_0)$ of regular simplicial meshes of quasi-uniform triangles or tetrahedra. For any $l \in \mathbb{N}_0$, we denote the mesh width of $\widetilde{\mathcal{T}}_l$ by

$$\tilde{h}_l = \max_{T \in \widetilde{\mathcal{T}}_l}\{\operatorname{diam}(T)\} =: \max_{T \in \widetilde{\mathcal{T}}_l}\{\tilde{h}_T\}.$$

Then, we define the micro FE spaces on a micro domain K_ε by

$$(4.1) \qquad S_\#^1(K_\varepsilon, \widetilde{\mathcal{T}}_l) = \{\phi \in H_\#^1(K_\varepsilon)/\mathbb{R};\ \phi|_T, \in \mathcal{P}^1(T),\ T \in \widetilde{\mathcal{T}}_l\}.$$

Here, $H_\#^1(K_\varepsilon)$ denotes the set of all functions in $H_{loc}^1(\mathbb{R}^d)$ which belong to $H^1(K_\varepsilon)$ and are ε-periodic. The space $H_\#^1(K_\varepsilon)$ sets the coupling condition between macro and micro functions in the FE-HMM (see (4.2) below) [3]. As we choose for the macro problem piecewise linear functions for the FE spaces, it is natural (owing to the convergence rates in [1]) to choose piecewise linear FEs for the micro problems as well.

Micro problems. Next, we determine for every element $K \in \mathcal{T}_l$, $l \in \mathbb{N}_0$, the additive contribution to the macro stiffness matrix by computing the solutions of the following *micro problem*: For given $\boldsymbol{z} \in U$ and $v_l(\boldsymbol{z}; \cdot) \in V_l = \mathcal{S}_0^1(D, \mathcal{T}_l)$, $l \in \mathbb{N}_0$, and $K_\varepsilon \subset K \in \mathcal{T}_l$, find $v_{l,K_\varepsilon}(\boldsymbol{z}; \cdot) : K_\varepsilon \mapsto \mathbb{R}$ such that $v_{l,K_\varepsilon}(\boldsymbol{z}; \cdot) - v_l(\boldsymbol{z}; \cdot) \in S_\#^1(K_\varepsilon, \widetilde{\mathcal{T}}_l)$ and for all $\varphi_l \in S_\#^1(K_\varepsilon, \widetilde{\mathcal{T}}_l)$

$$(4.2) \qquad \int_{K_\varepsilon} A(\boldsymbol{z}; x_K, x/\varepsilon)\nabla v_{l,K_\varepsilon}(\boldsymbol{z}; x) \cdot \nabla\varphi_l(x)dx = 0.$$

We observe that Equation (4.2) can equivalently be formulated as: For a given $v_l(\boldsymbol{z}; \cdot) \in V_l$, $l \in \mathbb{N}_0$, find $\phi_{\tilde{l}}(\boldsymbol{z}; \cdot) \in S_\#^1(K_\varepsilon, \widetilde{\mathcal{T}}_l)$ such that for all $\varphi_l \in S_\#^1(K_\varepsilon, \widetilde{\mathcal{T}}_l)$

$$\int_{K_\varepsilon} A(\boldsymbol{z}; x_K, x/\varepsilon)\nabla\phi_{\tilde{l}}(\boldsymbol{z}; \cdot) \cdot \nabla\varphi_l(x)dx = -\int_{K_\varepsilon} A(\boldsymbol{z}; x_K, x/\varepsilon)\nabla v_l(\boldsymbol{z}; \cdot) \cdot \nabla\varphi_l(x)dx,$$

which has a unique solution due to the Lax-Milgram Theorem (the above bilinear form is coercive on $S_\#^1(K_\varepsilon, \widetilde{\mathcal{T}}_l) \subset H_\#^1(K_\varepsilon)/\mathbb{R}$ if this space is equipped with the norm $\|\nabla v\|_{L^2(K_\varepsilon)}$).

The macroscopic solution of the FE–HMM which aims at approximating the homogenized solution $\bar{u}$ of Problem (2.24) is defined as follows: Find $u_{l,HMM} \in L^2(U; V_l)$ such that for all $v_l \in L^2(U; V_l)$

$$(4.3) \qquad B_l^\varepsilon(u_{l,HMM}, v_l) = F(v_l)$$

where $B_l^\varepsilon(\cdot, \cdot) : L^2(U; V_l) \times L^2(U; V_l) \to \mathbb{R}$ is, for $0 < \varepsilon \leq 1$ and $v_l, w_l \in L^2(U; V_l)$, given by

$$(4.4) \qquad B_l^\varepsilon(w_l, v_l) := \int_U b_l(\varepsilon, \boldsymbol{z}; w_l(\boldsymbol{z}, \cdot), v_l(\boldsymbol{z}; \cdot))\rho(d\boldsymbol{z}).$$

---

[3] Other boundary conditions based e.g. on homogeneous Dirichlet coupling could be chosen for appropriate choices of the micro domain (e.g, K_δ instead of K_ε with $\delta > \varepsilon$) [24, 3].

Here, we set

$$(4.5) \qquad b_l(\varepsilon, \boldsymbol{z}; w_l, v_l) := \sum_{K \in \mathcal{T}_l} \frac{|K|}{|K_\varepsilon|} \int_{K_\varepsilon} A(\boldsymbol{z}; x_K, x/\varepsilon) \nabla w_{l,K_\varepsilon}(x) \cdot \nabla v_{l,K_\varepsilon}(x) \, dx,$$

where the factor $1/|K_\varepsilon|$ ensures the correct upscaling of the integrals on the micro element K_ε to the macroscopic Finite Element K.

We shall also need the *parametric, deterministic weak formulation* corresponding to the variational formulation in Equation (4.3): for $\boldsymbol{z} \in U$, and for $0 < \varepsilon \le 1$, find $u_{l,HMM}(\boldsymbol{z}; \cdot) \in V_l$, $l \in \mathbb{N}_0$, such that for all $v_l \in V_l$

$$(4.6) \qquad b_l(\varepsilon, \boldsymbol{z}; u_{l,HMM}(\boldsymbol{z}; \cdot), v_l) = \int_D f v_l \, dx.$$

Remark 4.1. Observe that in Equation (4.5) and in Equation (4.2) we have collocated the slow variable x at the barycenters x_K for $K \in \mathcal{T}_l$, $l \in \mathbb{N}_0$. As noticed in [1] and [6] this leads to robust convergence rates for the FE–HMM, i.e., convergence rates independent of ε.

To show that the problem in Equation (4.3) is well posed, we need the following lemma, whose proof follows closely the proof for the deterministic FE–HMM (see [1, 24],[3, Sect. 3.3.1]).

Lemma 4.2. *For ρ-a.e. $\boldsymbol{z} \in U$ and every $0 < \varepsilon \le 1$ consider $v_l(\boldsymbol{z}; \cdot) \in V_l$, $l \in \mathbb{N}_0$ and its corresponding micro solution $v_{l,K_\varepsilon}(\boldsymbol{z}; \cdot) \in S^1_\#(K_\varepsilon, \widetilde{\mathcal{T}_l})$, defined in Equation (4.2), holds*

$$(4.7) \qquad \|\nabla v_l(\boldsymbol{z}; \cdot)\|_{L^2(K_\varepsilon)} \le \|\nabla v_{l,K_\varepsilon}(\boldsymbol{z}; \cdot)\|_{L^2(K_\varepsilon)} \le \sqrt{\frac{\beta}{\alpha}} \|\nabla v_l(\boldsymbol{z}; \cdot)\|_{L^2(K_\varepsilon)}$$

where α, β are independent of ε and were defined in Assumption 2.1.

We notice that Lemma 4.2 is valid for higher order micro FE spaces such as $S^q_\#(K_\delta, \widetilde{\mathcal{T}})$, for $q \ge 1$. We have the following well-posedness result.

Theorem 4.3. *Under Assumption 2.1 and 2.2, for every $f \in L^2(D)$ the variational formulation, Equation (4.3), admits ρ-a.s. a unique solution $u_{l,\text{HMM}} \in L^2(U; V_l)$, $l \in \mathbb{N}_0$. This solution satisfies, for ρ-a.e. $\boldsymbol{z} \in U$ and for every $l \in \mathbb{N}_0$, the a-priori estimate*

$$(4.8) \qquad \|u_{l,\text{HMM}}(\boldsymbol{z}; \cdot)\|_V \le \frac{C(D, \beta)}{\alpha} \|f\|_{L^2(D)}.$$

Proof. Lemma 4.2 and the Cauchy–Schwarz inequality imply, that there exist constants $C_i(D, \alpha, \beta) > 0$, $i = 1, 2$, which are independent of ε and of l such that, for every $v_l, w_l \in L^2(U; V_l)$ (see [1, 24],[3, Sect. 3.3.1])

$$(4.9) \qquad |B_l^\varepsilon(v_l, w_l)| \le C_2 \|v_l\|_{L^2(U;V)} \|w_l\|_{L^2(U;V)},$$

and such that for every $v_l \in L^2(U; V_l)$

$$(4.10) \qquad B_l^\varepsilon(v_l, v_l) \ge C_1 \|v_l\|^2_{L^2(U;V)}.$$

Moreover, for every $w_l \in L^2(U; V_l)$, $F(w_l)$ is bounded similarly as in Equation (2.21). The assertion then follows from the Lax-Milgram Lemma. The explicit bound of order α^{-1} follows by inspection. □

For the error analysis of the FE–HMM, we also introduce the following variational formulation (which will not be used for actual computations). We define $\bar{\bar{u}}_{l,FEM} \in L^2(U; V_l)$, the solution of the FE method with numerical quadrature for the homogenized problem, i.e., find $\bar{\bar{u}}_{l,FEM} \in L^2(U; V_l)$ such that for all $v_l \in L^2(U; V_l)$

$$(4.11) \qquad \bar{B}_l(\bar{\bar{u}}_{l,FEM}, v_l) = F(v_l),$$

where the integrated bilinear form $\bar{B}_l(\cdot, \cdot) : L^2(U; V_l) \times L^2(U; V_l) \to \mathbb{R}$ is given by

$$(4.12) \qquad \bar{B}_l(w_l, v_l) := \int_U \bar{b}_l(\boldsymbol{z}; w_l(\boldsymbol{z}, \cdot), v_l(\boldsymbol{z}, \cdot)) \, \rho(d\boldsymbol{z}).$$

The *parametric, bilinear form with numerical integration* $b_l(\boldsymbol{z}; \cdot, \cdot)$ is for $v_l, w_l \in V_l$ defined by

$$(4.13) \qquad \bar{b}_l(\boldsymbol{z}; w_l, v_l) := \sum_{K \in \mathcal{T}_l} |K| \bar{A}(\boldsymbol{z}; x_K) \nabla w_l \cdot \nabla v_l.$$

Here, $\bar{A}(\boldsymbol{z}; x)$ denotes the parametric, deterministic homogenized tensor, $|K|$ the measure of the element K with barycenter x_K.

As before, the solution $\bar{\bar{u}}_{l,FEM} \in L^2(U; V_l)$ of Equation (4.11) coincides with the parametric, deterministic multiscale solution, defined by: find $\bar{\bar{u}}_{l,FEM}(\boldsymbol{z}; \cdot) \in V_l$ such that (pointwise, for ρ-a.e. $\boldsymbol{z} \in U$) it holds that for all $v_l \in V_l$

$$(4.14) \qquad \bar{b}_l(\boldsymbol{z}; \bar{\bar{u}}_{l,FEM}(\boldsymbol{z}; \cdot), v_l) = \int_D f v_l dx.$$

Next, we establish error estimates for the FE–HMM.

Lemma 4.4. *Under Assumption 2.1 and 2.2, for ρ-a.e. $\boldsymbol{z} \in U$ the solutions $\bar{u}(\boldsymbol{z}; \cdot)$ and $u_{l,\mathrm{HMM}}(\boldsymbol{z}; \cdot)$, $l \in \mathbb{N}_0$, of the variational formulations given in Equation (2.23) and Equation (4.6), respectively, satisfy*

$$(4.15) \qquad \begin{aligned} \|\bar{u}(\boldsymbol{z}; \cdot) &- u_{l,\mathrm{HMM}}(\boldsymbol{z}; \cdot)\|_V \\ &\leq \|\bar{u}(\boldsymbol{z}; \cdot) - \bar{\bar{u}}_{l,FEM}(\boldsymbol{z}; \cdot)\|_V + C \sup_{v_l \in V_l} \frac{|b_l(\varepsilon, \boldsymbol{z}; u_{l,\mathrm{HMM}}, v_l) - \bar{b}_l(\boldsymbol{z}; u_{l,\mathrm{HMM}}, v_l)|}{\|v_l\|_V}, \end{aligned}$$

where $\bar{\bar{u}}_{l,FEM}(\boldsymbol{z}; \cdot)$ is the solution of the variational formulation given in Equation (4.14) and the constant $C > 0$ depends on α and β, but is independent of $l \in \mathbb{N}_0$ and ε.

Proof. The proof follows with the usual Strang-type argument from the triangle inequality

$$\|\bar{u}(\boldsymbol{z}; \cdot) - u_{l,HMM}(\boldsymbol{z}; \cdot)\|_V \leq \|\bar{u}(\boldsymbol{z}; \cdot) - \bar{\bar{u}}_{l,FEM}(\boldsymbol{z}; \cdot)\|_V + \|\bar{\bar{u}}_{l,FEM}(\boldsymbol{z}; \cdot) - u_{l,HMM}(\boldsymbol{z}; \cdot)\|_V,$$

and the inequality

$$\begin{aligned} \|\bar{\bar{u}}_{l,FEM}(\boldsymbol{z}; \cdot) - u_{l,HMM}(\boldsymbol{z}; \cdot)\|_V^2 \leq C \, \big| b_l\big(\varepsilon, \boldsymbol{z}; u_{l,HMM}(\boldsymbol{z}; \cdot), \bar{\bar{u}}_{l,FEM}(\boldsymbol{z}; \cdot) - u_{l,HMM}(\boldsymbol{z}; \cdot)\big) \\ - \bar{b}_l\big(\boldsymbol{z}; u_{l,HMM}(\boldsymbol{z}; \cdot), \bar{\bar{u}}_{l,FEM}(\boldsymbol{z}; \cdot) - u_{l,HMM}(\boldsymbol{z}; \cdot)\big) \big|. \end{aligned}$$

$\square$

To obtain convergence rates, we impose additional regularity assumptions on the homogenized and the micro problems.

Assumption 4.5. The homogenized coefficients satisfy, for ρ-a.e. $\boldsymbol{z} \in U$, $\bar{A}(\boldsymbol{z}; \cdot) \in C^1(\bar{D})^{d \times d}$ and $\|\bar{A}(\boldsymbol{z}; \cdot)\|_{C^1(\bar{D})^{d \times d}}$ is uniformly bounded. Further, for ρ-a.e. $\boldsymbol{z} \in U$, $\bar{u}(\boldsymbol{z}; \cdot) \in H^2(D)$ and $\|\bar{u}(\boldsymbol{z}; \cdot)\|_{H^2(D)}$ (see Remark 2.6) are uniformly bounded.

We note that if we assume that the domain D is convex and that $A(\boldsymbol{z}; x, \boldsymbol{y})$ has strong derivatives of order 2 with respect to x and y_j that are uniformly bounded for all $\boldsymbol{z} \in U$, Assumption 4.5 can be shown to hold [31, Proposition 4.2].

Regularity on the solution the following cell problem is also needed: For $\boldsymbol{z} \in U$ and $0 < \varepsilon \leq 1$ find $\psi_{K_\varepsilon}^j(\boldsymbol{z}; \cdot) \in H_\#^1(K_\varepsilon)/\mathbb{R}$, for $j = 1, \ldots, d$ such that for all $\varphi \in H_\#^1(K_\varepsilon)/\mathbb{R}$

$$(4.16) \qquad \int_{K_\varepsilon} A(\boldsymbol{z}; x_K, x/\varepsilon) \nabla \psi_{K_\varepsilon}^j(x) \cdot \nabla \varphi(x)\, dx = -\int_{K_\delta} A(\boldsymbol{z}; x_K, x/\varepsilon)\, \mathbf{e}_j(x) \cdot \nabla \varphi(x)\, dx,$$

where $\mathbf{e}_j$ is the j-th vector of the canonical basis of $\mathbb{R}^d$ and $x_K \in K$ is a quadrature node (located at the barycenter of K). The solution of Equation (4.16), $\psi_{K_\varepsilon}^j(\boldsymbol{z}; x_K, x/\varepsilon)$, is given by $\varepsilon \hat{\psi}^j(z; x_K, x/\varepsilon)$, where $\hat{\psi}^j(\boldsymbol{z}; x_K, y) \in H_\#^1(Y)/\mathbb{R}$ is the solution of the cell problem given by

$$(4.17) \qquad \int_Y A(\boldsymbol{z}; x_K, y) \nabla \hat{\psi}^j(y) \cdot \nabla \hat{\varphi}(y)\, dy = -\int_Y A(\boldsymbol{z}; x_K, y)\, \mathbf{e_j}(y) \cdot \nabla \hat{\varphi}(y)\, dy,$$

for all $\hat{\varphi} \in H_\#^1(Y)/\mathbb{R}$.

Assumption 4.6. Assume that, for ρ-a.e. $\boldsymbol{z} \in U$, the solution $\psi_{K_\varepsilon}^j(\boldsymbol{z}; \cdot)$ of the variational formulation given in Equation (4.16) satisfies $\psi_{K_\varepsilon}^j(\boldsymbol{z}; \cdot) \in H^2(K_\varepsilon)$, and that $\|\psi_{K_\varepsilon}^j(\boldsymbol{z}; \cdot)\|_{H^2(K_\varepsilon)}$ is uniformly bounded and that

$$(4.18) \qquad |\psi_{K_\varepsilon}^j(\boldsymbol{z}; x_K, \cdot)|_{H^2(K_\varepsilon)} \leq C\varepsilon^{-1} \sqrt{|K_\varepsilon|}.$$

We note that if we assume that $A(\boldsymbol{z}; x_K, \cdot) \in W^{1,\infty}(K_\varepsilon)$, then classical regularity results ensure that Equation (4.18) holds (see [5, Remark 5.1]).

The following result has been obtained in [1] (see also [3]) for deterministic problems.

Theorem 4.7. *Under Assumption 2.1, 2.2, 4.5 and 4.6 for ρ-a.e. $\boldsymbol{z} \in U$ the difference of the solutions $\bar{u}(\boldsymbol{z}; \cdot), u_{l,\mathrm{HMM}}(\boldsymbol{z}; \cdot), l \in \mathbb{N}_0$, of the variational formulations given in Equation (2.23) and Equation (4.6), respectively, satisfies*

$$(4.19) \qquad \|\bar{u}(\boldsymbol{z}; \cdot) - u_{l,\mathrm{HMM}}(\boldsymbol{z}; \cdot)\|_V \leq C \left(h_l + \left(\frac{\tilde{h}_l}{\varepsilon} \right)^2 \right) \|f\|_{L^2(D)},$$

where the constant C is independent of h_l, $\tilde{h}_l$, ε.

Proof. We use Lemma 4.4 and estimate both term in the right-hand side of Equation (4.15). Using Assumption 2.1, Assumption 4.5, and classical estimates for the FE method with numerical quadrature (see [16]) we may write

$$(4.20) \qquad \|\bar{u}(\boldsymbol{z}; \cdot) - \bar{\bar{u}}_{l,FEM}(\varepsilon, \boldsymbol{z}; \cdot)\|_V \leq Ch_l \|f\|_{L^2(D)}.$$

for the first term. Following the proof of [1, Lemma 3.3] (see also [5, Lemma 5.2]), using Assumptions 2.1 and Assumption 4.6, the second term of the right-hand side of Equation (4.15) is bounded by $C \left(\frac{\tilde{h}_l}{\varepsilon} \right)^2$. $\qquad \square$

We emphasize that $\frac{\tilde{h}_l}{\varepsilon}$ is independent of ε. Indeed, in terms of degrees of freedom of the micro FE method $N_{l,mic}$ we have $\tilde{h}_l = \frac{\varepsilon}{N_{l,mic}^{1/d}}$ and define

$$(4.21) \qquad \hat{h}_l := \frac{\tilde{h}_l}{\varepsilon} = \frac{1}{N_{l,mic}^{1/d}}.$$

Of interest is also the following post-processing procedure to approximate the fine scale solution u^ε. Let $u_{l,HMM}$ be the solution of Equation (4.3) and u_{l,K_ε} the solution of the variational formulation of the microproblem, given in Equation (4.2). We then consider the function $(u_{l,HMM} - u_{l,K_\varepsilon})(\boldsymbol{z};\cdot)$ available on each micro domain $K_\varepsilon \subset K$ centered in x_K, and the *gradient recovery approximation* $\mathcal{R}_l^\varepsilon$ of the solution $u_{l,HMM}$ defined, for all $x \in K \in \mathcal{T}_l$, by

$$(4.22) \qquad \mathcal{R}_l^\varepsilon u_{l,HMM}(\boldsymbol{z};x)|_K := u_{l,HMM} + (u_{l,HMM} - u_{l,K_\varepsilon})(\boldsymbol{z};x - [x]_{K_\varepsilon}),$$

where $[x]_{K_\varepsilon} \in \varepsilon \mathbb{Z}^d$ is such that $x - [x]_{K_\varepsilon} \in K_\varepsilon$. We note that

$$\mathcal{R}_l^\varepsilon u_{l,HMM} \in L^2(U; H^1(\mathcal{T}_l)),$$

where $H^1(\mathcal{T}_l)$ denotes the broken Sobolev space which is endowed with the norm

$$\|v\|_{H^1(\mathcal{T}_l)} := \left(\|v\|_{L^2(D)}^2 + \sum_{K \in \mathcal{T}_l} \|\nabla v\|_{L^2(K)}^2 \right)^{1/2}.$$

Theorem 4.8. *Let all assumptions of Theorem (4.7) be satisfied. Further, assume that $A \in L^\infty(U; C^1(\bar{D}; C_\#^1(Y)))$, $\bar{u} \in L^\infty(U, H^2(D))$, and that the solution $\hat{\psi}^j$ of Equation (4.17) satisfies $\hat{\psi}^j \in L^\infty(U; C^1(\bar{D}; H_\#^2(Y))) \cap L^\infty(U \times \bar{D}; C^1(\bar{Y}))$. Then, for ρ-a.e. $\boldsymbol{z} \in U$ and $0 < \varepsilon \leq 1$ holds the error bound*

$$\|u^\varepsilon(\boldsymbol{z};\cdot) - \mathcal{R}_l^\varepsilon u_{l,\mathrm{HMM}}(\boldsymbol{z};\cdot)\|_{H^1(\mathcal{T}_l)} \leq C \left(h_l + \hat{h}_l + \sqrt{\varepsilon} \right) \|f\|_{L^2(D)},$$

where the constant C is independent of $h_l, \tilde{h}_l$, and ε.

Proof. The proof follows from [1, Corollary 3.7, Theorem 3.11] and [31, Proposition 5.1]. $\square$

4.2. **Monte Carlo Finite Element Heterogeneous Multiscale Method.** We proceed with an analysis of the rate of convergence of the Monte Carlo method for the solution of the stochastic elliptic problem, given in Equation (2.16). The implementation of the estimator $E_M[u]$ in Equation (3.1) requires a Finite Element approximation of the copies u^i of u. We choose the Finite Element approximation to be the FE–HMM and we establish a first error estimate for the MC FE–HMM in the case when *the same* discretization $\mathcal{T}_l$, for $l \in \mathbb{N}_0$, is used for all samples. We estimate the expectation of the solution, $\mathbb{E}[\bar{u}]$, by

$$E_M[u_{l,HMM}] := \frac{1}{M} \sum_{i=1}^M u_{l,HMM}^i,$$

where $u_{l,HMM}^i$ are independent copies of the random solution $u_{l,HMM}$ of Equation (4.3). We further estimate $\mathbb{E}[u^\varepsilon]$ by

$$E_M[\mathcal{R}_l^\varepsilon u_{l,HMM}] := \frac{1}{M} \sum_{i=1}^M \mathcal{R}_l^\varepsilon u_{l,HMM}^i,$$

where $\mathcal{R}_l^\varepsilon u_{l,HMM}^i$ are independent copies of the random solution $\mathcal{R}_l^\varepsilon u_{l,HMM}$ defined in Equation (4.22).

Theorem 4.9. *Let all assumptions of Theorem 4.7 be satisfied. Then, it holds, for any $l \in \mathbb{N}_0$ and $M \in \mathbb{N}$, that*

$$(4.23) \qquad \|\mathbb{E}[\bar{u}] - E_M[u_{l,\mathrm{HMM}}]\|_{L^2(U;V)} \le C\left(\frac{1}{\sqrt{M}} + (h_l + \hat{h}_l^2)\right)\|f\|_{L^2(D)}.$$

If the assumptions of Theorem 4.8 are satisfied, it holds that

$$(4.24) \qquad \|\mathbb{E}[u^\varepsilon] - E_M[\mathcal{R}_l^\varepsilon u_{l,\mathrm{HMM}}]\|_{L^2(U;H^1(\mathcal{T}_l))} \le C\left(\frac{1}{\sqrt{M}} + (h_l + \hat{h}_l + \sqrt{\varepsilon})\right)\|f\|_{L^2(D)}.$$

In both estimates, the constant C is independent of h_l, $\hat{h}_l$, M and ε (where $\hat{h}_l$ is defined in Equation (4.21)).

Proof. We first prove Equation (4.23). To this end, we split the left hand side into

$$\|\mathbb{E}[\bar{u}] - E_M[u_{l,HMM}]\|_{L^2(U;V)} \le \|\mathbb{E}[\bar{u}] - \mathbb{E}[u_{l,HMM}]\|_{L^2(U;V)}$$
$$+ \|\mathbb{E}[u_{l,HMM}] - E_M[u_{l,HMM}]\|_{L^2(U;V)}.$$

The first term on the right hand side is bounded by Theorem 4.7 (since the bound is independent of $\boldsymbol{z} \in U$). The assertion follows with Lemma 3.1 and Theorem 4.3 for the second term. A similar argument shows the second assertion: by referring to Theorem 4.8 to bound the first term on the right hand side of the corresponding inequality and, finally, using the rescaling in Equation (4.21) for the error bound of Theorem 4.8, we obtain the stated result. $\qquad\square$

We notice that in the proof above, we need a bound on $\|\mathcal{R}_l^\varepsilon u_{l,HMM}\|_{L^2(U;V)}$. This follows from the representation

$$\mathcal{R}_l^\varepsilon u_{l,HMM}(\boldsymbol{z};\cdot) = u_{l,HMM}(\boldsymbol{z};\cdot) + \sum_{j=1}^d \psi_{K_\varepsilon}^j(\boldsymbol{z};\cdot)\frac{\partial u_{l,HMM}}{\partial x_j}(\boldsymbol{z};\cdot)$$

for $\mathcal{R}_l^\varepsilon u_{l,HMM}$ (see for example [3, Proposition 8]), the triangle inequality, Theorem 4.3 and Assumption 4.6. The optimal choice of sample size versus grid size is reached when the statistical and the discretization errors are equilibrated. Assume a uniform macro and micro mesh refinement for the FE–HMM with mesh width $h_l = 2^{-l}h_0$ (referred to as "macro mesh") and $\hat{h}_l = 2^{-l}\hat{h}_0$ (referred to "micro mesh") with $h_0, \hat{h}_0 = \mathrm{O}(1)$ (c.f. Equation (4.21)). In view of Equation (3.8) and Equation (4.24) the optimal choice of sample size versus grid size is given by $M^{-\frac{1}{2}} = \mathrm{O}(h_l) = \mathrm{O}(2^{-l})$ and $\hat{h}_l = \sqrt{h_l} = \mathrm{O}(2^{-l/2})$ for estimating the expectation of the solution $\mathbb{E}[\bar{u}]$ and $\hat{h}_l = h_l = \mathrm{O}(2^{-l})$ for estimating the expectation of the solution $\mathbb{E}[u^\varepsilon]$ (in the case that $\sqrt{\varepsilon} \le h_l$). In terms of the degrees of freedom of the FE–HMM, denoting by $N_{l,mac}$ the degrees of freedom of the macro FE space and by $N_{l,mic}$ the degrees of freedom of the micro FE space we obtain the basic relations

$$(4.25) \qquad M^{-\frac{1}{2}} = \mathrm{O}(N_{l,mac}^{-\frac{1}{d}}),$$

and for the micro degrees of freedom $N_{l,mic}$

$$(4.26) \qquad N_{l,mic} = \sqrt{N_{l,mac}},$$

for estimating $\mathbb{E}[\bar{u}]$ and

$$(4.27) \qquad N_{l,mic} = N_{l,mac},$$

for estimating $\mathbb{E}[u^\varepsilon]$. The total computational cost (in terms of the macro degrees of freedom and of the number of samples M) of the estimates $E_M[u_{l,HMM}]$ and $E_M[\mathcal{R}_l^\varepsilon u_{l,HMM}]$ is $\mathrm{O}(M \cdot N_{l,mac}^{3/2})$ and $\mathrm{O}(M \cdot N_{l,mac}^2)$.

With the previous calculation on the optimal sample size, i.e., Equation (4.25), Equation (4.26) and Equation (4.27), we may write for the computational cost $\mathrm{O}(2^{l(\frac{3}{2}d+2)})$ and $\mathrm{O}(2^{l(2d+2)})$, for computing $E_M[u_{l,HMM}]$ and $E_M[\mathcal{R}_l^\varepsilon u_{l,HMM}]$, respectively. We emphasize that the work estimates are independent of ε while the error estimates hold for arbitrary mesh sizes $h_l = 2^{-l}h_0$ that do not need to resolve the fine scale ε. For computing $E_M[\mathcal{R}_l^\varepsilon u_{l,HMM}]$ we notice however that the accuracy depends on ε (see Equation (4.24)).

We note at this point that MC FE–HMM with recovery is only applicable in the case where $\sqrt{\varepsilon} \leq h_l$ (for convergence of $\mathrm{O}(h_l)$), whereas the singlelevel Monte Carlo method requires scale resolution, i.e., $h_l < \varepsilon$. For a fixed ε the sets $\{h_l : h_l < \varepsilon\}$ and $\{h_l : h_l \geq \sqrt{\varepsilon}\}$ are disjoint and therefore, a comparison of the singlelevel MC FE–HMM with recovery and the (scale resolved) singlelevel Monte Carlo FE method is not viable. However, a Monte Carlo approximation with standard FE methods, would require a total work of the order $\mathrm{O}(\varepsilon^{-d}2^{l(2+d)})$ for an accuracy of $\mathrm{O}(h_l)$ (with the choice of $h_l = \varepsilon\, 2^{-l}$ in Equation (3.9)). Subsequently we generalize these calculations to the case of a multilevel Monte Carlo approximation of the expectation.

4.3. Multilevel Monte Carlo FE–HMM.

In the MLMC FE–HMM, we consider the hierarchic sequence $\mathcal{V} = (V_l = \mathcal{S}_0^1(D, \mathcal{T}_l), l \in \mathbb{N}_0)$ of finite dimensional sub–spaces of V with $V_0 \subset V_1 \subset \ldots \subset V_l \subset \ldots \subset V$. With the convention $u_{-1,HMM} \equiv 0$ we write for $L \in \mathbb{N}_0$

$$u_{L,HMM} = \sum_{l=0}^{L}(u_{l,HMM} - u_{l-1,HMM})$$

and, by linearity of the expectation,

$$(4.28) \qquad \mathbb{E}[u_{L,HMM}] = \sum_{l=0}^{L}\mathbb{E}[u_{l,HMM} - u_{l-1,HMM}].$$

In the MLMC FE–HMM, we estimate $\mathbb{E}[u_{l,HMM} - u_{l-1,HMM}]$ by a level dependent number M_l of samples, which implies that we may estimate $\mathbb{E}[\bar{u}]$ by

$$(4.29) \qquad E^L[u_{L,HMM}] := \sum_{l=0}^{L} E_{M_l}[u_{l,HMM} - u_{l-1,HMM}],$$

and $\mathbb{E}[u^\varepsilon]$ by the *gradient recovered MLMC FE–HMM* approximation defined by

$$(4.30) \qquad E^L[\mathcal{R}_L^\varepsilon u_{L,HMM}] := \sum_{l=0}^{L} E_{M_l}[\mathcal{R}_l^\varepsilon u_{l,HMM} - \mathcal{R}_{l-1}^\varepsilon u_{l-1,HMM}].$$

Here, the mesh- and ε-dependent recovery operators $\mathcal{R}_l^\varepsilon$, $l \in \mathbb{N}_0$, are defined in Equation (4.22).

The convergence of the MLMC–FE–HMM method is guaranteed by the following.

Lemma 4.10. *Let the assumptions of Theorem 4.7 be satisfied. Then, the MLMC FE–HMM approximations in Equation (4.29) of the expectation $\mathbb{E}[\bar{u}]$ of the solution $\bar{u}$ to the stochastic*

elliptic boundary value problem, presented in Equation (2.23), admit, for $L \in \mathbb{N}_0$, the error bound

$$\|\mathbb{E}[\bar{u}] - E^L[u_{L,\mathrm{HMM}}]\|_{L^2(U;V)}$$

(4.31)
$$\leq C \left(h_L + \hat{h}_L^2 + M_0^{-1/2} + \sum_{l=1}^{L} (h_l + \hat{h}_l^2) M_l^{-1/2} \right) \|f\|_{L^2(D)}.$$

Likewise, if the assumptions of Theorem 4.8 are satisfied, then the MLMC–FE–HMM approximations in Equation (4.30) of the expectation $\mathbb{E}[u^\varepsilon]$ of the solution u^ε presented in Equation (2.13) admits, for $L \in \mathbb{N}_0$, the error bound

$$\|\mathbb{E}[u^\varepsilon] - E^L[\mathcal{R}_L^\varepsilon u_{L,\mathrm{HMM}}]\|_{L^2(U;H^1(\mathcal{T}_L))}$$

(4.32)
$$\leq C \left(h_L + \hat{h}_L + \sqrt{\varepsilon} + M_0^{-1/2} + \sum_{l=1}^{L} (h_l + \hat{h}_l + \sqrt{\varepsilon}) M_l^{-1/2} \right) \|f\|_{L^2(D)}.$$

In either of the error bounds, the constant C is independent of h_l, for $l = 1, \ldots, L$, and ε.

Proof. We start by proving the first assertion and rewrite $\|\mathbb{E}[\bar{u}] - E^L[u_{L,\mathrm{HMM}}]\|_{L^2(U;V)}$ as

$$\|\mathbb{E}[\bar{u}] - E^L[u_{L,\mathrm{HMM}}]\|_{L^2(U;V)}$$

$$= \|\mathbb{E}[\bar{u}] - \mathbb{E}[u_{L,\mathrm{HMM}}] + \mathbb{E}[u_{L,\mathrm{HMM}}] - \sum_{l=0}^{L} E_{M_l}[u_{l,\mathrm{HMM}} - u_{l-1,\mathrm{HMM}}]\|_{L^2(U;V)}$$

$$\leq \|\mathbb{E}[\bar{u}] - \mathbb{E}[u_{L,\mathrm{HMM}}]\|_{L^2(U;V)}$$

$$\quad + \|\sum_{l=0}^{L} \left(\mathbb{E}[u_{l,\mathrm{HMM}} - u_{l-1,\mathrm{HMM}}] - E_{M_l}[u_{l,\mathrm{HMM}} - u_{l-1,\mathrm{HMM}}] \right)\|_{L^2(U;V)}$$

$$=: I + II.$$

Here, we adopted the convention that $u_{-1,\mathrm{HMM}} \equiv 0$. We calculate the error bounds for the terms I and II separately.

Term I: By Jensen's and the Cauchy-Schwarz inequality, for $L \in \mathbb{N}_0$, we get

$$I = \|\mathbb{E}[\bar{u} - u_{L,\mathrm{HMM}}]\|_{L^2(U;V)} \leq \|\bar{u} - u_{L,\mathrm{HMM}}\|_{L^2(U;V)} \leq C \left(h_L + \hat{h}_L^2 \right) \|f\|_{L^2(D)},$$

where we used Theorem 4.7 (the bound is independent of $\boldsymbol{z} \in U$) in the last inequality.

For Term II we use the triangle inequality and Lemma 3.1 to get

$$II \leq \sum_{l=0}^{L} \|(\mathbb{E} - E_{M_l})[u_{l,\mathrm{HMM}} - u_{l-1,\mathrm{HMM}}]\|_{L^2(U;V)}$$

$$\leq \sum_{l=0}^{L} \frac{1}{\sqrt{M_l}} \|u_{l,\mathrm{HMM}} - u_{l-1,\mathrm{HMM}}\|_{L^2(U;V)}.$$

For the first term of the sum ($l = 0$) we have with Theorem 4.3, and the convention $u_{-1,\mathrm{HMM}} \equiv 0$ the bound

$$\frac{1}{\sqrt{M_0}} \|u_{0,\mathrm{HMM}}\|_{L^2(U;V)} \leq C \frac{1}{\sqrt{M_0}} \|f\|_{L^2(D)}.$$

All the other terms in the sum ($l = 1, \ldots, L$) we bound with the triangle inequality and Theorem 4.7 as

$$
\frac{1}{\sqrt{M_l}} \|u_{l,HMM} - u_{l-1,HMM}\|_{L^2(U;V)} \leq \frac{1}{\sqrt{M_l}} \left(\|u_{l,HMM} - \bar{u}\|_{L^2(U;V)} + \|\bar{u} - u_{l-1,HMM}\|_{L^2(U;V)} \right)
$$

$$
\leq C \frac{1}{\sqrt{M_l}} (h_l + \hat{h}_l^2 + h_{l-1} + \hat{h}_{l-1}^2) \|f\|_{L^2(D)}
$$

$$
\leq C \frac{1}{\sqrt{M_l}} (h_l + \hat{h}_l^2 + 2h_l + 4\hat{h}_l^2) \|f\|_{L^2(D)}
$$

$$
\leq C \frac{1}{\sqrt{M_l}} (h_l + \hat{h}_l^2) \|f\|_{L^2(D)} .
$$

Here, we used that $h_l = 2^{-1} h_{l-1} = 2^{-l} h_0$ and $\hat{h}_l = 2^{-1} \hat{h}_{l-1} = 2^{-l} \hat{h}_0$ with $h_0, \hat{h}_0 = O(1)$. Summing these estimates from $l = 1, ..., L$ and combining the bounds for term I and term II completes the proof.

The second assertion can be obtained similarly using Theorem 4.8 instead of Theorem 4.7.
$\square$

The preceding result gives an error bound for the MLMC FE–HMM approximation, for *any* distribution of samples ($M_l, l \in \mathbb{N}_0$) over the mesh levels. Like in the singlelevel Monte Carlo approximation one is interested in the optimal ratio of sample size versus grid size in every level, i.e., how M_l relates to h_l to achieve an overall convergence rate of $O(h_L)$, for $L \in \mathbb{N}_0$ and $l = 0, \ldots, L$.

Theorem 4.11. *Assume that the spaces ($V_l, l \in \mathbb{N}_0$) have mesh width $h_l = 2^{-l} h_0$, where h_0 denotes the mesh width of the coarsest triangulation. Further, assume that the mesh width of the micro FE spaces is $\hat{h}_l = \sqrt{h_l}$, for all $l \in \mathbb{N}_0$. Then, for $L \in \mathbb{N}_0$, the MLMC FE–HMM approximation given in Equation* (4.29) *of the expectation of the solution of the stochastic elliptic boundary value problem given in Equation* (2.23) *with ($M_l, l = 0, \ldots, L$) samples given by*

$$
M_0 = O(2^{2L} h_0) \quad and \quad M_l = O(2^{2(L-l)} h_0 \, l^{2+2\eta}) \; for \; l = 1, 2, ..., L,
$$

where $\eta > 0$ is arbitrarily small, admits the error bound

(4.33) $$\|\mathbb{E}[\bar{u}] - E^L[u_{L,\mathrm{HMM}}]\|_{L^2(U;V)} \leq C h_L \|f\|_{L^2(D)}.$$

If, at each level $l = 0, \ldots, L$ the equations for each sample of $u_{l,\mathrm{HMM}}$ in the estimator $E_M[u_{l,\mathrm{HMM}}]$ are solved approximately with a full Multigrid method to accuracy $O(h_l)$ in the energy norm, the total work $\mathcal{W}(L)$ for computing $\mathbb{E}[\bar{u}]$ approximately to accuracy $O(h_L)$ is bounded by

$$
\mathcal{W}(L) = \begin{cases} O\left((N_{L,mac})^2\right) & for \; d = 1 \\ O\left((N_{L,mac})^{\frac{3}{2}} (\log_2 N_{L,mac}/d)^{2+2\eta}\right) & for \; d = 2, 3, \end{cases}
$$

where the constant C depends on η and on d but is independent of L and ε.

Proof. We fix $L \in \mathbb{N}_0$. All constants C in this proof are independent of ε and of L. With this convention, with the choice $M_0 = O(h_L^{-2}) = O(2^{2L} h_0)$ and with the choice

(4.34) $$M_l = O((h_l/h_L)^2 l^{2+2\eta}) = O(2^{2(L-l)} l^{2+2\eta}) \; for \; l = 1, \ldots, L,$$

for some $\eta > 0$, we obtain from Equation (4.31), Equation (4.32) the asserted error bound in Equation (4.33).

Indeed, for $\eta > 0$, Equation (4.31) yields

$$\sum_{l=0}^{L}(h_l + \hat{h}_l^2)M_l^{-1/2} \leq C\sum_{l=0}^{L} h_l M_l^{-1/2} \leq C\sum_{l=1}^{L} h_l \frac{h_L}{h_l} l^{-(1+\eta)} \leq C(\eta)h_L$$

and

$$M_0^{-1/2} \leq Ch_L.$$

To estimate the work, we observe that the approximate solution of $\mathbb{E}[\bar{u}]$ given by the FE–HMM at mesh level $l = 0, \ldots, L$ to accuracy h_l is of complexity $\mathrm{O}(N_{l,mac}^{3/2})$, in the number of degrees of freedom $N_{l,mac}$ at mesh level l (since by $\hat{h}_l = \sqrt{h_l}$ we have $N_{l,mic} = N_{l,mac}^{1/2}$). For M_l samples (possibly in parallel) this requires a total of $\mathrm{O}(M_l N_{l,mac}^{3/2})$ computational work. For the approximation of $\mathbb{E}[\bar{u}]$, this amounts to the following bound for the overall work for the MLMC FE–HMM at level L

$$\mathcal{W}(L) \lesssim \sum_{l=0}^{L} M_l (N_{l,mac})^{3/2} \leq C\left(h_L^{-2} + \sum_{l=1}^{L}(\frac{h_l}{h_L})^2 l^{2+2\eta} h_l^{-\frac{3}{2}d}\right)$$

$$\leq C h_L^{-2}\left(1 + \sum_{l=1}^{L} l^{2(1+\eta)} h_l^{2-\frac{3}{2}d}\right)$$

In the case $d = 1$ the overall computational complexity is bounded by $\mathcal{W}(L) = \mathrm{O}(N_{L,mac}^2)$, since then $\sum_{l=1}^{L} l^{2(1+\eta)} h_l^{\frac{1}{2}}$ is bounded independently of L. In the case $d = 2, 3$ we have

$$\mathcal{W}(L) \leq C h_L^{-2}\left(1 + h_L^{2-\frac{3}{2}d} \sum_{l=1}^{L} l^{2(1+\eta)} \left(\frac{h_l}{h_L}\right)^{2-\frac{3}{2}d}\right)$$

$$\leq C h_L^{-\frac{3}{2}d}\left(1 + L^{2(1+\eta)} \sum_{l=1}^{L} \left(\frac{h_L}{h_l}\right)^{\frac{3}{2}d-2}\right)$$

$$\leq C h_L^{-\frac{3}{2}d} L^{2(1+\eta)},$$

where we used that the sum $\sum_{l=1}^{L} \left(\frac{h_L}{h_l}\right)^{\frac{3}{2}d-2}$ is bounded by $2^{\frac{3}{2}d-2}$ for $h_L/h_l = 2^{-(L-l)}$. As $L = \log_2(N_{L,mac})/d$ we obtain for $d = 1, 2, 3$ the asserted work estimates. $\qquad \square$

We have a similar result for the approximation of $\mathbb{E}[u^\varepsilon]$.

Theorem 4.12. *Let the assumptions of Theorem 4.11 be satisfied. Assume that the mesh width of the micro FE spaces is $\hat{h}_l = h_l$, for all $l \in \mathbb{N}_0$. Then, for $L \in \mathbb{N}_0$, the gradient recovered MLMC FE–HMM approximation given in Equation (4.30) of the expectation of the solution of the stochastic elliptic boundary value problem given in Equation (2.13) with $(M_l, l = 0, \ldots, L)$ samples given by*

$$M_0 = \mathrm{O}(2^{2L}h_0) \quad and \quad M_l = \mathrm{O}(2^{2(L-l)}h_0 \, l^{2+2\eta}) \text{ for } l = 1, 2, ..., L,$$

where $\eta > 0$ is arbitrarily small, admit the error bound

(4.35) $$\|\mathbb{E}[u^\varepsilon] - E^L[\mathcal{R}_L^\varepsilon u_{L,\mathrm{HMM}}]\|_{L^2(U;H^1(\mathcal{T}_L))} \leq C(h_L + \sqrt{\varepsilon})\|f\|_{L^2(D)}.$$

The total work $\mathcal{W}(L)$ for computing $\mathbb{E}[u^\varepsilon]$ approximately to accuracy $\mathrm{O}(h_L)$ is bounded by

$$(4.36) \qquad \mathcal{W}(L) = \begin{cases} \mathrm{O}\left((N_{L,mac})^2(\log_2 N_{L,mac})^{3+2\eta}\right) & \text{for } d = 1 \\ \mathrm{O}\left((N_{L,mac})^2(\log_2 N_{L,mac}/d)^{2+2\eta}\right) & \text{for } d = 2, 3, \end{cases}$$

where the constant C depends on η and d but is independent of L and ε.

Proof. The proof is nearly identical to the proof of Theorem 4.11. For $\eta > 0$, and with the choice $\hat{h}_l = h_l$, Equation (4.32) yields

$$\sum_{l=0}^{L}(h_l + \hat{h}_l + \sqrt{\varepsilon})M_l^{-1/2} \leq C\left(\sum_{l=0}^{L}(h_l + \sqrt{\varepsilon})M_l^{-1/2}\right)$$

$$\leq C(\eta)\left(\sqrt{\varepsilon} + \sum_{l=1}^{L}h_l\frac{h_L}{h_l}l^{-(1+\eta)}\right) \leq C(\eta)(h_L + \sqrt{\varepsilon})$$

and

$$M_0^{-1/2} \leq C(h_L + \sqrt{\varepsilon}).$$

For the approximation of $\mathbb{E}[u^\varepsilon]$, this amounts to the following bound for the overall work for the MLMC FE–HMM at level L

$$\mathcal{W}(L) \lesssim \sum_{l=0}^{L} M_l(N_{l,mac})^2 \leq C\left(h_L^{-2} + \sum_{l=1}^{L}\left(\frac{h_l}{h_L}\right)^2 l^{2+2\eta}h_l^{-2d}\right)$$

$$\leq C\,h_L^{-2}\left(1 + \sum_{l=1}^{L}l^{2(1+\eta)}h_l^{2-2d}\right).$$

In the case $d = 1$, we observe that $\sum_{l=1}^{L}l^{2(1+\eta)} \leq CL^{3+2\eta} \leq C\log_2(N_{L,mac})^{3+2\eta}$ and the overall computational complexity is bounded by $\mathcal{W}(L) = \mathrm{O}(N_{L,mac}^2 \log_2(N_{L,mac})^{3+2\eta})$. In the case $d = 2, 3$ similar calculation as in the proof of Theorem 4.11 yield

$$\mathcal{W}(L) \leq C\,h_L^{-(2d)}L^{2(1+\eta)},$$

with C independent of L and ε. Using again $L = \log_2(N_{L,mac})/d$ we obtain the asserted work estimates. $\qquad\square$

We see in Theorem 4.11 and Theorem 4.12 that the work of the MLMC FE–HMM for the approximation of $\mathbb{E}[u^\varepsilon]$ and $\mathbb{E}[\bar{u}]$ is independent of the scale parameter ε and equals, up to a logarithmic term, the work for solving one deterministic elliptic multiscale problem with the FE–HMM at the finest grid $h_L = \mathrm{O}(2^{-L})$ (except for the approximation of $\mathbb{E}[\bar{u}]$ for $d = 1$).

Remark 4.13. We notice that for the approximation of $\mathbb{E}[u^\varepsilon]$, there is a lower bound on the accuracy that can be obtained, namely $\mathrm{O}(\sqrt{\varepsilon})$. This accuracy is reached with a "macro mesh" of size $h_L = \mathrm{O}(\sqrt{\varepsilon})$. If an accuracy smaller than $\mathrm{O}(\sqrt{\varepsilon})$ is needed, we switch to the fine scale solver in the FE–HMM. We emphasize that in this setting, only the (few) solves on the finest grid will be done with a resolved FE method.

5. Numerical Experiments

In this section we present numerical simulations of the approximations of the expectation of the solution of the elliptic equation with stochastic multiscale coefficient given by

$$(5.1) \qquad -\mathrm{div}(A^{\varepsilon}(\omega, x)\nabla u^{\varepsilon}(\omega, x)) = f(x),$$

for x in $D = (0, 1)$. We choose a model problem in one space dimension and develop all numerical experiments for the case of only one "fast" spatial length scale, i.e., we assume that $n = 1$ in Equation (2.9). This is to be able to compute with reasonable effort "reference" solutions for convergence tests, and also to be able to investigate the performance of the methods in the so-called "resonance" cases, when the meshwidth either over- or under-resolves the length-scale ε. Further, we impose Dirichlet boundary conditions and set $f \equiv 1$ for all $x \in D$. For the stochastic multiscale coefficient we consider two choices: First, the *modulating tensor*

$$(5.2) \qquad A^{\varepsilon}(\omega, x) = 1 + \sin(\pi x) + \sin(\pi(\frac{x}{\varepsilon} + 1/2)) \cdot (2\, Y(\omega) - 1),$$

and second the *amplitude tensor*

$$(5.3) \qquad A^{\varepsilon}(\omega, x) = 1 + \sin(\pi x) + \sin(\pi(\frac{x}{\varepsilon} + 1/2)) \cdot Y(\omega).$$

For both choices Y is a uniformly distributed random variable taking values in $[0, 1]$. We plotted the realization of each tensor for $Y(\omega) = 0.6$ in Figure 1 as well as empirical expectations and variances over 10000 samples. We note that the variance of the modulating tensor is higher than of the amplitude tensor. Whereas the mean of the modulating tensor exhibits a smaller amplitude. With these choices of the coefficient the simulations are not exposed to an additional error truncating the Karhunen–Lóeve expansion.

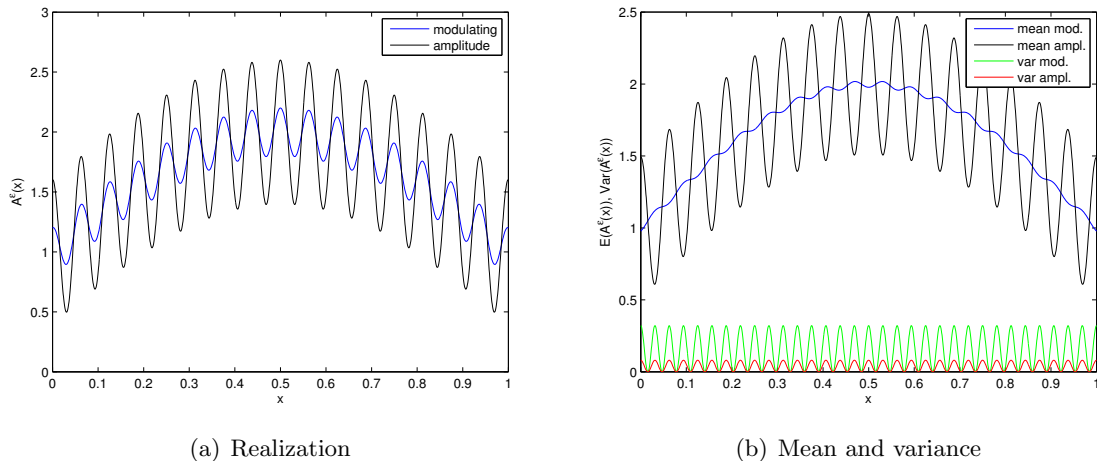(a) Realization

(b) Mean and variance

Figure 1. The realization for $Y(\omega) = 0.6$ for the modulating and amplitude tensor and empirical mean and variance.

In the simulation we compare the performance of the singlelevel Monte Carlo Finite Element method with the multilevel Monte Carlo Finite Element method and the singlelevel Monte Carlo Finite Element Heterogeneous Multiscale method with the multilevel Monte

Carlo Finite Element Heterogeneous Multiscale method. As expected from the theoretical results, the singlelevel and the multilevel Monte Carlo Finite Element methods are converging to the solution $\mathbb{E}(u^\varepsilon)$, whereas the singlelevel and multilevel Monte Carlo Finite Element Heterogeneous Multiscale methods are converging to the solution $\mathbb{E}(\bar{u})$.

For the convergence analysis, we calculate numerically an approximate reference solution $\mathbb{E}(u^\varepsilon)$ through

$$\mathbb{E}(u^\varepsilon(x)) = \int_\Omega u^\varepsilon(\omega, x)\, d\mathbb{P}(\omega) = \int_0^1 u^\varepsilon(z, x)\, dz.$$

This last integral is then approximated by a Gauss–Legendre quadrature with 50 quadrature nodes. We note that this is only possible due to the model problem having only one term in the Karhunen–Loève expansion; if we have a full Karhunen–Loève expansion, this approach is not feasible. The quadrature weights and nodes were computed according to the function `gauleg` in [40]. For a given y the integrand $u^\varepsilon(y, x)$ is approximated by the Finite Element method using a equidistant mesh with 8193 degrees of freedom (which corresponds to discretization level $L = 13$). We denote this approximation

$$\mathbb{E}(u^\varepsilon) \approx \mathrm{GQ}_{50}[u^\varepsilon_{13,FEM}] \,.$$

The same quadrature rule is used to calculate the reference solution $\mathbb{E}(\bar{u})$. Then, for a given y, the integrand $\bar{u}(y, x)$ is approximated by the Finite Element Heterogeneous Multiscale method applied to Equation (5.1) with $\varepsilon = 10^{-8}$ on a equidistant grid with 8193 degrees of freedom, i.e.,

$$\mathbb{E}(\bar{u}) \approx \mathrm{GQ}_{50}[u^{10^{-8}}_{13,HMM}].$$

So we use so-called "overkill" reference solutions in the domain D, (nearly) exactly integrated in the stochastic domain to approximate the expectation.

For the numerical simulation the Finite Element method implemented uses a hat basis and the 8-point Gauss–Legendre quadrature rule to integrate the stiffness matrix. The quadrature weights and nodes are computed according to [40].

For the convergence analysis we consider all levels up to $L = 10$. In all simulations we set $\varepsilon = 2^{-5}$. With this choice we have $h_l \geq \varepsilon$ for $l = 0, \ldots, 5$ and $h_l < \varepsilon$ for $l = 6, \ldots, L$. For the singlelevel Monte Carlo methods we have $N_L = 2^L + 1$ degrees of freedom (including the 2 boundary points). The number of samples is chosen as $M_L = 2^{2L}$, according to Equation (3.9) (in the case of scale resolution) and Equation (3.11). The multilevel Monte Carlo methods with levels $l = L_{\min}, \ldots, L$, where $L_{\min} \in \{0, \ldots, L\}$, use $N_l = 2^l + 1$ degrees of freedom and sample numbers according to the preceding sections. The parameter η in the number of samples is chosen to be $\eta = 0.001$ (see Equation (3.17), Theorem 4.11 and Theorem 4.12). The Finite Element Heterogeneous Multiscale method uses $N_{\mathrm{mic}} = 10$ degrees of freedom of the micro problem and $\delta = 2\varepsilon$.

We calculate the mean square error for the singlelevel Monte Carlo Finite Element and the multilevel Monte Carlo Finite Element method by

$$\|\mathbb{E}(u^\varepsilon) - E^L(u^\varepsilon_{L,FEM})\|_{L^2(\Omega; L^2(D))} \approx \frac{1}{S}\Big(\sum_{s=1}^S \Big(\frac{h_{\mathrm{Ref}}}{2} \sum_{i=1}^{N_{\mathrm{Ref}}} \frac{1}{3}\Delta E_L(x_i)^2$$
$$+ \frac{4}{3}\left(\frac{\Delta E_L(x_i) + \Delta E_L(x_{i+1})}{2}\right)^2 + \frac{1}{3}\Delta E_L(x_{i+1})^2\Big)\Big)^{\frac{1}{2}},$$

where $\Delta E_L := \mathbb{E}(u^\varepsilon) - P_L E^L(u_L^{\varepsilon,k})$ and P_L is the exact prolongation to the grid on which the reference solution was computed. Here, h_{Ref} is the meshwidth of the grid of the reference solution and N_{Ref} its degrees of freedom. The variance of the multilevel Monte Carlo method is approximated by a Monte Carlo method with S samples. In all simulations we chose $S = 30$. The L^2-error for the singlelevel Monte Carlo Finite Element Heterogeneous Multiscale and multilevel Monte Carlo Finite Element Heterogeneous Multiscale method is computed similarly.

The error in the V-norm is calculated by

$$\|\mathbb{E}(u^\varepsilon) - E^L(u_L^\varepsilon)\|_{L^2(\Omega;V)} \approx \frac{1}{S} \left(\sum_{s=1}^S h_{\mathrm{Ref}} \sum_{i=1}^{N_{\mathrm{Ref}}} \left(\frac{\Delta E_L(x_{i+1}) - \Delta E_L(x_i)}{h_{\mathrm{Ref}}} \right)^2 \right)^{\frac{1}{2}}.$$

The $H^1(D)$-error is the approximation of

$$\|\mathbb{E}(u^\varepsilon) - E^L(u_L^\varepsilon)\|_{L^2(\Omega;H^1(D))} = \left(\|\mathbb{E}(u^\varepsilon) - E^L(u_L^\varepsilon)\|_{L^2(\Omega;L^2(D))}^2 + \|\mathbb{E}(u^\varepsilon) - E^L(u_L^\varepsilon)\|_{L^2(\Omega;V)}^2 \right)^{\frac{1}{2}}.$$

The computations were carried out on an AMD 8x16 Opteron, and used less than 1GB of memory per core. The wall-clock time used for all experiments (at a fixed value of ε) ranged form $723 - 739$ minutes per each of the tensors A^ε. We use a C++-implementation with some C inserts for MPI-related code, as a vector type we use Eigen 3 (3.0.4). The sparse linear system is solved with UMFPACK from SuiteSparse (V3.7.0) (see [22, 20, 21]), called through the package Eigen. OpenMPI (1.4.5) (see [25]) was used to parallelize over samples. The work was distributed among processors using a greedy algorithm under the assumption that the runtime of Finite Element and Heterogeneous Multiscale methods scale linearly in the number of degrees of freedom. More precisely, we predict the runtime per core and assign as many samples of the largest level as possible without exceeding the projected runtime per core, successively fill up with smaller levels, so that the last processor handles the remainder.
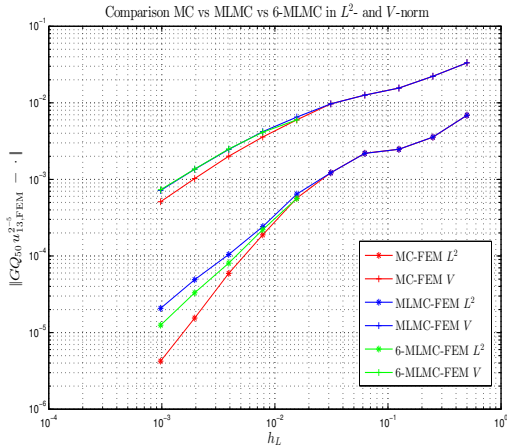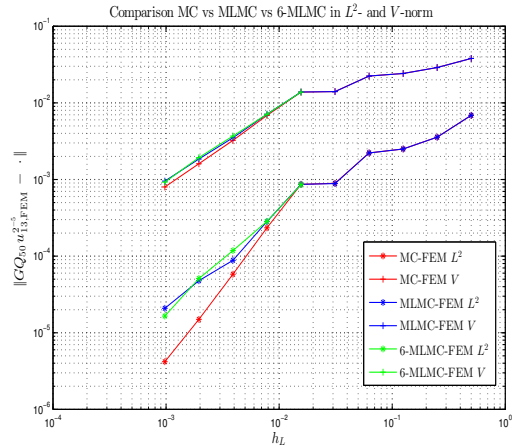
(a) Modulating tensor (Equation (5.2)) (b) Amplitude tensor (Equation (5.3))

FIGURE 2. Estimated $L^2(\Omega;L^2(D))$-errors and $L^2(\Omega;V)$-errors of Finite Element based methods for the modulating and the amplitude tensor.

In Figure 2 and Figure 3 we see the $L^2(D)$- and V-norms of the Finite Element based methods. Figure 4 shows the convergence of the Heterogeneous Multiscale based methods to $\mathbb{E}[\bar{u}]$.

In Figure 2 we observe convergence of the singlelevel Monte Carlo Finite Element approximation, multilevel Monte Carlo Finite Element approximation and a multilevel Monte Carlo Finite Element approximation starting at level 6 ($L_{\min} = 6$) to the reference solution approximating $\mathbb{E}[u^\varepsilon]$ with respect to h_l. The level 6 multilevel Monte Carlo Finite Element method has resolved solutions on all its levels, whereas the multilevel Monte Carlo method has under-resolved samples on the first levels. We observe convergence of order h_l in the resolved region for all methods. However, in the under-resolved region, where we have no convergence of the Finite Element error, we have convergence of order $1/2$, due to the convergence of the Monte Carlo error. Further, both tensors give similar results. The differences in the convergence for the under-resolved levels, however, are probably due to the difference in the variances of the two tensors (see Figure 1(b)). In Figure 2 we see further the effect of the choice of samples for the multilevel method on under-resolved levels. We chose here less samples than in the analysis in [10], because of the dependence on ε in Theorem 3.4. For the multilevel Monte Carlo Finite Element method the many under-resolved samples do not seem to affect the convergence as soon as the finest level of the method is resolved. The number of samples on the under-resolved levels seem to outweigh the error of the Finite Element method in the space domain of each sample.
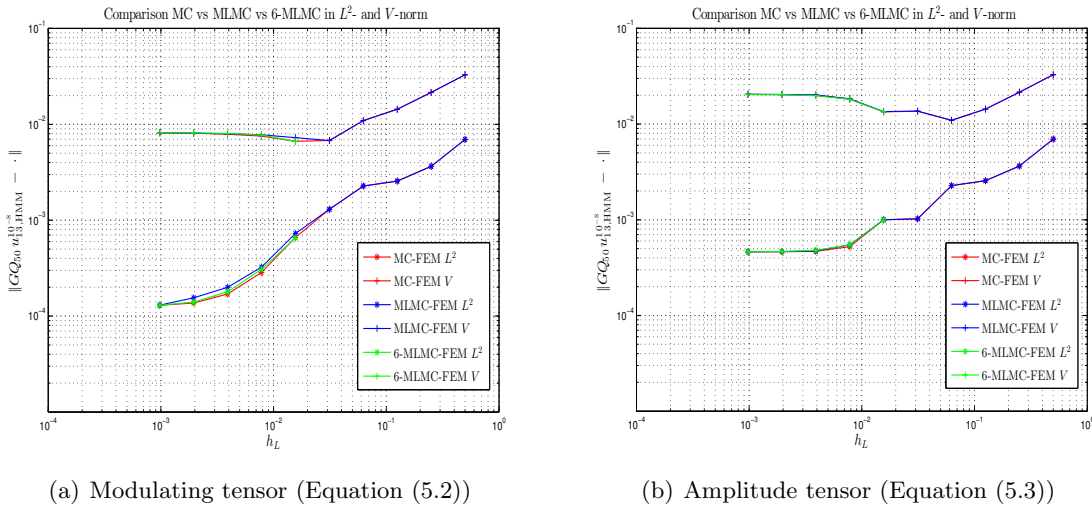

(a) Modulating tensor (Equation (5.2))   (b) Amplitude tensor (Equation (5.3))

FIGURE 3. Estimated $L^2(\Omega; L^2(D))$-errors and $L^2(\Omega; V)$-errors with respect to the homogenized solution $\mathbb{E}[\bar{u}]$ of Finite Element based methods for the modulating and the amplitude tensor.

Figure 3 shows the convergence of the singlelevel Monte Carlo Finite Element approximation, multilevel Monte Carlo Finite Element approximation and again a multilevel Monte Carlo Finite Element approximation starting at level 6 to the expectation of the homogenized solution $\mathbb{E}[\bar{u}]$. For under-resolved levels we have convergence of order $1/2$. Convergence is in

this case bounded by

$$\mathbb{E}\|E^L[u^\varepsilon_{L,FEM}] - \mathbb{E}[\bar{u}]\|^2_V \leq \mathbb{E}\|E^L[u^\varepsilon_{L,FEM}] - \mathbb{E}[u^\varepsilon]\|^2_V + \mathbb{E}\|\mathbb{E}[u^\varepsilon - \bar{u}]\|^2_V.$$

The first term on the right hand side converges with Theorem 3.4 with rate $1/2$ for the under-resolved levels if the numbers of samples are chosen accordingly. The second term is constant, depending on the choice of A^ε. In Figure 3 we display the difference in these constants. The expectation of the modulating tensor seems to be closer to the expectation of the homogenized solution than for the amplitude tensor, presumably due to the smaller amplitude of the expectation of the tensor (see Figure 1(b)). Depending on this constant, this method could be appropriate to sample the expectation of the homogenized solution just by a multilevel Monte Carlo Finite Element approximation.

In Figure 4 convergence of the singlelevel Monte Carlo Finite Element Heterogeneous Multiscale method, the multilevel Monte Carlo Finite Element Heterogeneous Multiscale method and the level 6 multilevel Monte Carlo Finite Element Heterogeneous Multiscale method in the $L^2(D)$-norm is depicted. The methods approximate $E[\bar{u}]$. Therefore, a comparison with the singlelevel and multilevel Monte Carlo Finite Element methods is not viable. However, we observe that the Heterogeneous Multiscale method does not introduce a bias in the convergence of rate $\mathrm{O}(h_l)$ for under-resolved levels. The numerical experiments are in agreement with the theoretical findings. Further, if we compare the error for the first levels in Figure 3 with the first levels in Figure 4 we have similar qualitative behavior of the errors.
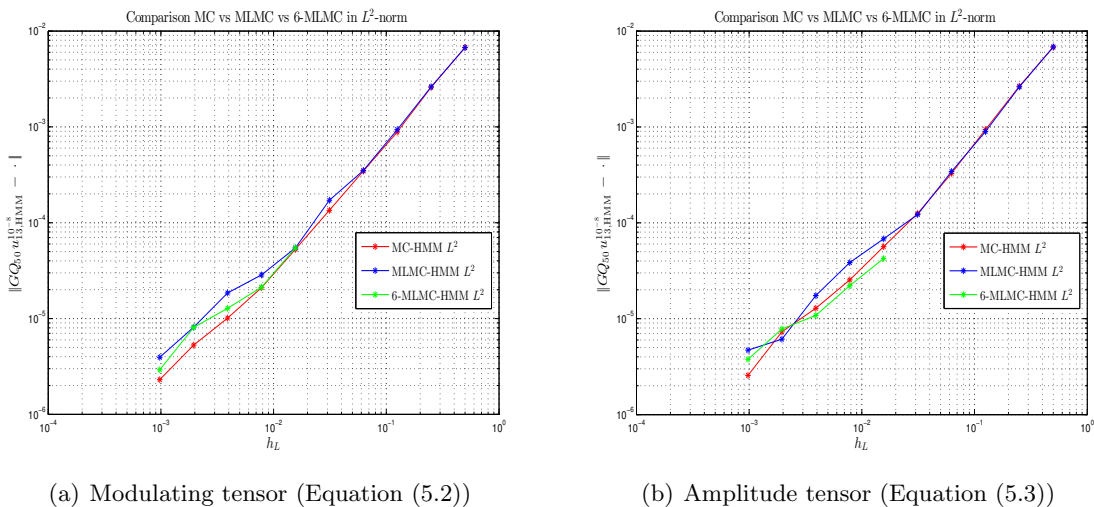

(a) Modulating tensor (Equation (5.2))   (b) Amplitude tensor (Equation (5.3))

FIGURE 4. Estimated $L^2(\Omega; L^2(D))$-errors of Heterogeneous Multiscale based methods for the modulating and the amplitude tensor.

6. CONCLUSIONS

We presented a theory and algorithms for the efficient approximation of a class of elliptic boundary value problems with random diffusion coefficients. The random coefficients were assumed to exhibit multiple, separated spatial length scales, and were assumed to be $\mathbb{P}$-a.s. coercive and continuous. Based on the theoretical results in [31], we introduced a class of

multilevel Monte Carlo Finite Element methods for the efficient numerical computation of the expectation of the solution at small, positive values of the scale parameter $\varepsilon_n(\varepsilon)$.

The principal conclusions are, that the multilevel Monte Carlo approach with a Finite Element approximation ensures the full convergence rate if the finest Finite Element mesh used in the multilevel Monte Carlo Finite Element simulation does resolve all physical length scales. However, the multilevel Monte Carlo Finite Element approach, on the other hand, allows for *mean square convergence of under-resolved physical solutions* in the preasymptotic regime. The under-resolved physical length scales of (the large number of) samples on coarse grids do not pollute the multilevel Monte Carlo estimate of the expected value of the solution, provided that the finest mesh in the hierarchy of discretizations does resolve the length scales. In addition, our analysis revealed reduced sample numbers M_l on discretization levels without scale resolution (i.e., for mesh levels l such that $h_l > \varepsilon_n(\varepsilon)$).

Further, a combination of a multilevel Monte Carlo method with the recently developed Finite Element Hierarchic Multiscale method results in a discretization scheme which allows the efficient numerical estimation of the ensemble average of the solution, even if the physical length scales are not resolved on the finest mesh of the hierarchy. The multilevel Monte Carlo Finite Element Heterogeneous Multiscale method can generate, in stochastic elliptic PDE problems with two separated length scales, numerical approximations to the unknown solution's mean field with work versus accuracy independent of the physical scale parameter ε. The complexity of the multilevel Monte Carlo Finite Element Heterogeneous Multiscale method is proved to be proportional to the complexity of the Finite Element Heterogeneous Multiscale method applied to one single, deterministic elliptic multiscale problem (similar to the multilevel Monte Carlo method based on a standard Finite Element method).

In the present paper, we reformulated the stochastic multiscale problem by means of a parametric Karhunen–Loève expansion to an infinite-dimensional, parametric, deterministic problem, given in Equation (2.15). This was done here solely for purposes of numerical analysis (to derive the random homogenized limiting problem, among others). However, the parametric, deterministic formulation in Equation (2.15) may also be used as starting point for an approximation based on *Quasi Monte-Carlo methods* along the lines of [34].

In the present work and complexity analysis, we assumed that approximate samples of the random, multiscale coefficient could be drawn at cost O(1). The results in this paper cover, therefore, in particular the case of "finite-dimensional" noise. They can be extended straightforwardly to the case when the Karhunen–Loève expansion is truncated at a level-dependent truncation dimension.

The formalism in the present paper and the analysis in [31] allows for generalizations to random coefficients which are not uniformly elliptic, but rather are only elliptic $\mathbb{P}$-a.s., with ellipticity constants which degenerate with small probability. This occurs, for example, in inputs with so-called lognormal probability laws.

Appendix A. Multiscale Karhunen–Loève expansion

For numerical simulations as well as for the analytical developments in the present paper, the random field A in Equation (2.10) must be represented parametrically. Here, we briefly present the Karhunen–Loève expansion of A from [31]. We give a particular example of a parametric expansion (2.4), the *Karhunen–Loève* expansion of a random matrix function $A(\cdot; x, \boldsymbol{y})$. We give, in particular, sufficient conditions in order for Assumption 2.2 to hold. We formulate these conditions in terms of the smoothness of the covariance of the matrix

function $A(\cdot; x, \boldsymbol{y})$, which is given by the fourth order tensor

$$\mathbf{Cov}[A]_{iji'j'}(x, \boldsymbol{y}, x', \boldsymbol{y}') = \int_\Omega (A_{ij}(\omega; x, \boldsymbol{y}) - \mu_{ij}(x, \boldsymbol{y}))(A_{i'j'}(\omega; x', \boldsymbol{y}') - \mu_{i'j'}(x', \boldsymbol{y}'))d\mathbb{P},$$

for $i, j, i', j' = 1, \ldots, d$. Then we have $\mathbf{Cov}[A]_{iji'j'} \in L^\infty((D \times \mathbf{Y}) \times (D \times \mathbf{Y}), \mathbb{R})$, for all i, j, i', j'. We also define the corresponding *covariance operator* $\mathbf{Q}_A : L^2(D \times \mathbf{Y})^{d \times d}_{\text{sym}} \to L^2(D \times \mathbf{Y})^{d \times d}_{\text{sym}}$ as

$$(\mathbf{Q}_A \Phi)_{ij}(x, \boldsymbol{y}) := \int_D \int_\mathbf{Y} \mathbf{Cov}[A]_{iji'j'}(x, \boldsymbol{y}, x', \boldsymbol{y}')\Phi_{i'j'}(x', \boldsymbol{y}')d\boldsymbol{y}'dx'.$$

Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$ denote the eigenvalues of the operator $\mathbf{Q}_A$ and let $\Phi_k \in L^2(D \times \mathbf{Y})^{d \times d}$ denote the corresponding eigenvectors. We assume that $\|\Phi_k\|_{L^2(D \times \mathbf{Y})^{d \times d}} = 1$ for all $k \in \mathbb{N}$. Any random field $A \in L^2(\Omega; L^2(D \times \mathbf{Y})^{d \times d}_{\text{sym}})$ can be represented by a Karhunen–Loève expansion

$$(A.1) \qquad A(\omega; x, \boldsymbol{y}) = \mu(x, \boldsymbol{y}) + \sum_{k=1}^\infty \sqrt{\lambda_k}\Phi_k(x, \boldsymbol{y})Z_k(\omega),$$

where Z_k are pairwise uncorrelated random variables that satisfy

$$Z_k(\omega) = \frac{1}{\sqrt{\lambda_k}} \int_{D \times \mathbf{Y}} (A_{ij} - \mu_{ij})(\Phi_k)_{ij} d\boldsymbol{y}dx.$$

By Assumption 2.2, the random coefficients Z_k in (A.1) are uniformly bounded, $\mathbb{P}$-a.s. for all $k \in \mathbb{N}$. Note also that, due to the normalization assumption $\|\Phi_k\|_{L^2(D \times \mathbf{Y})^{d \times d}} = 1$ the probability densities of the random variables Z_k are not necessarily supported in $[-1, 1]$. We will use the following Propositions whose proof can be found in [31].

Proposition A.1. *Assume that the random diffusion matrix A in (2.1) satisfies*

$$(A.2) \qquad\qquad A \in L^2(\Omega; H^t_{mix}(D \times \mathbf{Y}))$$

where, for $t \geq 0$, the space $H^t_{mix}(D \times \mathbf{Y})$ is defined by $H^t_{mix}(D \times \mathbf{Y}) = H^t(D) \otimes H^t_\#(Y_1) \otimes \ldots \otimes H^t_\#(Y_n)$ with $\otimes$ denoting the tensor product of separable Hilbert spaces and with H^t denoting, for noninteger values of t, the fractional order Sobolev space (see, e.g, [42]).

Then it holds that $\mathbf{Cov}[A] \in (H^t_{mix}(D \times \mathbf{Y}) \otimes H^t_{mix}(D \times \mathbf{Y}))$. Moreover, for all $\varepsilon > 0$, there exists a constant $c = c(\varepsilon) > 0$ such that for all $k \geq 1$ holds $\lambda_k \leq c(\varepsilon)k^{-t/d+\varepsilon}$.

Proposition A.2. *Assume that the random coefficient A in Equation (2.1) satisfies Equation (A.2) for some $t > d/2$. Then for every $d/2 < t^* < t$ there is a constant $c > 0$ independent of k such that*

$$\sum_{i,j=1}^d \|(\Phi_k)_{ij}\|^2_{L^\infty(D \times \mathbf{Y})} \leq c\lambda_k^{-2t^*/t}.$$

In the expansion (A.1), let $\Psi_k = \sqrt{\lambda_k}\Phi_k$. We then find that there exists a constant $c > 0$ (depending on t, t^* and on d) such that for all k

$$\sum_{i,j=1}^d \|(\Psi_k)_{ij}\|^2_{L^\infty(D \times \mathbf{Y})} \leq c\lambda_k^{1-2t^*/t}.$$

From Proposition A.1, we find that

$$\sum_{i,j=1}^{d} \|(\Psi_k)_{ij}\|_{L^\infty(D\times Y)}^2 \le ck^{(-t/d+\varepsilon)(1-2t^*/t)}.$$

For each vector $\xi \in \mathbb{R}^d$, we have

$$|(\Psi_k)_{ij}(x,\boldsymbol{y})\xi_i\xi_j|^2 \le (\sum_{i,j=1}^{d} \|(\Psi_k)_{ij}\|_{L^\infty(D\times\mathbf{Y})}^2)(\sum_{i,j=1}^{d} \xi_i^2\xi_j^2) \le ck^{(-t/d+\varepsilon)(1-2t^*/t)}|\xi|^4.$$

Therefore we may choose

(A.3) $$\beta_k = ck^{(-t/d+\varepsilon)(1/2-t^*/t)}.$$

When t is sufficiently large, e.g., $(t/d-\varepsilon)(1/2-t^*/t) > 1$, this implies that $\beta = \{\beta_k\}_{k\ge 1} \in \ell^1$. Assuming that the random variables Z_k in the expansion (A.1) are uniformly bounded, we can assume that they are rescaled so that the support of their laws equals $[-1,1]$. Assumption 2.2 holds when the constant α_0 is sufficiently large.

References

[1] A. ABDULLE, *On a priori error analysis of fully discrete heterogeneous multiscale fem*, SIAM, Multiscale Model. Simul., 4 (2005), pp. 447–459.

[2] ———, *Analysis of a heterogeneous multiscale FEM for problems in elasticity*, Math. Models Methods Appl. Sci., 16 (2006), pp. 615–635.

[3] ———, *The finite element heterogeneous multiscale method: a computational strategy for multiscale pdes*, GAKUTO Int. Ser. Math. Sci. Appl., 31 (2009), pp. 135–184.

[4] ———, *A priori and a posteriori error analysis for numerical homogenization: a unified framework*, Ser. Contemp. Appl. Math. CAM, 16 (2011), pp. 280–305.

[5] ———, *Discontinuous galerkin finite element heterogeneous multiscale method for elliptic problems with multiple scales*, Math. Comp., 81 (2012), pp. 687–713.

[6] A. ABDULLE AND C. SCHWAB, *Heterogeneous multiscale FEM for diffusion problems on rough surfaces*, Multiscale Model. Simul., 3 (2004/05), pp. 195–220.

[7] A. BARTH, *A Finite Element Method for martingale-driven stochastic partial differential equations*, Comm. Stoch. Anal., 4 (2010), pp. 355–375.

[8] A. BARTH AND A. LANG, L^p *and almost sure convergence of a Milstein scheme for stochastic partial differential equations*. SAM report 2011-15, June 2009.

[9] ———, *Simulation of stochastic partial differential equations using Finite Element methods*. to appear in Stochastics, 2010.

[10] A. BARTH, C. SCHWAB, AND N. ZOLLINGER, *Multi-level monte carlo finite element method for elliptic pdes with stochastic coefficients*, Numerische Mathematik, 119 (2011), pp. 123–161.

[11] A. BENSOUSSAN, J.-L. LIONS, AND G. PAPANICOLAOU, *Asymptotic analysis for periodic structures*, vol. 5 of Studies in Mathematics and its Applications, North-Holland Publishing Co., Amsterdam, 1978.

[12] A. BOURGEAT, A. MIKELIĆ, AND S. WRIGHT, *Stochastic two-scale convergence in the mean and applications*, J. Reine Angew. Math., 456 (1994), pp. 19–51.

[13] D. BRAESS, *Finite Elemente*, Springer, Berlin, 3rd ed., 2002.

[14] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer, Berlin, 2nd ed., 2002.

[15] P. CIARLET, *The finite element method for elliptic problems.*, vol. 40 of Classics Appl. Math., SIAM, Philadelphia, 2002.

[16] P. CIARLET AND P. RAVIART, *The combined effect of curved boundaries and numerical integration in isoparametric finite element methods*, Math. Foundation of the FEM with Applications to PDE, (1972), pp. 409–474.

[17] A. COHEN, R. DEVORE, AND C. SCHWAB, *Convergence rates of best N-term Galerkin approximations for a class of elliptic spdes*, Found. Comp. Math., 10 (2010), pp. 615–646.

[18] ———, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs*, to appear J. Analysis and Applications, (2011).

[19] G. DA PRATO AND J. ZABCZYK, *Stochastic equations in infinite dimensions*, vol. 44 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1992.

[20] T. A. DAVIS, *Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method*, ACM Trans. Math. Software, 30 (2004), pp. 196–199.

[21] ———, *A column pre-ordering strategy for the unsymmetric-pattern multifrontal method*, ACM Trans. Math. Software, 30 (2004), pp. 167–195.

[22] T. A. DAVIS AND I. S. DUFF, *A combined unifrontal/multifrontal method for unsymmetric sparse matrices*, ACM Trans. Math. Software, 25 (1999), pp. 1–20.

[23] W. E AND B. ENGQUIST, *The heterogeneous multiscale methods*, Commun. Math. Sci., 1 (2003), pp. 87–132.

[24] W. E, P. MING, AND P. ZHANG, *Analysis of the heterogeneous multiscale method for elliptic homogenization problems*, J. Amer. Math. Soc., 18 (2005), pp. 121–156.

[25] E. GABRIEL, G. E. FAGG, G. BOSILCA, T. ANGSKUN, J. J. DONGARRA, J. M. SQUYRES, V. SAHAY, P. KAMBADUR, B. BARRETT, A. LUMSDAINE, R. H. CASTAIN, D. J. DANIEL, R. L. GRAHAM, AND T. S. WOODALL, *Open MPI: Goals, concept, and design of a next generation MPI implementation*, in Proceedings, 11th European PVM/MPI Users' Group Meeting, Budapest, Hungary, September 2004, pp. 97–104.

[26] M. GEISSERT, M. KOVÁCS, AND S. LARSSON, *Rate of weak convergence of the Finite Element Method for the stochastic heat equation with additive noise*, BIT Numerical Mathematics, 49 (2009), pp. 343–356.

[27] M. B. GILES, *Improved multilevel Monte Carlo convergence using the Milstein scheme*, Preprints NA-06/22, Oxford University Computing Laboratory, Parks Road, Oxford, U.K., 2006.

[28] ———, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.

[29] P. GRISVARD, *Elliptic problems in nonsmooth domains*, vol. 24 of Monographs and Studies in Mathematics, Pitman (Advanced Publishing Program), Boston, MA, 1985.

[30] S. HEINRICH, *Multilevel Monte Carlo Methods*, in Large-Scale Scientific Computing, Third International Conference, LSSC 2001, Sozopol, Bulgaria, June 6-10, 2001, Revised Papers, S. Margenov, J. Wasniewski, and P. Y. Yalamov, eds., vol. 2179 of Lecture Notes in Computer Science, Springer, 2001, pp. 58–67.

[31] V.-H. HOANG AND C. SCHWAB, *Analytic regularity and polynomial approximation of stochastic, parametric elliptic multiscale pdes*, Analysis and Applications (Singapore) (to appear), (2012).

[32] V. V. JIKOV, S. M. KOZLOV, AND O. A. OLEĬNIK, *Homogenization of differential operators and integral functionals*, Springer-Verlag, Berlin, 1994. Translated from the Russian by G. A. Yosifian [G. A. Iosifyan].

[33] M. KOVÁCS, S. LARSSON, AND F. SAEDPANAH, *Finite Element approximation of the linear stochastic wave equation with additive noise*, SIAM J. Numer. Anal., 48 (2010), pp. 408–427.

[34] F. KUO, C. SCHWAB, AND I. SLOAN, *Quasi-monte carlo finite element methods for a class of elliptic partial differe ntial equations with random coefficients*, SIAM Journ. Numer. Analysis, (2012), p. (to appear).

[35] O. LADYZHENSKAYA, *The boundary value problems of mathematical physics*, vol. 49 of Applied Mathematical Sciences, Springer-Verlag, New York, 1985.

[36] S. MISHRA AND C. SCHWAB, *Sparse tensor Multi-Level Monte Carlo Finite Volume Methods for hyperbolic conservation laws with random intitial data*, Tech. Rep. 2010–24, Seminar for Applied Mathematics, ETH Zurich, September 2010. To appear in Math. Comp. (2012).

[37] S. MISHRA, C. SCHWAB, AND J. SUKYS, *Multi-level monte carlo finite volume methods for shallow water equations with uncertain topography in multi-dimensions*, Tech. Rep. 2011/70, Seminar for Applied Mathematics, ETH Zürich, 2011. In review.

[38] R. L. NAFF, D. F. HALEY, AND E. A. SUDICKY, *High-resolution Monte Carlo simulation of flow and conservative transport in heterogeneous porous media 1. Methodology and flow results*, Water Resour. Res., 34 (1998), pp. 663–677.

[39] ———, *High-resolution Monte Carlo simulation of flow and conservative transport in heterogeneous porous media 2. Transport results*, Water Resour. Res., 34 (1998), pp. 679–697.

[40] W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING, AND B. P. FLANNERY, *Numerical recipes*, Cambridge University Press, Cambridge, third ed., 2007. The art of scientific computing.

[41] C. SCHWAB AND C. J. GITTELSON, *Sparse tensor discretizations of high-dimensional parametric and stochastic pdes*, Acta Numerica, 20 (2011), pp. 291–467.

[42] H. TRIEBEL, *Theory of Function Spaces, 2nd Ed.*, Joh. Ambrosius Barth, Leipzig, Germany, 1995.

(Assyr Abdulle)
ANMC, Mathematics Section
École Polytechnique Fédérale de Lausanne
1015 Lausanne
   E-mail address: `assyr.abdulle@math.epfl.ch`

(Andrea Barth)
ETH, Seminar für Angewandte Mathematik
Rämistrasse 101
8092 Zürich
   E-mail address: `andrea.barth@math.ethz.ch`

(Christoph Schwab)
ETH, Seminar für Angewandte Mathematik
Rämistrasse 101
8092 Zürich
   E-mail address: `schwab@math.ethz.ch`

27.2012 C. EFFENBERGER:
Robust successive computation of eigenpairs for nonlinear eigenvalue problems

28.2012 C. LUBICH, T. ROHWEDDER, R. SCHNEIDER, B. VANDEREYCKEN:
Dynamical approximation of hierarchical Tucker and tensor-train tensors

29.2012 A. ABDULLE, Y. BAI:
Adaptive reduced basis finite element heterogeneous multiscale method

30.2012 L. KARLSSON, D. KRESSNER:
Optimally packed chains of bulges in multishift QR algorithms

31.2012 F. BONIZZONI, A. BUFFA, F. NOBILE:
Moment equations for the mixed formulation of the Hodge Laplacian with stochastic data

32.2012 M. MOTAMED, F. NOBILE, R. TEMPONE:
Analysis and computation of the elastic wave equation with random coefficients

33.2012 T. LASSILA, A. MANZONI, A. QUARTERONI, G. ROZZA:
Generalized reduced basis methods and n-width estimates for the approximation of the solution manifold of parametric PDEs

34.2012 P.CHEN, A. QUARTERONI, G. ROZZA:
Comparison between reduced basis and stochastic collocation methods for elliptic problems

35.2012 A. ABDULLE, G. VILMART, K.C. ZYGALAKIS:
Mean-square A-stable diagonally drift-implicit integrators of weak second order for stiff Itˆo stochastic differential equations

36.2012 L. DEDÈ, A. QUARTERONI:
Isogeometric analysis for second order partial differential equations on surfaces

37.2012 L. DEDÈ, A. QUARTERONI, A. TAGLIABUE:
Isogeometric analysis for high-order partial differential equations: error estimation and problems in fluid dynamics

38.2012 A. ABDULLE, P. LIN, A.V. SHAPEEV:
A priori and posteriori $W^{1,\infty}$ error analysis of a QC method for complex lattices

39.2012 A. ABDULLE, G. VILMART:
PIROCK: a Swiss-knife partitioned implicit-explicit orthogonal Runge-Kutta Chebyshev integrator for stiff diffusion-advection-reaction problems with or without noise

40.2012 F. NEGRI, G. ROZZA, A. MANZONI, A. QUARTERONI:
Reduced basis method for parametrized elliptic optimal control problems

41.2012 A. ABDULLE, A. BARTH, X. SCHWAB:
Multilevel Monte Carlo methods for stochastic elliptic multiscale PDES