

Multilevel accelerated quadrature for PDEs with log-normally distributed diffusion coefficient*

Helmut Harbrecht[†], Michael Peters[†], and Markus Siebenmorgen[†]

Abstract. This article is dedicated to multilevel quadrature methods for the rapid solution of stochastic partial differential equations with a log-normally distributed diffusion coefficient. The key idea of such approaches is a sparse-grid approximation of the occurring product space between the stochastic and the spatial variable. We develop the mathematical theory and present error estimates for the computation of the solution’s moments with focus on the mean and the variance. Especially, the present framework covers the multilevel Monte Carlo method and the multilevel quasi-Monte Carlo method as special cases. The theoretical findings are supplemented by numerical experiments.

Key words. multilevel quadrature, PDEs with stochastic data, log-normal diffusion, Karhunen-Loève expansion, finite element method

AMS subject classifications. 65N30, 65D32, 60H15, 60H35

1. Introduction. In this article, we consider multilevel quadrature methods to compute the moments of the solution to elliptic partial differential equations with log-normally distributed diffusion coefficient. The basic idea of the multilevel quadrature is a sparse-grid-like discretization of the underlying Bochner space $L^2_{\mathbb{P}}(\Omega; H^1_0(D))$. The spatial variable is discretized by a classical finite element method whereas the stochastic variable is treated by an appropriately chosen quadrature rule, which naturally leads to a *non-intrusive* method. Since the problem’s solution provides the necessary mixed Sobolev regularity, the approximation errors on the different levels of resolution can be equilibrated in a sparse-grid-like fashion, cf. [8, 20, 42]. This idea has already been proposed for different quadrature strategies in case of uniformly elliptic diffusion coefficients in [23]. The well-known *Multilevel Monte Carlo Method* (MLMC), as introduced in [3, 15, 16, 26, 27], and also the *Randomized Multilevel Quasi-Monte Carlo Method*, as introduced in [30], only provide probabilistic error estimates in the mean-square sense. To avoid this drawback, two fully deterministic methods have been proposed in [23], namely the *Multilevel Quasi-Monte Carlo Method* (MLQMC) and the *Multilevel Polynomial Chaos Method* (MLPC).

The multilevel Monte Carlo method has been considered at first for a log-normal diffusion coefficient in [12] and further been analyzed in [11, 40]. However, for deterministic quadrature methods, the log-normal case is much more involved due to the unboundedness of the domain of integration, i.e. \mathbb{R}^m for some $m \in \mathbb{N}$, in combination with the stronger regularity requirements on the integrand. This makes the analysis of the quadrature error difficult. In particular, special regularity results are required which extend those of [2, 10, 29].

For a finite stochastic dimension, we show that the multilevel quasi-Monte Carlo quadrature is feasible also for a log-normal diffusion coefficients if an auxiliary density is introduced.

*This research has been supported by the Swiss National Science Foundation (SNSF) through the project “Rapid Solution of Boundary Value Problems on Stochastic Domains”.

[†]University of Basel, Department of Mathematics and Computer Science, Spiegelgasse 1, 4051 Basel, Switzerland (helmut.harbrecht@unibas.ch, michael.peters@unibas.ch, markus.siebenmorgen@unibas.ch)

If the stochastic dimension tends to infinity, this technique is no longer applicable since the error estimates involve a *discrepancy* which grows exponentially in the stochastic dimension m . A possibility to bypass this obstruction has recently been proposed in [19] by means of randomly shifted lattice rules. There, however, only probabilistic error estimates are obtained in the mean-square sense. In case of the Halton sequence, cf. [33], it is possible to adapt the ideas presented in [36] to show an almost dimension independent, i.e. up to a constant which depends linearly on the dimension, and *deterministic* convergence rate, see [24] for the details. We emphasize that the mixed regularity estimates provided here directly carry over to multilevel methods based on those quasi-Monte Carlo quadrature formulae.

A log-normally distributed diffusion coefficient depends non-linearly on the stochastic parameter. Thus, MLPC is no longer feasible since a polynomial chaos expansion, cf. [13, 14], would yield a fully coupled system of partial differential equations. This can be avoided by the application of stochastic collocation, cf. [2, 5, 34]. If statistics of the solution, like the mean or the variance, are desired rather than the solution itself, the stochastic collocation coincides with a quadrature rule based on polynomial interpolation. Especially, for the log-normal case, quadrature formulae based on the Hermite polynomials are convenient. This yields the *Multilevel Gaussian Quadrature Method* (MLGQ) which we also analyze in this article.

We mention that a sparse-grid-like discretization of the Bochner space $L^2_{\mathbb{P}}(\Omega; H^1_0(D))$ has already been proposed in [6, 18] for partial differential equations with uniformly elliptic diffusion coefficient in the context of a stochastic Galerkin method. Then, however, one arrives again at a huge, fully coupled system of partial differential equations which is hard to solve. In particular, this approach is intrusive. In [5], however, a sparse collocation method for the approximation in the stochastic variable is combined with a wavelet Galerkin discretization in the spatial variable. This decouples the spatial and the stochastic approximation. In our approach, we do not apply an explicit representation with respect to appropriate detail spaces. Instead, we mimic this representation with differences of Galerkin projections on consecutive grids of refinement. This has the advantage that we can employ a standard finite element method in the spatial discretization. Nevertheless, the convergence analysis requires that the differences of Galerkin projections provide sufficient regularity. The related result is proven in Section 8 and constitutes one of the main novelties of the present article.

The rest of this article is organized as follows. Section 2 specifies the diffusion problem under consideration. In particular, the parametric reformulation as a high dimensional deterministic problem is performed here. The solution's moments are hereby transformed into Bochner integrals over a high dimensional parameter domain. In Section 3, we give a brief outline of the multilevel finite element method which we will employ for the spatial discretization later on. The multilevel quadrature method is presented in Section 4. This method equilibrates the quadrature error and the finite element error on each level of refinement. Of course, one could also equilibrate the number of unknowns or even the overall work, cf. [20]. In Section 5, we derive crucial regularity estimates for the solution of the stochastic diffusion problem under consideration. Especially, we present regularity results for the powers of the solution which are the integrands in case of the computation of the moments. Section 6 provides the theoretical background for the Gaussian quadrature in the stochastic variable and follows closely the arguments of the stochastic collocation method in [2]. Here, the regularity estimates from the preceding section are employed to determine the order of quadrature for

each stochastic dimension which is required to guarantee the convergence of the quadrature. Section 7 is concerned with the Monte Carlo and the quasi-Monte Carlo quadrature. In case of the quasi-Monte Carlo quadrature, we introduce an auxiliary density to bound the quadrature error. In Section 8, the previously specified spatial and stochastic discretizations are combined to get convergence results for the multilevel quadrature applied to the solution's moments. In order to apply the convergence results of the single-level quadrature methods to the multilevel quadrature method, the regularity results of Section 5 are extended to the differences of Galerkin projections on two consecutive levels of refinement. Section 9 provides the complexities of the different multilevel quadrature methods under the finite dimensional noise assumption. Finally, in Section 10, the theoretical findings are validated by numerical examples.

In the following, in order to avoid the repeated use of generic but unspecified constants, by $C \lesssim D$ we mean that C can be bounded by a multiple of D , independently of parameters which C and D may depend on. Obviously, $C \gtrsim D$ is defined as $D \lesssim C$, and we write $C \approx D$ if $C \lesssim D$ and $C \gtrsim D$, simultaneously.

2. Problem setting. In the following, let $D \subset \mathbb{R}^n$ for $n = 2, 3$ be a polygonal or polyhedral domain and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with σ -field $\mathcal{F} \subset 2^\Omega$ and probability measure \mathbb{P} . We intend to compute the random function $u(\omega) \in H_0^1(D)$ which solves for almost every $\omega \in \Omega$ the stochastic diffusion problem

$$(2.1) \quad -\operatorname{div}_{\mathbf{x}}(a(\omega)\nabla_{\mathbf{x}}u(\omega)) = f \text{ in } D$$

in the weak sense. Throughout this article, we shall assume that the load f is purely deterministic and belongs to $L^2(D)$. Furthermore, we assume that the logarithm of the diffusion coefficient is a centered Gaussian field which can be represented by a Karhunen-Loève expansion, cf. [31],

$$(2.2) \quad b(\mathbf{x}, \omega) := \log(a(\mathbf{x}, \omega)) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \varphi_k(\mathbf{x}) \psi_k(\omega).$$

Here, $\{\varphi_k\}_k \subset L^\infty(D)$ are pairwise orthonormal functions and $\{\psi_k\}_k$ are independent, standard normally distributed random variables, i.e. $\psi_k(\omega) \sim \mathcal{N}(0, 1)$. For the uniform convergence of the series in (2.2), it is sufficient that the sequence

$$(2.3) \quad \gamma_k := \sqrt{\lambda_k} \|\varphi_k\|_{L^\infty(D)}$$

satisfies $\{\gamma_k\}_k \in \ell^1(\mathbb{N})$, which we assume in the sequel.

In practice, one has of course to compute the expansion (2.2) from the given covariance kernel $\operatorname{Cov}_b(\mathbf{x}, \mathbf{y}) := \int_{\Omega} b(\mathbf{x}, \omega)b(\mathbf{y}, \omega) d\mathbb{P}(\omega)$. Thus, the Karhunen-Loève expansion is either finite of length m or needs to be appropriately truncated after m terms. We will assume this in the following. The respective truncation error has been discussed in [10].

The fact that the random variables $\{\psi_k(\omega)\}_k$ are stochastically independent implies that the pushforward measure $\mathbb{P}_\psi := \mathbb{P} \circ \psi^{-1}$ with respect to the measurable mapping

$$\psi: \Omega \rightarrow \mathbb{R}^m, \quad \omega \mapsto \psi(\omega) := (\psi_1(\omega), \dots, \psi_m(\omega))$$

is given by the joint density function with respect to the Lebesgue measure

$$(2.4) \quad \boldsymbol{\rho}(\mathbf{y}) := \prod_{k=1}^m \rho(y_k), \quad \text{where} \quad \rho(y) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right).$$

With this representation at hand, we can reformulate the stochastic problem (2.1) as a parametric deterministic problem. To that end, we replace the space $L_{\mathbb{P}}^2(\Omega)$ by $L_{\boldsymbol{\rho}}^2(\mathbb{R}^m)$ and substitute the random variables ψ_k by the coordinates $y_k \in \mathbb{R}$. Thus, we obtain the parameterized and truncated diffusion coefficient $b: D \times \mathbb{R}^m \rightarrow \mathbb{R}$ via

$$(2.5) \quad b(\mathbf{x}, \mathbf{y}) := \sum_{k=1}^m \sqrt{\lambda_k} \varphi_k(\mathbf{x}) y_k \quad \text{and} \quad a(\mathbf{x}, \mathbf{y}) := \exp(b(\mathbf{x}, \mathbf{y}))$$

for all $\mathbf{x} \in D$ and $\mathbf{y} = (y_1, y_2, \dots, y_m) \in \mathbb{R}^m$. In particular, the variational formulation for the parametric diffusion problem reads as follows:

$$(2.6) \quad \begin{aligned} &\text{find } u \in L_{\boldsymbol{\rho}}^2(\mathbb{R}^m; H_0^1(D)) \text{ such that} \\ &\quad -\operatorname{div}_{\mathbf{x}}(a(\mathbf{x}, \mathbf{y}) \nabla_{\mathbf{x}} u(\mathbf{x}, \mathbf{y})) = f(\mathbf{x}) \text{ in } D \text{ for all } \mathbf{y} \in \mathbb{R}^m. \end{aligned}$$

Here and in the sequel, for a given Banach space X , the Bochner space $L_{\boldsymbol{\rho}}^p(\mathbb{R}^m; X)$, $1 \leq p \leq \infty$, consists of all equivalence classes of strongly measurable functions $v: \mathbb{R}^m \rightarrow X$ whose norm

$$\|v\|_{L_{\boldsymbol{\rho}}^p(\mathbb{R}^m; X)} := \begin{cases} \left(\int_{\mathbb{R}^m} \|v(\cdot, \mathbf{y})\|_X^p \boldsymbol{\rho}(\mathbf{y}) \, d\mathbf{y} \right)^{1/p}, & p < \infty \\ \operatorname{ess\,sup}_{\mathbf{y} \in \mathbb{R}^m} \|v(\cdot, \mathbf{y})\|_X, & p = \infty \end{cases}$$

is finite. If $p = 2$ and X is a separable Hilbert space, then the Bochner space is isomorphic to the tensor product space $L_{\boldsymbol{\rho}}^2(\mathbb{R}^m) \otimes X$. Note that, for notational convenience, we will always write $v(\mathbf{x}, \mathbf{y})$ instead of $(v(\mathbf{y}))(\mathbf{x})$ if $v \in L_{\boldsymbol{\rho}}^p(\mathbb{R}^m; X)$.

The stochastic diffusion coefficient $a(\mathbf{x}, \mathbf{y})$ is neither uniformly bounded away from zero nor uniformly bounded from above for all $\mathbf{y} \in \mathbb{R}^m$. Consequently, it is impossible to show unique solvability in the classical way for elliptic boundary value problems. Especially the Lax-Milgram theorem does not directly apply to the problem (2.1). Nevertheless, in [37], it is shown that the set

$$(2.7) \quad \Gamma := \left\{ \mathbf{y} \in \mathbb{R}^m : \sum_{k=1}^m \gamma_k |y_k| < \infty \right\}$$

is of measure $\mathbb{P}_{\boldsymbol{\psi}}(\Gamma) = 1$ for all $m \leq \infty$. Moreover, for all $\mathbf{y} \in \Gamma$, the diffusion coefficient satisfies

$$(2.8) \quad 0 < a_{\min}(\mathbf{y}) \leq \operatorname{ess\,inf}_{\mathbf{x} \in D} a(\mathbf{x}, \mathbf{y}) \leq \operatorname{ess\,sup}_{\mathbf{x} \in D} a(\mathbf{x}, \mathbf{y}) \leq a_{\max}(\mathbf{y}) < \infty$$

with

$$(2.9) \quad a_{\min}(\mathbf{y}) = \exp\left(-\sum_{k=1}^m \gamma_k |y_k|\right) \quad \text{and} \quad a_{\max}(\mathbf{y}) = \exp\left(\sum_{k=1}^m \gamma_k |y_k|\right).$$

It is convenient to make in the sequel use of the abbreviation

$$(2.10) \quad \kappa(\mathbf{y}) := \frac{a_{\max}(\mathbf{y})}{a_{\min}(\mathbf{y})} = \exp\left(2 \sum_{k=1}^m \gamma_k |y_k|\right).$$

Remark 2.1. *In the framework of [37], the situation $m \rightarrow \infty$ is considered. Then, the restriction (2.3) of the parameter domain ensures for all $\mathbf{y} \in \Gamma$ that $|b(\mathbf{x}, \mathbf{y})| < \infty$ holds uniformly in $\mathbf{x} \in D$. Obviously, we have $\Gamma = \mathbb{R}^m$ for all $m < \infty$.*

Remark 2.2. *In the following, the Sobolev space $H_0^1(D)$ is considered to be equipped with the norm $\|\cdot\|_{H^1(D)} := \|\nabla \cdot\|_{L^2(D)}$. Likewise, we use corresponding norms for the Sobolev spaces $W_0^{1,p}(D)$, i.e. $\|\cdot\|_{W^{1,p}(D)} := \|\nabla \cdot\|_{L^p(D)}$. Since we take here only homogenous Dirichlet data into account, by Sobolev's norm equivalence theorem, cf. [1], they all induce norms that are equivalent to the standard norms for these spaces. Of course, all results are straightforwardly extendable to the case of non-homogenous Dirichlet problems.*

Due to (2.8), for every fixed $\mathbf{y} \in \mathbb{R}^m$, the problem (2.6) is elliptic and admits a unique solution $u(\mathbf{y}) \in H_0^1(D)$ which satisfies

$$(2.11) \quad \|u(\mathbf{y})\|_{H^1(D)} \lesssim \frac{1}{a_{\min}(\mathbf{y})} \|f\|_{L^2(D)}.$$

We refer the reader to e.g. [37] for a proof of this result. The constant involved here arises from the estimate $\|f\|_{H^{-1}(D)} \leq C(D) \|f\|_{L^2(D)}$.

The goal of this article is to compute the moments $\mathcal{M}^p u(\mathbf{x}) := \mathbb{E}[u(\mathbf{x}, \mathbf{y})^p]$ of the solution to (2.6). Especially, the solution's mean

$$(2.12) \quad \mathbb{E}_u(\mathbf{x}) := \mathbb{E}[u(\mathbf{x}, \mathbf{y})] = \int_{\mathbb{R}^m} u(\mathbf{x}, \mathbf{y}) \boldsymbol{\rho}(\mathbf{y}) \, d\mathbf{y} \in H_0^1(D)$$

and its variance

$$(2.13) \quad \mathbb{V}_u(\mathbf{x}) := \mathbb{E}_{u^2}(\mathbf{x}) - \mathbb{E}_u^2(\mathbf{x}) = \int_{\mathbb{R}^m} u^2(\mathbf{x}, \mathbf{y}) \boldsymbol{\rho}(\mathbf{y}) \, d\mathbf{y} - \mathbb{E}_u^2(\mathbf{x}) \in W_0^{1,1}(D)$$

are of interest to us. They correspond to the first and the second (centered) moment of the solution u . Notice that the knowledge of all moments is sufficient to determine the solution's distribution.

Remark 2.3. *In contrast to e.g. [11, 12, 40], we do not aim at evaluating functionals of the solution. Our goal is the approximation of its p -th moment which is the trace of the p -point correlation. The computation of the p -point correlation by the multilevel Monte-Carlo method is analyzed in [3]. Therein, a sparse tensor product approximation is used for the p -point correlation since it arises from the p -fold tensorization of equation (2.1) and is therefore contained in $[H_0^1(D)]^p$. Of course, the p -point correlation contains more information than the p -th moment of u , but the computational complexity can be reduced by sampling the moments of u directly if one is only interested in them. Hence, it is necessary to analyze the regularity of the p -th power of u which is performed in Section 5.*

3. Finite element approximation in the spatial variable. For the spatial discretization of the diffusion problem under consideration, we will employ multilevel finite elements. To that end, we consider a coarse grid triangulation $\mathcal{T}_0 = \{\tau_{0,k}\}$ of the domain D . Then, for $\ell \geq 1$, a uniform and shape regular triangulation $\mathcal{T}_\ell = \{\tau_{\ell,k}\}$ is recursively obtained by uniformly refining each simplex $\tau_{\ell-1,k}$ into 2^n simplices with diameter $h_\ell \approx 2^{-\ell}$. For $d \geq 1$, we define the finite element spaces on level ℓ by

$$\mathcal{S}_\ell^d(D) := \{v \in C(D) : v|_{\partial D} = 0 \text{ and } v|_\tau \in \mathcal{P}_d \text{ for all } \tau \in \mathcal{T}_\ell\} \subset H_0^1(D),$$

where \mathcal{P}_d denotes the space of all polynomials of total degree d . In the subsequent analysis, we restrict ourselves to piecewise linear finite elements, i.e. $d = 1$. Nevertheless, we emphasize that, by performing obvious modifications, all results remain valid also for higher order finite elements.

Now, given $\mathbf{y} \in \mathbb{R}^m$, we define the Galerkin projection $G_\ell(\mathbf{y}) : H_0^1(D) \rightarrow \mathcal{S}_\ell^1(D)$ via the Galerkin orthogonality:

$$\int_D a(\mathbf{y}) \nabla(v - G_\ell(\mathbf{y})v) \nabla w \, d\mathbf{x} = 0 \text{ for all } w \in \mathcal{S}_\ell^1(D).$$

Moreover, we set $G_{-1}(\mathbf{y}) := 0$ for all $\mathbf{y} \in \mathbb{R}^m$. In the sequel, letters in the German type setting will always refer to a Galerkin projection, i.e.

$$\mathbf{v}_\ell(\mathbf{y}) := G_\ell(\mathbf{y})v \in \mathcal{S}_\ell^1(D).$$

The Galerkin projection $\mathbf{u}_\ell(\mathbf{y})$ of the solution $u(\mathbf{y})$ to the diffusion problem (2.6) is known to fulfill the following error estimate.

Lemma 3.1. *Let the domain D be convex, $f \in L^2(D)$ and $a(\mathbf{y}) \in C^1(\overline{D})$. Then, the Galerkin projection $\mathbf{u}_\ell(\mathbf{y}) \in \mathcal{S}_\ell^1(D)$ of the diffusion problem (2.6) satisfies the error estimate*

$$(3.1) \quad \|u(\mathbf{y}) - \mathbf{u}_\ell(\mathbf{y})\|_{H^1(D)} \lesssim 2^{-\ell} \sqrt{\kappa(\mathbf{y})} \|u(\mathbf{y})\|_{H^2(D)},$$

where $\kappa(\mathbf{y})$ is given by (2.10). Moreover, if $f \in L^p(D)$ for given $p \geq 2$, then it holds $\mathbf{u}_\ell^p(\mathbf{y}) \in \mathcal{S}_\ell^p(D)$ with

$$(3.2) \quad \|(u^p - \mathbf{u}_\ell^p)(\mathbf{y})\|_{W^{1,1}(D)} \lesssim 2^{-\ell} \kappa(\mathbf{y})^p \|u(\mathbf{y})\|_{W^{2,p}(D)}^p.$$

Here, the constants hidden in (3.1) depend on D and in (3.2) additionally on p but not on $\mathbf{y} \in \mathbb{R}^m$.

Proof. The parametric diffusion problem (2.6) is H^2 -regular since D is convex and $f \in L^2(D)$. Hence, the first error estimate immediately follows from the standard finite element theory. We further find that for functions $v_1, \dots, v_p \in W^{1,p}(D)$ it holds $\|v_1 \cdots v_p\|_{W^{1,1}(D)} \leq \sum_{k=1}^p \|\nabla v_k \prod_{i \neq k} v_i\|_{L^1(D)}$. By the generalized Hölder inequality, we obtain $\|v_1 \cdots v_p\|_{W^{1,1}(D)} \lesssim \prod_{k=1}^p \|v_k\|_{W^{1,p}(D)}$, where the hidden constant depends on p and D . This yields

$$\begin{aligned} \|(u^p - \mathbf{u}_\ell^p)(\mathbf{y})\|_{W^{1,1}(D)} &= \left\| (u - \mathbf{u}_\ell)(\mathbf{y}) \sum_{i=0}^{p-1} u^i(\mathbf{y}) \mathbf{u}_\ell^{p-1-i}(\mathbf{y}) \right\|_{W^{1,1}(D)} \\ &\leq \sum_{i=0}^{p-1} \|(u - \mathbf{u}_\ell)(\mathbf{y})\|_{W^{1,p}(D)} \|u(\mathbf{y})\|_{W^{1,p}(D)}^i \|\mathbf{u}_\ell(\mathbf{y})\|_{W^{1,p}(D)}^{p-i-1}. \end{aligned}$$

By using the estimate $\|(u - \mathbf{u}_\ell)(\mathbf{y})\|_{W^{1,p}(D)} \lesssim 2^{-\ell} \kappa(\mathbf{y}) \|u(\mathbf{y})\|_{W^{2,p}(D)}$, cf. [7], it follows

$$\|\mathbf{u}_\ell(\mathbf{y})\|_{W^{1,p}(D)} \leq \|u(\mathbf{y})\|_{W^{1,p}(D)} + \|(u - \mathbf{u}_\ell)(\mathbf{y})\|_{W^{1,p}(D)} \lesssim (1 + \kappa(\mathbf{y})2^{-\ell}) \|u(\mathbf{y})\|_{W^{2,p}(D)}.$$

Putting this into the previous estimate, we finally arrive at (3.2). ■

4. Multilevel quadrature. Let X denote some Banach space of functions which are defined on the domain D , for example $X = H_0^1(D)$ or $X = W_0^{1,1}(D)$. For X and a weight $w : \mathbb{R}^m \rightarrow \mathbb{R}$, we define the weighted space, cf. [2],

$$(4.1) \quad C_w^0(\mathbb{R}^m; X) := \left\{ v : \mathbb{R}^m \rightarrow X : v \text{ is continuous and } \sup_{\mathbf{y} \in \mathbb{R}^m} \|w(\mathbf{y})v(\cdot, \mathbf{y})\|_X < \infty \right\},$$

equipped with the norm $\|v\|_{C_w^0(\mathbb{R}^m; X)} := \sup_{\mathbf{y} \in \mathbb{R}^m} \|w(\mathbf{y})v(\cdot, \mathbf{y})\|_X$. Here, for the sake of simplicity, it suffices to consider $w \equiv 1$. Later on, we need to specify the weight to establish regularity results for the solution u to (2.6).

The crucial idea of the multilevel quadrature is a finite dimensional approximation of the mapping

$$\mathbb{E} : C_w^0(\mathbb{R}^m; X) \rightarrow X, \quad v \mapsto \mathbb{E}_v.$$

To that end, we have to combine an appropriate quadrature rule for the stochastic variable with the multilevel finite element discretization in the spatial variable. More precisely, for a function $v \in C_w^0(\mathbb{R}^m; X)$, we perform a multilevel splitting of \mathbb{E}_v in X and approximate each level with a level dependent quadrature accuracy. This accuracy is chosen contraddiractional to the approximation power of the finite element spaces for the spatial domain.

For the approximation in the stochastic variable \mathbf{y} , we shall hence provide a sequence of quadrature formulae $\{\mathbf{Q}_\ell\}$ for the Bochner integral $\mathbb{E}_v(\mathbf{x}) = \int_{\mathbb{R}^m} v(\mathbf{x}, \mathbf{y}) \rho(\mathbf{y}) \, d\mathbf{y}$ of the form

$$\mathbf{Q}_\ell : C_w^0(\mathbb{R}^m; X) \rightarrow X, \quad \mathbf{Q}_\ell v = \sum_{i=1}^{N_\ell} \omega_{\ell,i} v(\cdot, \boldsymbol{\xi}_{\ell,i}).$$

For our purposes, we assume that the number of points N_ℓ of the quadrature formula \mathbf{Q}_ℓ is chosen such that the corresponding accuracy is

$$(4.2) \quad \varepsilon_\ell = 2^{-\ell}.$$

The multilevel quadrature of the mean of a function $v \in C_w^0(\mathbb{R}^m; X)$ is now defined by

$$(4.3) \quad \mathbb{E}_v(\mathbf{x}) \approx \sum_{\ell=0}^j \mathbf{Q}_{j-\ell}(\mathbf{v}_\ell - \mathbf{v}_{\ell-1})(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^j \sum_{i=0}^{N_{j-\ell}} \omega_{j-\ell,i} (\mathbf{v}_\ell - \mathbf{v}_{\ell-1})(\mathbf{x}, \boldsymbol{\xi}_{j-\ell,i}).$$

The higher order moments are approximated in complete analogy by

$$(4.4) \quad \mathcal{M}_v^p(\mathbf{x}) \approx \sum_{\ell=0}^j \mathbf{Q}_{j-\ell}(\mathbf{v}_\ell^p - \mathbf{v}_{\ell-1}^p)(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^j \sum_{i=0}^{N_{j-\ell}} \omega_{j-\ell,i} (\mathbf{v}_\ell^p - \mathbf{v}_{\ell-1}^p)(\mathbf{x}, \boldsymbol{\xi}_{j-\ell,i}).$$

Since the multilevel quadrature (4.3) and (4.4) can be interpreted as a sparse-grid approach, cf. [23], it is known that mixed regularity results of the integrand have to be provided.

5. Regularity of the solution. As motivated before, we consider in this section the regularity of u and its square u^2 . Under certain regularity assumptions it is also possible to obtain bounds for arbitrary powers of the solution, i.e. u^p for $p \in \mathbb{N}$. The regularity of the solution u has already been addressed in [2, 10, 29, 37]. We will adapt here some of the results, which originate from those articles, for our framework.

Remark 5.1. *The constants hidden in this section appear due to the repeated application of Poincaré's inequality. The dependency of the constants on p is also suppressed. In most cases, they depend exponentially on p but they are independent of the stochastic dimension m .*

Since the diffusion coefficient $a(\mathbf{x}, \mathbf{y})$ is not uniformly elliptic with respect to \mathbf{y} , we cannot expect the solution to be uniformly bounded in \mathbf{y} . Thus, the solution u to (2.6) may not be contained in the Bochner space $C_w^k(\mathbb{R}^m; H_0^1(D))$ with $w \equiv 1$. Nonetheless, we can multiply u by an appropriate auxiliary weight and end up with a bounded product in the sense of (4.1).

Definition 5.2. *For $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m) \in \mathbb{R}_+^m$, we define the auxiliary weight $\boldsymbol{\sigma}(\mathbf{y}) := \prod_{k=1}^m \sigma_k(y_k)$, where $\sigma_k(y_k) := \exp(-\beta_k |y_k|)$. For the special case $\boldsymbol{\beta} = \boldsymbol{\gamma}$ with $\boldsymbol{\gamma} := (\gamma_1, \dots, \gamma_m)$, cf. (2.3), the auxiliary weight is denoted by $\boldsymbol{\sigma}_{\min}(\mathbf{y})$.*

If we choose $w = \boldsymbol{\sigma}$ in (4.1), the space $C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; X)$ is a subset of $L_{\boldsymbol{\rho}}^p(\mathbb{R}^m; X)$ for all $p \in \mathbb{N}$. This fact issues from

$$\|v\|_{L_{\boldsymbol{\rho}}^p(\mathbb{R}^m; X)} \leq \|v\|_{C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; X)} \left(\int_{\mathbb{R}^m} (\boldsymbol{\sigma}(\mathbf{y}))^{-p} \boldsymbol{\rho}(\mathbf{y}) \, d\mathbf{y} \right)^{\frac{1}{p}} < \infty$$

because of $p\beta_k y_k = \mathfrak{o}(y_k^2)$ for $y_k \rightarrow \infty$ and the integrability of the normal distribution's tails. Especially, we will employ several times the continuity of the mapping

$$\mathbb{E} : C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; X) \rightarrow X, \quad v \mapsto \mathbb{E}_v,$$

which satisfies

$$\|\mathbb{E}v\|_X \leq C(\boldsymbol{\sigma}) \|v\|_{C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; X)} \quad \text{with} \quad C(\boldsymbol{\sigma}) = \int_{\mathbb{R}^m} (\boldsymbol{\sigma}(\mathbf{y}))^{-1} \boldsymbol{\rho}(\mathbf{y}) \, d\mathbf{y}.$$

The constant $C(\boldsymbol{\sigma})$ depends on the choice of $\boldsymbol{\beta}$ and may grow exponentially in m since the domain of integration is \mathbb{R}^m . This is however not necessarily the case here. If we integrate $\boldsymbol{\sigma}_{\min}$ or even the p -th power of $\boldsymbol{\sigma}_{\min}$ with respect to the Gaussian measure, we get, cf. [17],

$$(5.1) \quad \left(\int_{\mathbb{R}^m} (\boldsymbol{\sigma}_{\min}(\mathbf{y}))^{-p} \boldsymbol{\rho}(\mathbf{y}) \, d\mathbf{y} \right)^{\frac{1}{p}} \leq \exp \left(p^2 \sum_{k=1}^m \gamma_k^2 + p \sqrt{\frac{2}{\pi}} \sum_{k=1}^m \gamma_k \right).$$

This expression depends only on the decay of the sequence $\{\gamma_k\}_k$ and is bounded independently of the dimension m due to (2.3).

Proposition 5.3. *For all weights $\boldsymbol{\sigma}$ satisfying $\boldsymbol{\beta} \geq \boldsymbol{\gamma}$, the solution u of (2.6) is contained in $C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; H_0^1(D))$ and satisfies $u^2 \in C_{\boldsymbol{\sigma}_2}^0(\mathbb{R}^m; W_0^{1,1}(D))$. In particular, it holds*

$$\|u\|_{C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; H^1(D))} \lesssim \|f\|_{L^2(D)}, \quad \|u^2\|_{C_{\boldsymbol{\sigma}_2}^0(\mathbb{R}^m; W^{1,1}(D))} \lesssim \|f\|_{L^2(D)}^2.$$

Proof. In view of inequality (2.11) and since $1/a_{\min}(\mathbf{y}) = \exp\left(\sum_{k=1}^m \gamma_k |y_k|\right)$, we conclude the first estimate

$$\boldsymbol{\sigma}(\mathbf{y})\|u(\mathbf{y})\|_{H^1(D)} \lesssim \exp\left(\sum_{k=1}^m (\gamma_k - \beta_k)|y_k|\right) \|f\|_{L^2(D)} \leq \|f\|_{L^2(D)}$$

while the second estimate follows from

$$\boldsymbol{\sigma}^2(\mathbf{y})\|u^2(\mathbf{y})\|_{W^{1,1}(D)} \leq \boldsymbol{\sigma}^2(\mathbf{y})\|2u(\mathbf{y})\nabla u(\mathbf{y})\|_{L^1(D)} \lesssim (\boldsymbol{\sigma}(\mathbf{y})\|u(\mathbf{y})\|_{H^1(D)})^2 \lesssim \|f\|_{L^2(D)}^2.$$

■

The differentiability of u follows from the differentiability of $a(\mathbf{x}, \mathbf{y})$. An estimate on multivariate derivatives has been proven in [4] and was extended to our problem setting in [29]. We use here the related lemma from [29], which is adjusted for our purposes.

Lemma 5.4. *For every $\mathbf{y} \in \mathbb{R}^m$, it holds that*

$$\|\partial_{y_k}^j u(\mathbf{y})\|_{H^1(D)} \leq j! \left(\frac{\gamma_k}{\log 2}\right)^j \sqrt{\kappa(\mathbf{y})} \|u(\mathbf{y})\|_{H^1(D)}.$$

With this lemma at hand, we are able to show the following result.

Proposition 5.5. *For all weights $\boldsymbol{\sigma}$ with $\boldsymbol{\beta} \geq 2\boldsymbol{\gamma}$, the partial derivatives of the solution u to (2.6) satisfy*

$$(5.2) \quad \|\partial_{y_k}^j u\|_{C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; H^1(D))} \lesssim j! \left(\frac{\gamma_k}{\log 2}\right)^j \|f\|_{L^2(D)}.$$

Epecially, it holds $\partial_{y_k}^j u \in C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; H_0^1(D))$.

Proof. From Lemma 5.4, we obtain

$$\|\partial_{y_k}^j u\|_{C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; H^1(D))} = \sup_{\mathbf{y} \in \mathbb{R}^m} \|\boldsymbol{\sigma}(\mathbf{y}) \partial_{y_k}^j u(\mathbf{y})\|_{H^1(D)} \leq j! \left(\frac{\gamma_k}{\log 2}\right)^j \sup_{\mathbf{y} \in \mathbb{R}^m} \sqrt{\kappa(\mathbf{y})} \boldsymbol{\sigma}(\mathbf{y}) \|u(\mathbf{y})\|_{H^1(D)}.$$

In view of (2.10), we have $\sqrt{\kappa(\mathbf{y})} = \exp\left(\sum_{k=1}^m \gamma_k |y_k|\right)$. This yields the desired estimate as follows

$$\begin{aligned} \|\partial_{y_k}^j u\|_{C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; H^1(D))} &\leq j! \left(\frac{\gamma_k}{\log 2}\right)^j \sup_{\mathbf{y} \in \mathbb{R}^m} \exp\left(\sum_{k=1}^m (\gamma_k - \beta_k)|y_k|\right) \|u(\mathbf{y})\|_{H^1(D)} \\ &\leq j! \left(\frac{\gamma_k}{\log 2}\right)^j \sup_{\mathbf{y} \in \mathbb{R}^m} \exp\left(\sum_{k=1}^m -\gamma_k |y_k|\right) \|u(\mathbf{y})\|_{H^1(D)} \lesssim j! \left(\frac{\gamma_k}{\log 2}\right)^j \|f\|_{L^2(D)}. \end{aligned}$$

■

The previous result shows the regularity of the solution u . In the following proposition, we consider the regularity of u^2 .

Proposition 5.6. *The partial derivatives of u^2 , where u is the solution of (2.6), satisfy $\partial_{y_k}^j u^2 \in C_{\sigma^2}^0(\mathbb{R}^m; W_0^{1,1}(D))$ for all σ with $\beta \geq 2\gamma$. Especially, it holds*

$$(5.3) \quad \|\partial_{y_k}^j u^2\|_{C_{\sigma^2}^0(\mathbb{R}^m; W^{1,1}(D))} \lesssim (j+1)! \left(\frac{\gamma_k}{\log 2}\right)^j \|f\|_{L^2(D)}^2 \leq j! \left(\frac{2\gamma_k}{\log 2}\right)^j \|f\|_{L^2(D)}^2.$$

Proof. By the Leibniz rule, we obtain

$$(5.4) \quad \|\partial_{y_k}^j u^2(\mathbf{y})\|_{W^{1,1}(D)} \leq \sum_{\ell=0}^j \binom{j}{\ell} \|\partial_{y_k}^\ell u(\mathbf{y}) \partial_{y_k}^{j-\ell} u(\mathbf{y})\|_{W^{1,1}(D)}.$$

Each summand on the right hand side can be estimated as follows:

$$\begin{aligned} & \|\partial_{y_k}^\ell u(\mathbf{y}) \partial_{y_k}^{j-\ell} u(\mathbf{y})\|_{W^{1,1}(D)} \\ &= \|\nabla \partial_{y_k}^\ell u(\mathbf{y}) \partial_{y_k}^{j-\ell} u(\mathbf{y}) + \partial_{y_k}^\ell u(\mathbf{y}) \nabla \partial_{y_k}^{j-\ell} u(\mathbf{y})\|_{L^1(D)} \\ &\leq \|\nabla \partial_{y_k}^\ell u(\mathbf{y})\|_{L^2(D)} \|\partial_{y_k}^{j-\ell} u(\mathbf{y})\|_{L^2(D)} + \|\partial_{y_k}^\ell u(\mathbf{y})\|_{L^2(D)} \|\nabla \partial_{y_k}^{j-\ell} u(\mathbf{y})\|_{L^2(D)} \\ &\lesssim \|\partial_{y_k}^\ell u(\mathbf{y})\|_{H^1(D)} \|\partial_{y_k}^{j-\ell} u(\mathbf{y})\|_{H^1(D)} + \|\partial_{y_k}^\ell u(\mathbf{y})\|_{H^1(D)} \|\partial_{y_k}^{j-\ell} u(\mathbf{y})\|_{H^1(D)}. \end{aligned}$$

Application of Lemma 5.4 yields

$$\|\partial_{y_k}^\ell u(\mathbf{y}) \partial_{y_k}^{j-\ell} u(\mathbf{y})\|_{W^{1,1}(D)} \lesssim 2\ell!(j-\ell)! \left(\frac{\gamma_k}{\log 2}\right)^j \kappa(\mathbf{y}) \|u(\mathbf{y})\|_{H^1(D)}^2.$$

Inserting this inequality into (5.4) results in

$$\|\partial_{y_k}^j u^2(\mathbf{y})\|_{W^{1,1}(D)} \lesssim 2 \sum_{\ell=0}^j j! \left(\frac{\gamma_k}{\log 2}\right)^j \kappa(\mathbf{y}) \|u(\mathbf{y})\|_{H^1(D)}^2 \leq 2(j+1)! \left(\frac{\gamma_k}{\log 2}\right)^j \kappa(\mathbf{y}) \|u(\mathbf{y})\|_{H^1(D)}^2.$$

The estimate (5.3) is now obtained analogously to estimate (5.2) in Proposition 5.5. \blacksquare

Remark 5.7. *Given that the load satisfies $f \in L^p(D)$ for $p > 2$, the solution $u(\mathbf{y})$ to (2.6) is contained in $W_0^{1,p}(D)$ and satisfies the regularity estimate*

$$\|u(\mathbf{y})\|_{W^{1,p}(D)} \lesssim \frac{1}{a_{\min}(\mathbf{y})} \|f\|_{L^p(D)}.$$

The derivatives of u with respect to the parametric variable \mathbf{y} can be estimated by

$$\|\partial_{y_k}^j u(\mathbf{y})\|_{W^{1,p}(D)} \lesssim j! \left(\frac{C(p, D)\gamma_k}{\log 2}\right)^j \kappa(\mathbf{y}) \|u(\mathbf{y})\|_{W^{1,p}(D)}.$$

The constant $C(p, D)$ stems from the dense embedding of $W_0^{1,q}(D)$ into $[L^q(D)]^d$ by the mapping $v \mapsto \nabla v$, cf. [38], for the dual exponent q of p . Additionally, the derivatives of the powers u^p with respect to the parametric variable \mathbf{y} fulfill

$$\|\partial_{y_k}^j u^p(\mathbf{y})\|_{W^{1,1}(D)} \lesssim j! \left(\frac{pC(p, D)\gamma_k}{\log 2}\right)^j \kappa(\mathbf{y})^p \|u(\mathbf{y})\|_{W^{1,p}(D)}^p.$$

For all weights σ with $\beta \geq 3\gamma$ this leads to

$$(5.5) \quad \|\partial_{y_k}^j u^p\|_{C_{\sigma^p}^0(\mathbb{R}^m; W^{1,1}(D))} \lesssim j! \left(\frac{pC(p, D)\gamma_k}{\log 2} \right)^j \|f\|_{L^p(D)}.$$

A proof of this remark is found in [24] for multivariate derivatives.

Lemma 5.4 provides only a bound on the derivatives of $\partial_{y_k}^j u$ when the spatial regularity is measured in $H_0^1(D)$. We shall thus complete this section by a result from [29] which establishes estimates on $\partial_{y_k}^j u$ when the spatial regularity is measured in the space $\mathcal{W} := H^2(D) \cap H_0^1(D)$. This result guarantees us the mixed regularity which is necessary for the sparse-grid construction between the spatial and the stochastic variable. To establish this result, we shall assume that the corresponding eigenfunctions of the Karhunen-Loève expansion (2.2) belong to $W^{1,\infty}(D)$ which is for example fulfilled in case of a Gaussian covariance. If we then replace γ_k by

$$(5.6) \quad \tilde{\gamma}_k := \gamma_k + \sqrt{\lambda_k} \|\nabla \varphi_k\|_{L^\infty(D)} = \sqrt{\lambda_k} (\|\varphi_k\|_{L^\infty(D)} + \|\nabla \varphi_k\|_{L^\infty(D)})$$

in the definition of the set Γ , cf. (2.7), and if we assume that $\{\tilde{\gamma}_k\}_k \in \ell^1(\mathbb{N})$, we still have the parameter domain \mathbb{R}^m for each finite $m \in \mathbb{N}$. Hence, this sharpened condition induces no further restriction to the parameter domain. In analogy to (2.9) and (2.10), it is useful to define

$$(5.7) \quad \tilde{a}_{\min}(\mathbf{y}) := \exp\left(-\sum_{k=1}^m \tilde{\gamma}_k |y_k|\right), \quad \tilde{a}_{\max}(\mathbf{y}) := \exp\left(\sum_{k=1}^m \tilde{\gamma}_k |y_k|\right), \quad \tilde{\kappa}(\mathbf{y}) := \frac{\tilde{a}_{\max}(\mathbf{y})}{\tilde{a}_{\min}(\mathbf{y})}.$$

Furthermore, we will employ spaces $C_{\sigma}^0(\mathbb{R}^m; X)$, see (4.1) and Definition 5.2, where the auxiliary weight σ is defined with respect to $\tilde{\gamma}$ instead of γ .

For convex or sufficiently smooth curved domains, a norm on \mathcal{W} is given by

$$\|u\|_{\mathcal{W}} := \|\nabla u\|_{L^2(D)} + \|\Delta u\|_{L^2(D)},$$

cf. [29]. Along the lines of [29], we have the following

Proposition 5.8. *For all $\mathbf{y} \in \mathbb{R}^m$, the solution to problem (2.6) satisfies*

$$\|\sqrt{a(\mathbf{y})} \Delta \partial_{y_k}^j u(\mathbf{y})\|_{L^2(D)} \lesssim j! \left(\frac{2\tilde{\gamma}_k}{\log 2} \right)^j \left(\|\sqrt{a(\mathbf{y})}^{-1} f\|_{L^2(D)} + 2g(\mathbf{y}) \|\sqrt{a(\mathbf{y})} u(\mathbf{y})\|_{H^1(D)} \right)$$

with $g(\mathbf{y}) := 1 + 2 \sum_{k=1}^m |y_k| \sqrt{\lambda_k} \|\nabla \varphi_k\|_{L^\infty(D)} < \infty$.

This proposition implies the estimate

$$\|\Delta \partial_{y_k}^j u(\mathbf{y})\|_{L^2(D)} \lesssim \sqrt{\kappa(\mathbf{y})} j! \left(\frac{2\tilde{\gamma}_k}{\log 2} \right)^j \left(\|f\|_{L^2(D)} + 2g(\mathbf{y}) \|u(\mathbf{y})\|_{H^1(D)} \right),$$

which can be further estimated by

$$\|\Delta \partial_{y_k}^j u(\mathbf{y})\|_{L^2(D)} \lesssim \kappa(\mathbf{y}) g(\mathbf{y}) j! \left(\frac{2\tilde{\gamma}_k}{\log 2} \right)^j \|f\|_{L^2(D)}.$$

due to (2.9) and (2.11). It follows together with Lemma 5.4 and with

$$\begin{aligned} \kappa(\mathbf{y})g(\mathbf{y}) &= \exp\left(2\sum_{k=1}^m \gamma_k |y_k|\right) \left(1 + 2\sum_{k=1}^m |y_k| \sqrt{\lambda_k} \|\nabla \varphi_k\|_{L^\infty(D)}\right) \\ &\leq \exp\left(2\sum_{k=1}^m \gamma_k |y_k|\right) \exp\left(2\sum_{k=1}^m |y_k| \sqrt{\lambda_k} \|\nabla \varphi_k\|_{L^\infty(D)}\right) = \tilde{\kappa}(\mathbf{y}) \end{aligned}$$

that

$$(5.8) \quad \|\partial_{y_k}^j u(\mathbf{y})\|_{H^2(D)} \lesssim \|\partial_{y_k}^j u(\mathbf{y})\|_{\mathcal{W}} \lesssim \tilde{\kappa}(\mathbf{y}) j! \left(\frac{2\tilde{\gamma}_k}{\log 2}\right)^j \|f\|_{L^2(D)}.$$

This establishes the following

Proposition 5.9. *The solution u to (2.6) is contained in $C_\sigma^0(\mathbb{R}^m; H^2(D))$ for all σ with $\beta \geq 2\tilde{\gamma}$ and it holds*

$$(5.9) \quad \|\partial_{y_k}^j u\|_{C_\sigma^0(\mathbb{R}^m; H^2(D))} \lesssim j! \left(\frac{2\tilde{\gamma}_k}{\log 2}\right)^j \|f\|_{L^2(D)}.$$

Remark 5.10. *All estimates in this section are given for univariate derivatives of the solution u to (2.6). They can easily be extended to multidimensional derivatives $\partial_{\mathbf{y}}^\alpha u$. The factor $j!(\gamma_k/\log 2)^j$ which appears in the estimates for the j -th derivative in the k -th direction in Lemma 5.4 is then replaced by $|\alpha|!(\gamma/\log 2)^\alpha$, see [4, 29]. Furthermore, multivariate versions of Proposition 5.6 and Remark 5.7 are proven in [24] and of Proposition 5.8 in [29], respectively.*

6. Gauss-Hermite quadrature for the stochastic variable. For given sets of points

$$(6.1) \quad \{\eta_0^{(k)}, \dots, \eta_{N_k}^{(k)}\} \subset \mathbb{R}, \quad N_k \in \mathbb{N}, \quad k = 1, \dots, m,$$

the Lagrangian basis polynomials $L_0^{(k)}, \dots, L_{N_k}^{(k)}$ of degree N_k are uniquely determined by the property $L_i^{(k)}(\eta_j^{(k)}) = \delta_{i,j}$. Thus, for a multi-index $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathcal{J}$ with $\mathcal{J} := \times_{k=1}^m \{0, \dots, N_k\}$, we define the corresponding tensor product Lagrangian basis polynomial

$$\mathbf{L}_\alpha(\mathbf{y}) := \prod_{k=1}^m L_{\alpha_k}^{(k)}(y_k) \quad \text{with} \quad \mathbf{L}_\alpha(\boldsymbol{\eta}_{\alpha'}) = \delta_{\alpha, \alpha'} \quad \text{for} \quad \boldsymbol{\eta}_\alpha := (\eta_{\alpha_1}^{(1)}, \dots, \eta_{\alpha_m}^{(m)}).$$

Given a continuous function $v: \mathbb{R}^m \rightarrow \mathbb{R}$, we introduce the associated interpolation operator by

$$(\mathbf{\Pi}_{\mathcal{J}} v)(\mathbf{y}) := \sum_{\alpha \in \mathcal{J}} v(\boldsymbol{\eta}_\alpha) \mathbf{L}_\alpha(\mathbf{y}).$$

If we choose for each k the sets (6.1) to be the roots of the Hermite polynomials of degree $N_k + 1$, which are known to be orthogonal with respect to the inner product

$$(q, r)_{L_\rho^2(\mathbb{R})} = \int_{\mathbb{R}} q(y)r(y)\rho(y) \, dy, \quad q, r \in L_\rho^2(\mathbb{R}),$$

we straightforwardly derive the associated Gaussian quadrature formula from the interpolation operator. These quadrature rules are known to be exact of degree $2N_k + 1$. Further, one easily verifies

$$(L_i^{(k)}, L_j^{(k)})_{L_\rho^2(\mathbb{R})} = \omega_i^{(k)} \delta_{i,j} \quad \text{with } \omega_i^{(k)} = (L_i^{(k)}, 1)_{L_\rho^2(\mathbb{R})}.$$

By tensor product construction, we get the multivariate weights $\boldsymbol{\omega}_\alpha := \prod_{k=1}^m \omega_{\alpha_k}^{(k)}$. Then, for $\boldsymbol{\alpha}, \boldsymbol{\alpha}' \in \mathcal{J}$, we have the corresponding multivariate relations

$$(\mathbf{L}_\alpha, \mathbf{L}_{\alpha'})_{L_\rho^2(\mathbb{R}^m)} = \boldsymbol{\omega}_\alpha \delta_{\alpha, \alpha'} \quad \text{and } \boldsymbol{\omega}_\alpha = (\mathbf{L}_\alpha, 1)_{L_\rho^2(\mathbb{R}^m)}.$$

Now, we can interpolate the solution $u \in L_\rho^2(\mathbb{R}^m; H_0^1(D))$ of (2.6) in the stochastic variable

$$(6.2) \quad u(\mathbf{x}, \mathbf{y}) \approx (\text{Id} \otimes \Pi_{\mathcal{J}} u)(\mathbf{x}, \mathbf{y}) = \sum_{\boldsymbol{\alpha} \in \mathcal{J}} u(\mathbf{x}, \boldsymbol{\eta}_\alpha) \mathbf{L}_\alpha(\mathbf{y}).$$

With (6.2) at hand, we can approximate the solution's mean (2.12) and its variance (2.13) by

$$\mathbb{E}_u(\mathbf{x}) \approx \sum_{\boldsymbol{\alpha} \in \mathcal{J}} u(\mathbf{x}, \boldsymbol{\eta}_\alpha) \boldsymbol{\omega}_\alpha \quad \text{and} \quad \mathbb{V}_u(\mathbf{x}) \approx \sum_{\boldsymbol{\alpha} \in \mathcal{J}} u(\mathbf{x}, \boldsymbol{\eta}_\alpha)^2 \boldsymbol{\omega}_\alpha - \left(\sum_{\boldsymbol{\alpha} \in \mathcal{J}} u(\mathbf{x}, \boldsymbol{\eta}_\alpha) \boldsymbol{\omega}_\alpha \right)^2.$$

The remainder of this section is dedicated to the analysis of the occurring quadrature error for the Gauss-Hermite quadrature. To that end, we define for a function $v \in C_\sigma^0(\mathbb{R}^m; X)$ with σ as in Definition 5.2 the quadrature operator

$$\mathbf{Q}_{\mathcal{J}}: C_\sigma^0(\mathbb{R}^m; X) \rightarrow X, \quad (\mathbf{Q}_{\mathcal{J}} v)(\mathbf{x}) := \sum_{\boldsymbol{\alpha} \in \mathcal{J}} \boldsymbol{\omega}_\alpha v(\mathbf{x}, \boldsymbol{\eta}_\alpha).$$

Of course, it is possible to consider quadrature operators with respect to other weighted spaces $C_w^0(\mathbb{R}^m; X)$ analogously.

In the following, we adopt the analysis and the notation presented in [2], where the polynomial approximation error in case of stochastic collocation for uniformly elliptic equations is analyzed. In [10], this analysis has been extended to the case of log-normal diffusion coefficients. According to [2], we shall introduce the one-dimensional Gaussian auxiliary measure $\sqrt{\rho(y)} \approx \exp(-y^2/4)$ and the corresponding space $C_{\sqrt{\rho}}^0(\mathbb{R}; X)$. This norm is weaker than the norm of $C_\sigma^0(\mathbb{R}; X)$ from the previous section, that is $C_\sigma^0(\mathbb{R}; X) \subset C_{\sqrt{\rho}}^0(\mathbb{R}; X)$. On $C_{\sqrt{\rho}}^0(\mathbb{R}; X)$, we can now define the univariate quadrature operator of degree $N \in \mathbb{N}$ by

$$Q_N: C_{\sqrt{\rho}}^0(\mathbb{R}; X) \rightarrow X, \quad (Q_N v)(\mathbf{x}) := \sum_{i=0}^N \omega_i v(\mathbf{x}, \eta_i),$$

where η_0, \dots, η_N are again the $N + 1$ roots of the Hermite polynomial of degree $N + 1$.

The following two lemmata imply that the univariate quadrature error is bounded by the polynomial approximation error. They are modifications of the corresponding lemmata in [2] for the polynomial interpolation.

Lemma 6.1. *The quadrature operator $Q_N: C_{\sqrt{\rho}}^0(\mathbb{R}; X) \rightarrow X$ is continuous.*

Proof. Consider $v \in C_{\sqrt{\rho}}^0(\mathbb{R}; X)$. By using the triangle inequality and exploiting the positivity of the weights w_k of the Gauss-Hermite quadrature, we have

$$\begin{aligned} \|Q_N v\|_X &= \left\| \sum_{i=0}^N \omega_i v(\eta_i) \right\|_X \leq \sum_{i=0}^N \omega_i \|v(\eta_i)\|_X = \sum_{i=0}^N \frac{\omega_i}{\sqrt{\rho(\eta_i)}} \|\sqrt{\rho(\eta_i)} v(\eta_i)\|_X \\ &\leq \max_{i=0, \dots, N} \|\sqrt{\rho(\eta_i)} v(\eta_i)\|_X \sum_{i=0}^N \frac{\omega_i}{\sqrt{\rho(\eta_i)}} \lesssim \|v\|_{C_{\sqrt{\rho}}^0(\mathbb{R}; X)}. \end{aligned}$$

The last inequality follows from [41], where the convergence

$$\sum_{i=0}^N \frac{\omega_i}{\sqrt{\rho(\eta_i)}} \xrightarrow{N \rightarrow \infty} \int_{\mathbb{R}} \frac{\rho(y)}{\sqrt{\rho(y)}} dy < \infty$$

is shown. ■

The next lemma relates the quadrature error to the best approximation error in $C_{\sqrt{\rho}}^0(\mathbb{R}; X)$. We denote by $\mathcal{P}_d(\mathbb{R})$ the space of polynomials of degree at most d .

Lemma 6.2. *For every $v \in C_{\sqrt{\rho}}^0(\mathbb{R}; X)$, the quadrature error of the $(N+1)$ -point Gauss-Hermite quadrature is bounded by $\|\mathbb{E}_v - Q_N v\|_X \lesssim \inf_{w \in \mathcal{P}_{2N+1}(\mathbb{R}) \otimes X} \|v - w\|_{C_{\sqrt{\rho}}^0(\mathbb{R}; X)}$.*

Proof. Since the $(N+1)$ -point Gauss-Hermite quadrature has degree of exactness $2N+1$, it holds $Q_N w = \mathbb{E}_w$ for all $w \in \mathcal{P}_{2N+1}(\mathbb{R}) \otimes X$. Thus, for arbitrary $w \in \mathcal{P}_{2N+1}(\mathbb{R}) \otimes X$, we have

$$\begin{aligned} \|\mathbb{E}_v - Q_N v\|_X &\leq \|\mathbb{E}_{v-w}\|_X + \|Q_N(v-w)\|_X \lesssim \|v-w\|_{L^1_{\rho}(\mathbb{R}; X)} + \|v-w\|_{C_{\sqrt{\rho}}^0(\mathbb{R}; X)} \\ &\lesssim \|v-w\|_{C_{\sqrt{\rho}}^0(\mathbb{R}; X)}. \end{aligned}$$

■

In the next step, we show that functions $v \in C_{\sigma}^0(\mathbb{R}^m; X)$ admit an analytic extension under certain decay properties of their derivatives. This is crucial to bound the error of the best approximation in the polynomial space. Following the notation in [2], we introduce

$$\rho_k^*(\mathbf{y}_k^*) := \prod_{\substack{i=1 \\ i \neq k}}^m \rho(y_i) \quad \text{and} \quad \mathbf{y}_k^* := (y_1, \dots, y_{k-1}, y_{k+1}, \dots, y_m) \in \mathbb{R}^{m-1}.$$

Lemma 6.3. *Let $(y_k, \mathbf{y}_k^*) \in \mathbb{R}^m$, $v \in C_{\sigma}^0(\mathbb{R}^m; X)$, and assume that there holds*

$$\|\partial_{y_k}^j v\|_{C_{\sigma}^0(\mathbb{R}^m; X)} \lesssim j! \mu_k^j$$

with some constant $\mu_k \in (0, \infty)$. Then, for $\tau_k \in (0, 1/\mu_k)$, the function

$$v: \mathbb{R} \rightarrow C_{\sigma_k^*}^0(\mathbb{R}^{m-1}; X), \quad y_k \mapsto v(\mathbf{x}, y_k, \mathbf{y}_k^*)$$

admits an analytic extension $v(\mathbf{x}, z, \mathbf{y}_k^*)$ for $z \in \Sigma(\tau_k) := \{z \in \mathbb{C}, \text{dist}(z, \mathbb{R}) \leq \tau_k\}$. Moreover, the function v is bounded in the norm

$$(6.3) \quad \|v\|_{C_\sigma^0(\Sigma(\mathbb{R}, \tau_k); C_{\sigma_k^*}^0(\mathbb{R}^{m-1}; X))} := \sup_{z \in \Sigma(\mathbb{R}, \tau_k)} \sigma_k(\text{Re}(z)) \|v(z)\|_{C_{\sigma_k^*}^0(\mathbb{R}^{m-1}; X)}.$$

Proof. For given $y_k \in \mathbb{R}$, we define a formal Taylor expansion in $z \in \mathbb{C}$:

$$v(\mathbf{x}, z, \mathbf{y}_k^*) = \sum_{j=0}^{\infty} \frac{(z - y_k)^j}{j!} \partial_{y_k}^j v(\mathbf{x}, y_k, \mathbf{y}_k^*).$$

Thus, given an arbitrary $y_k \in \mathbb{R}$, we can estimate

$$\begin{aligned} \sigma_k(y_k) \|v(z)\|_{C_{\sigma_k^*}^0(\mathbb{R}^{m-1}; X)} &\leq \sum_{j=0}^{\infty} \frac{|z - y_k|^j}{j!} \sigma_k(y_k) \|\partial_{y_k}^j v(y_k)\|_{C_{\sigma_k^*}^0(\mathbb{R}^{m-1}; X)} \\ &\leq \sum_{j=0}^{\infty} \frac{|z - y_k|^j}{j!} \|\partial_{y_k}^j v\|_{C_\sigma^0(\mathbb{R}^m; X)} \lesssim \sum_{j=0}^{\infty} (|z - y_k| \mu_k)^j. \end{aligned}$$

The last expression converges for all $|z - y_k| \leq \tau_k < 1/\mu_k$. Hence, since we can cover $\Sigma(\tau_k)$ by the union of balls $|z - y_k| \leq \tau_k$, the function v can be extended analytically to the whole region $\Sigma(\tau_k)$. ■

Finally, we have to bound the approximation error of a function v which admits an analytic extension. This was done in [2], where the next lemma has been proven.

Lemma 6.4. *Suppose that $v \in C_\sigma^0(\mathbb{R}; X)$ admits an analytic extension in $\Sigma(\mathbb{R}, \tau)$ for some $\tau > 0$ and recall that $\sigma(y) = \exp(-\beta|y|)$, cf. Definition 5.2. Then, the error of the best approximation by polynomials of degree at most d can be bounded by*

$$(6.4) \quad \inf_{w \in \mathcal{P}_d(\mathbb{R}) \otimes X} \|v - w\|_{C_{\sqrt{\rho}}^0(\mathbb{R}; X)} \leq C(\tau) \sqrt{d} \exp(-\tau \sqrt{d}) \|v\|_{C_\sigma^0(\Sigma(\mathbb{R}, \tau); X)}.$$

Remark 6.5. *Notice that the constant $C(\tau)$ may grow extremely fast in τ , namely like $\exp(\tau^2/2)$ as $\tau \rightarrow \infty$ and, thus, (6.4) is not useful to exploit anisotropies in the integrand.*

We are now able to estimate the error of the tensor product Gaussian quadrature for functions $v \in C_\sigma^0(\mathbb{R}^m; X)$ which fulfill the condition of Lemma 6.3. For this purpose, we define the following tensor product integral operator:

$$\mathbf{I} := \bigotimes_{k=1}^m I_k \quad \text{with} \quad I_k v = \int_{\mathbb{R}} v(y_k) \rho(y_k) dy_k.$$

Note that $\mathbf{I}v$ coincides with \mathbb{E}_v due to the product structure of the measure $\rho d\mathbf{x}$.

Theorem 6.6. *Let $v \in C_\sigma^0(\mathbb{R}^m; X)$ satisfy the conditions of Lemma 6.4 in every direction, i.e. for all $k = 1, \dots, m$ there exists a $\tau_k > 0$ such that $v(\mathbf{x}, y_k, \mathbf{y}_k^*)$ admits as a function of y_k an analytic extension in $\Sigma(\tau_k)$. Furthermore, let the tensor product Gaussian quadrature operator*

$\mathbf{Q}_{\mathcal{J}}: C_{\sigma}^0(\Gamma; X) \rightarrow X$ as well as the integration operator $\mathbf{I}: C_{\sigma}^0(\Gamma; X) \rightarrow X$ be continuous and let $0 < \varepsilon < 1$. If we choose the number of quadrature points N_k such that

$$(6.5) \quad N_k \geq \frac{|\log \varepsilon|^2}{2(1-\delta)^2} - \frac{1}{2}$$

for some $\delta > 0$, then the quadrature error is of order $\mathcal{O}(\varepsilon)$. More precisely, it holds the estimate

$$(6.6) \quad \|(\mathbf{I} - \mathbf{Q}_{\mathcal{J}})v\|_X \lesssim \varepsilon \max_{k=1, \dots, m} \|v\|_{C_{\sigma_k}^0(\Sigma(\Gamma_k, \tau_k); C_{\sigma_k^*}^0(\Gamma_k^*; X))}.$$

Proof. Let $v \in C_{\sigma}^0(\mathbb{R}^m; X)$ be a function which fulfills the conditions of Lemma 6.3 for all directions y_k , $k = 1, \dots, m$. We estimate the tensor product quadrature error as usual by the sum of the one dimensional quadrature errors:

$$(6.7) \quad \|(\mathbf{I} - \mathbf{Q}_{\mathcal{J}})v\|_X \leq \sum_{k=1}^m \|(\mathbf{Q}_{N_1} \otimes \dots \otimes \mathbf{Q}_{N_{k-1}} \otimes (I_k - \mathbf{Q}_{N_k}) \otimes I_{k+1} \otimes \dots \otimes I_m)v\|_X.$$

With the continuity of the multivariate integral as well as the multivariate Gaussian quadrature operator, we can further deduce

$$\begin{aligned} & \|(\mathbf{Q}_{N_1} \otimes \dots \otimes \mathbf{Q}_{N_{k-1}} \otimes (I_k - \mathbf{Q}_{N_k}) \otimes I_{k+1} \otimes \dots \otimes I_m)v\|_X \\ & \lesssim \sup_{y_1 \in \mathbb{R}} \sigma_1(y_1) \|(\mathbf{Q}_{N_2} \otimes \dots \otimes \mathbf{Q}_{N_{k-1}} \otimes (I_k - \mathbf{Q}_{N_k}) \otimes I_{k+1} \otimes \dots \otimes I_m)v(\cdot, y_1)\|_X \\ & \lesssim \sup_{\mathbf{y}_k^* \in \mathbb{R}^{m-1}} \sigma_k^*(\mathbf{y}_k^*) \|(I_k - \mathbf{Q}_{N_k})v(\cdot, \mathbf{y}_k^*)\|_X \end{aligned}$$

By the Lemmata 6.2, 6.3 and 6.4, we conclude

$$(6.8) \quad \begin{aligned} & \|(\mathbf{Q}_{N_1} \otimes \dots \otimes \mathbf{Q}_{N_{k-1}} \otimes (I_k - \mathbf{Q}_{N_k}) \otimes I_{k+1} \otimes \dots \otimes I_m)v\|_X \\ & \lesssim \sup_{\mathbf{y}_k^* \in \mathbb{R}^{m-1}} \sigma_k^*(\mathbf{y}_k^*) \min_{w \in \mathcal{P}_{2N_k+1} \otimes X} \|v(\mathbf{y}_k^*) - w\|_{C_{\sqrt{\rho}}^0(\mathbb{R}; X)} \\ & \lesssim \sqrt{2N_k + 1} \exp(-\sqrt{2N_k + 1}) \|v\|_{C_{\sigma_k}^0(\Sigma(\Gamma_k, \tau_k); C_{\sigma_k^*}^0(\Gamma_k^*; X))}. \end{aligned}$$

With the choice (6.5) for the number of quadrature points, the assertion follows from

$$\sqrt{2N_k + 1} \exp(-\sqrt{2N_k + 1}) \lesssim \exp(-\sqrt{2N_k + 1}(1-\delta)) \leq \varepsilon$$

and summing up the terms in (6.8) with respect to (6.7). ■

Remark 6.7. The norm in the error estimate (6.6) obviously depends on the choice of τ_k and tends to infinity if τ_k comes close to the boundary of the analyticity region, i.e. if $\tau_k \rightarrow 1/\mu_k$, as can be seen from the proof of Lemma 6.3. Moreover, note that the regularity result (5.2) for the solution u to (2.6), the result (5.3) for the second moment u^2 , and, if $f \in L^p(D)$, the result (5.5) for the higher order moments u^p imply the conditions of Lemma 6.3 and, therefore, of Theorem 6.6.

7. (Quasi)-Monte Carlo quadrature for the stochastic variable. In this section, we discuss the use of Monte Carlo and quasi-Monte Carlo quadrature rules. These quadrature rules are classically of the form

$$\mathbf{Q}_{(Q)MC}v = \frac{1}{N} \sum_{i=1}^N v(\boldsymbol{\xi}_i),$$

where N denotes the number of *samples* and $\boldsymbol{\xi}_i \in \mathbb{R}^m$ is a *sample point*. In case of the Monte Carlo quadrature, the sample points are chosen randomly. Therefore, we need a (pseudo-) random number generator which produces m -dimensional normally distributed random vectors. There are two main advantages of the Monte Carlo quadrature: the method does not suffer from the curse of dimensionality and requires very weak regularity assumptions on the integrand. The drawback of this method is that it produces only probabilistic error estimates, also known as *root mean square error*, cf. [9], and that it converges only with the rate $\mathcal{O}(N^{-1/2})$. More precisely, given a Hilbert space X , one has

$$(7.1) \quad (\mathbb{E}\|(\mathbf{I} - \mathbf{Q}_{MC})v\|_X^2)^{\frac{1}{2}} \lesssim N^{-\frac{1}{2}} \|v\|_{L^2_\rho(\mathbb{R}^m; X)}.$$

For the error estimation of the quasi-Monte Carlo method, it is required that the integrand has integrable, mixed first order derivatives. Then, the error of a quasi-Monte Carlo method over the unit cube $[0, 1]^m$ is bounded by means of the L^∞ -*star discrepancy*

$$\mathcal{D}_\infty^*(\Xi) := \sup_{\mathbf{t} \in [0, 1]^m} \left| \text{Vol}([\mathbf{0}, \mathbf{t})) - \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{[\mathbf{0}, \mathbf{t})}(\boldsymbol{\xi}_i) \right|$$

of the set of sample points $\Xi = \{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N\} \subset [0, 1]^m$, where $\text{Vol}([\mathbf{0}, \mathbf{t}))$ denotes the Lebesgue measure of the cuboid $[\mathbf{0}, \mathbf{t})$, cf. [33]. In case of certain point sequences, like the Halton sequence, see [21], this discrepancy is typically estimated to be of the order $\mathcal{O}(N^{-1}(\log N)^m)$.

To obtain a quasi-Monte Carlo method for the domain of integration \mathbb{R}^m , the sample points have to be mapped to \mathbb{R}^m by the inverse distribution function. Numerically, this can be done very efficiently by employing a rational interpolant of the inverse distribution function, cf. [32]. It has been shown in e.g. [28] that the error can again be bounded by $\mathcal{D}_\infty^*(\Xi)$ for a certain set of functions. To specify this, we define the Bochner space $W_{\text{mix}}^{1,1}(\mathbb{R}^m; X)$ which consists of all equivalence classes of strongly measurable functions $v : \mathbb{R}^m \rightarrow X$ with finite norm

$$(7.2) \quad \|v\|_{W_{\text{mix}}^{1,1}(\mathbb{R}^m; X)} := \sum_{\|\mathbf{q}\|_\infty \leq 1} \int_{\mathbb{R}^m} \|\partial_{\mathbf{y}}^{\mathbf{q}} v(\mathbf{y})\|_X \, d\mathbf{y} < \infty.$$

Then, the error of the quasi-Monte Carlo method is typically estimated by

$$\|(\mathbf{I} - \mathbf{Q}_{QMC})v\|_X \lesssim \mathcal{D}_\infty^*(\Xi) \|v\|_{W_{\text{mix}}^{1,1}(\mathbb{R}^m; X)},$$

cf. [28], which is an extension of the *Koksma-Hlawka* inequality, cf. [33], to unbounded domains. The condition that the norm $\|v\|_{W_{\text{mix}}^{1,1}(\mathbb{R}^m; X)}$ is bounded is in general very restrictive

and is not necessarily fulfilled in our application. Hence, we follow a suggestion of [28] and rewrite the integral $\mathbf{I}v(\mathbf{x})$ according to

$$(7.3) \quad \mathbf{I}v(\mathbf{x}) = \int_{\mathbb{R}^m} v(\mathbf{x}, \mathbf{y}) \rho(\mathbf{y}) \, d\mathbf{y} = \bar{\rho} \int_{\mathbb{R}^m} v(\mathbf{x}, \mathbf{y}) \sqrt{\rho(\mathbf{y})} \frac{\sqrt{\rho(\mathbf{y})}}{\bar{\rho}} \, d\mathbf{y},$$

with the scaling factor $\bar{\rho}$ being defined by $\bar{\rho} := \int_{\mathbb{R}^m} \sqrt{\rho(\mathbf{y})} \, d\mathbf{y}$. Now, we employ a quasi-Monte Carlo method with respect to the auxiliary density function $\sqrt{\rho(\mathbf{y})}/\bar{\rho}$ and obtain the error estimate

$$(7.4) \quad \|(\mathbf{I} - \mathbf{Q}_{QMC})v\|_X \lesssim \mathcal{D}_\infty^*(\Xi) \|v\sqrt{\rho}\|_{W_{\text{mix}}^{1,1}(\mathbb{R}^m; X)}.$$

Herein, the last norm is finite (with a constant which depends on m but not on N) in case of the moment computation as it is proven in the next theorem.

Theorem 7.1. *For the solution u to (2.6), the following bound is valid*

$$\|u\sqrt{\rho}\|_{W_{\text{mix}}^{1,1}(\mathbb{R}^m; H^1(D))} \lesssim \left(\sum_{\|\mathbf{q}\|_\infty \leq 1} \frac{1}{2^{|\mathbf{q}|}} \sum_{\boldsymbol{\alpha} \leq \mathbf{q}} \left(\frac{2\gamma}{\log 2} \right)^\alpha |\boldsymbol{\alpha}|! \right) \|f\|_{L^2(D)} < \infty, \quad p = 1.$$

Furthermore, if $f \in L^p(D)$, it holds for the p -th power u^p of u that

$$\|u^p\sqrt{\rho}\|_{W_{\text{mix}}^{1,1}(\mathbb{R}^m; W^{1,1}(D))} \lesssim \left(\sum_{\|\mathbf{q}\|_\infty \leq 1} \frac{1}{2^{|\mathbf{q}|}} \sum_{\boldsymbol{\alpha} \leq \mathbf{q}} \left(\frac{2p\gamma}{\log 2} \right)^\alpha |\boldsymbol{\alpha}|! \right) \|f\|_{L^p(D)}^p < \infty, \quad p \geq 2.$$

Proof. Each summand in the expression

$$\|v\sqrt{\rho}\|_{W_{\text{mix}}^{1,1}(\mathbb{R}^m; X)} = \sum_{\|\mathbf{q}\|_\infty \leq 1} \int_{\mathbb{R}^m} \|\partial_{\mathbf{y}}^{\mathbf{q}}(v(\mathbf{y})\sqrt{\rho(\mathbf{y})})\|_X \, d\mathbf{y}$$

can be estimated by

$$\int_{\mathbb{R}^m} \|\partial_{\mathbf{y}}^{\mathbf{q}}(v(\mathbf{y})\sqrt{\rho(\mathbf{y})})\|_X \, d\mathbf{y} = \sum_{\boldsymbol{\alpha} \leq \mathbf{q}} \frac{\mathbf{q}!}{\boldsymbol{\alpha}!(\mathbf{q}-\boldsymbol{\alpha})!} \int_{\mathbb{R}^m} \|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} v(\mathbf{y}) \partial_{\mathbf{y}}^{\mathbf{q}-\boldsymbol{\alpha}} \sqrt{\rho(\mathbf{y})}\|_X \, d\mathbf{y}.$$

Due to the product structure of the auxiliary density and since we consider only mixed first derivatives, we find $\partial_{\mathbf{y}}^{\mathbf{q}-\boldsymbol{\alpha}} \sqrt{\rho(\mathbf{y})} = \frac{(-1)^{|\mathbf{q}-\boldsymbol{\alpha}|} \mathbf{y}^{\mathbf{q}-\boldsymbol{\alpha}}}{2^{|\mathbf{q}-\boldsymbol{\alpha}|}} \sqrt{\rho(\mathbf{y})}$ and $\mathbf{q}! = 1$, $\boldsymbol{\alpha}! = 1$, $(\mathbf{q}-\boldsymbol{\alpha})! = 1$. Hence, we arrive at

$$\int_{\mathbb{R}^m} \|\partial_{\mathbf{y}}^{\mathbf{q}}(v(\mathbf{y})\sqrt{\rho(\mathbf{y})})\|_X \, d\mathbf{y} = \sum_{\boldsymbol{\alpha} \leq \mathbf{q}} \frac{1}{2^{|\mathbf{q}-\boldsymbol{\alpha}|}} \int_{\mathbb{R}^m} \|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} v(\mathbf{y})\|_X \mathbf{y}^{\mathbf{q}-\boldsymbol{\alpha}} \sqrt{\rho(\mathbf{y})} \, d\mathbf{y}.$$

For all functions v whose mixed first order derivatives grow at most exponentially in $\|\mathbf{y}\|$, the norm $\|v\sqrt{\rho}\|_{W_{\text{mix}}^{1,1}(\mathbb{R}^m; X)}$ is bounded since $\sqrt{\rho(\mathbf{y})} = \prod_{k=1}^m \exp(-y_k^2/4)$ decays double exponentially in $\|\mathbf{y}\|$. Thus, the integrals on the right hand side of this equation are all finite.

For the solution u , the multivariate version of (5.2) is

$$(7.5) \quad \|\partial_{\mathbf{y}}^{\alpha} u\|_{C_{\sigma}^0(\mathbb{R}^m; H^1(D))} \lesssim |\alpha|! \left(\frac{\gamma}{\log 2}\right)^{\alpha} \|f\|_{L^2(D)}.$$

Therefore, we obtain the first assertion

$$\begin{aligned} \int_{\mathbb{R}^m} \|\partial_{\mathbf{y}}^{\mathbf{q}}(u(\mathbf{y})\sqrt{\rho(\mathbf{y})})\|_{H^1(D)} \, d\mathbf{y} &\lesssim \sum_{\alpha \leq \mathbf{q}} \frac{|\alpha|!}{2^{|\mathbf{q}-\alpha|}} \left(\frac{\gamma}{\log 2}\right)^{\alpha} \|f\|_{L^2(D)} \int_{\mathbb{R}^m} \sigma^{-1}(\mathbf{y}) \mathbf{y}^{\mathbf{q}-\alpha} \sqrt{\rho(\mathbf{y})} \, d\mathbf{y} \\ &\lesssim \sum_{\alpha \leq \mathbf{q}} \frac{|\alpha|!}{2^{|\mathbf{q}-\alpha|}} \left(\frac{\gamma}{\log 2}\right)^{\alpha} \|f\|_{L^2(D)}. \end{aligned}$$

Note that the last step holds since $\int_{\mathbb{R}^m} \sigma^{-\ell}(\mathbf{y}) \mathbf{y}^{\mathbf{q}-\alpha} \sqrt{\rho(\mathbf{y})} \, d\mathbf{y} < \infty$ for all $\ell \in \mathbb{N}$.

For the p -th power u^p of the solution, the multivariate version of (5.5) reads

$$(7.6) \quad \|\partial_{\mathbf{y}}^{\alpha} u^p\|_{C_{\sigma^p}^0(\mathbb{R}^m; W^{1,1}(D))} \lesssim |\alpha|! \left(\frac{p\gamma}{\log 2}\right)^{\alpha} \|f\|_{L^p(D)}^p.$$

By setting $v = u^p$, we thus arrive at

$$\begin{aligned} \int_{\mathbb{R}^m} \|\partial_{\mathbf{y}}^{\mathbf{q}}(u^p(\mathbf{y})\sqrt{\rho(\mathbf{y})})\|_{W^{1,1}(D)} \, d\mathbf{y} &\lesssim \sum_{\alpha \leq \mathbf{q}} \frac{|\alpha|!}{2^{|\mathbf{q}-\alpha|}} \left(\frac{p\gamma}{\log 2}\right)^{\alpha} \|f\|_{L^p(D)}^p \int_{\mathbb{R}^m} \sigma^{-p}(\mathbf{y}) \mathbf{y}^{\mathbf{q}-\alpha} \sqrt{\rho(\mathbf{y})} \, d\mathbf{y} \\ &\lesssim \sum_{\alpha \leq \mathbf{q}} \frac{|\alpha|!}{2^{|\mathbf{q}-\alpha|}} \left(\frac{p\gamma}{\log 2}\right)^{\alpha} \|f\|_{L^p(D)}^p. \end{aligned}$$

This implies the second assertion. ■

Remark 7.2. *The estimation of the discrepancy of a set $\Xi \subset [0, 1]^m$, especially for high dimensions m , has been the topic of many publications in the past fifteen years. The aim is to avoid the factor $(\log N)^{m-1}$ in the estimation of the discrepancy which grows exponentially in the dimension m . This exponential dependence is called intractability in the literature, cf. [35, 39]. To avoid intractability, further regularity assumptions on the integrand are necessary. Such assumptions are not imposed in the analysis presented above. In particular, the constant occurring in the previous theorem can only be bounded by*

$$\sum_{\|\mathbf{q}\|_{\infty} \leq 1} \frac{1}{2^{|\mathbf{q}|}} \sum_{\alpha \leq \mathbf{q}} \left(\frac{2p\gamma}{\log 2}\right)^{\alpha} |\alpha|! \lesssim \exp(cm \log m)$$

with some constant $c > 0$ if we do not take into account the decay of the sequence $\{\gamma_k\}_k$. As a consequence, for large values of m , one has to modify the analysis of the quasi-Monte Carlo quadrature. This can be done by extending the ideas in [36] and additionally exploiting the decay of the sequence $\{\gamma_k\}_k$, see [24] for the case of the Halton-sequence.

8. Analysis of the multilevel quadrature. We now want to use the results from Sections 6 and 7 to analyze the error of the multilevel quadrature approximation to the mean and the moments of the solution to (2.6). As we have seen, the error analysis in case of the Gaussian quadrature or the quasi-Monte Carlo quadrature is based on the derivatives of the integrand. Since the integrands in the multilevel quadrature are given by terms of the form $(\mathbf{u}_\ell^p - \mathbf{u}_{\ell-1}^p)(\mathbf{y})$, cf. (4.4), we shall show that their derivatives exhibit a behaviour similar to the derivatives of $u^p(\mathbf{y})$ but provide an additional factor $2^{-\ell}$.

Lemma 8.1. *For the error $\delta_\ell(\mathbf{y}) := (\mathbf{u}_\ell - u)(\mathbf{y})$ of the Galerkin projection, there holds the estimate*

$$(8.1) \quad \|\partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y})\|_{H^1(D)} \lesssim 2^{-\ell} |\alpha|! \tilde{\kappa}(\mathbf{y})^2 \left(\frac{3\tilde{\gamma}}{\log 2} \right)^\alpha \|f\|_{L^2(D)} \quad \text{for all } |\alpha| \geq 0.$$

Therefore, we have for the detail projections $\theta_\ell(\mathbf{y}) := (\mathbf{u}_\ell - \mathbf{u}_{\ell-1})(\mathbf{y})$ the estimate

$$(8.2) \quad \|\partial_{\mathbf{y}}^\alpha \theta_\ell(\mathbf{y})\|_{H^1(D)} \lesssim 3 \cdot 2^{-\ell} |\alpha|! \tilde{\kappa}(\mathbf{y})^2 \left(\frac{3\tilde{\gamma}}{\log 2} \right)^\alpha \|f\|_{L^2(D)} \quad \text{for all } |\alpha| \geq 0.$$

Proof. Since the Galerkin projection satisfies $(a(\mathbf{y}) \nabla_{\mathbf{x}} \delta_\ell(\mathbf{y}), \nabla_{\mathbf{x}} v)_{L^2(D)} = 0$ for all $v \in \mathcal{S}_\ell^1(D)$, it follows by differentiation that

$$-\int_D a(\mathbf{y}) \nabla_{\mathbf{x}} \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y}) \nabla_{\mathbf{x}} v \, d\mathbf{x} = \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \int_D \partial_{\mathbf{y}}^\beta a(\mathbf{y}) \nabla_{\mathbf{x}} \partial_{\mathbf{y}}^{\alpha-\beta} \delta_\ell(\mathbf{y}) \nabla_{\mathbf{x}} v \, d\mathbf{x}$$

for all $v \in \mathcal{S}_\ell^1(D)$. For an arbitrary function $v \in \mathcal{S}_\ell^1(D)$, we therefore obtain:

$$\begin{aligned} \|\sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y})\|_{L^2(D)}^2 &= \int_D a(\mathbf{y}) |\nabla_{\mathbf{x}} \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y})|^2 \, d\mathbf{x} \\ &= \int_D a(\mathbf{y}) \nabla_{\mathbf{x}} \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y}) [\nabla_{\mathbf{x}} \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y}) - \nabla_{\mathbf{x}} v] \, d\mathbf{x} \\ &\quad + \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \int_D \partial_{\mathbf{y}}^\beta a(\mathbf{y}) \nabla_{\mathbf{x}} \partial_{\mathbf{y}}^{\alpha-\beta} \delta_\ell(\mathbf{y}) [\nabla_{\mathbf{x}} \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y}) - \nabla_{\mathbf{x}} v] \, d\mathbf{x} \\ &\quad - \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \int_D \partial_{\mathbf{y}}^\beta a(\mathbf{y}) \nabla_{\mathbf{x}} \partial_{\mathbf{y}}^{\alpha-\beta} \delta_\ell(\mathbf{y}) \nabla_{\mathbf{x}} \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y}) \, d\mathbf{x}. \end{aligned}$$

From (2.5), we derive $\|\partial_{\mathbf{y}}^\beta a(\mathbf{y})/a(\mathbf{y})\|_{L^\infty(D)} \leq \gamma^\beta$. Hence, we can further estimate

$$\begin{aligned} \|\sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y})\|_{L^2(D)}^2 &\leq \int_D a(\mathbf{y}) \nabla_{\mathbf{x}} \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y}) [\nabla_{\mathbf{x}} \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y}) - \nabla_{\mathbf{x}} v] \, d\mathbf{x} \\ &\quad + \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \gamma^\beta \int_D a(\mathbf{y}) |\nabla_{\mathbf{x}} \partial_{\mathbf{y}}^{\alpha-\beta} \delta_\ell(\mathbf{y}) [\nabla_{\mathbf{x}} \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y}) - \nabla_{\mathbf{x}} v]| \, d\mathbf{x} \\ &\quad + \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \gamma^\beta \int_D a(\mathbf{y}) |\nabla_{\mathbf{x}} \partial_{\mathbf{y}}^{\alpha-\beta} \delta_\ell(\mathbf{y}) \nabla_{\mathbf{x}} \partial_{\mathbf{y}}^\alpha \delta_\ell(\mathbf{y})| \, d\mathbf{x}. \end{aligned}$$

The Cauchy-Schwarz inequality yields

$$\begin{aligned}
(8.3) \quad & \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)}^2 \leq \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \left\| \sqrt{a(\mathbf{y})} \nabla (\partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) - v) \right\|_{L^2(D)} \\
& + \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \gamma^{\beta} \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha-\beta} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \left\| \sqrt{a(\mathbf{y})} \nabla (\partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) - v) \right\|_{L^2(D)} \\
& + \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \gamma^{\beta} \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha-\beta} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)}.
\end{aligned}$$

Since $v \in \mathcal{S}_{\ell}^1(D)$ can be chosen arbitrarily, the bound holds also for the infimum

$$\begin{aligned}
& \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)}^2 \leq \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \inf_{v \in \mathcal{S}_{\ell}^1(D)} \left\| \sqrt{a(\mathbf{y})} \nabla (\partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) - v) \right\|_{L^2(D)} \\
& + \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \gamma^{\beta} \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha-\beta} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \inf_{v \in \mathcal{S}_{\ell}^1(D)} \left\| \sqrt{a(\mathbf{y})} \nabla (\partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) - v) \right\|_{L^2(D)} \\
& + \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \gamma^{\beta} \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha-\beta} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)}.
\end{aligned}$$

The expression containing the infimum is now estimated in two different ways. On the one hand, the approximation property of the finite element space $\mathcal{S}_{\ell}^1(D)$ implies

$$\inf_{v \in \mathcal{S}_{\ell}^1(D)} \left\| \sqrt{a(\mathbf{y})} \nabla (\partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) - v) \right\|_{L^2(D)} \lesssim 2^{-\ell} \sqrt{a_{\max}(\mathbf{y})} \left\| \partial_{\mathbf{y}}^{\alpha} u(\mathbf{y}) \right\|_{H^2(D)}.$$

On the other hand, due to $0 \in \mathcal{S}_{\ell}^1(D)$, we find

$$\inf_{v \in \mathcal{S}_{\ell}^1(D)} \left\| \sqrt{a(\mathbf{y})} \nabla (\partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) - v) \right\|_{L^2(D)} \leq \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)}.$$

Dividing (8.3) by $\left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)}$ leads in combination with both estimates to

$$\begin{aligned}
& \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \lesssim 2^{-\ell} \sqrt{a_{\max}(\mathbf{y})} \left\| \partial_{\mathbf{y}}^{\alpha} u(\mathbf{y}) \right\|_{H^2(D)} \\
& + 2 \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \gamma^{\beta} \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha-\beta} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)}.
\end{aligned}$$

In view of (5.8), we obtain

$$\begin{aligned}
(8.4) \quad & \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \lesssim 2^{-\ell} \sqrt{a_{\max}(\mathbf{y})} \tilde{\kappa}(\mathbf{y}) |\alpha|! \left(\frac{2\tilde{\gamma}}{\log 2} \right)^{\alpha} \|f\|_{L^2(D)} \\
& + 2 \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} \gamma^{\beta} \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha-\beta} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)}.
\end{aligned}$$

We conclude now by induction that

$$\left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \lesssim 2^{-\ell} \frac{a_{\max}(\mathbf{y})}{\sqrt{a_{\min}(\mathbf{y})}} \tilde{\kappa}(\mathbf{y}) (3\tilde{\gamma})^{\alpha} B_{|\alpha|} \|f\|_{L^2(D)}.$$

Herein, B_k denotes the k -th *ordered Bell-number* which satisfies the recurrence relation $B_k = \sum_{i=0}^{k-1} \binom{k}{i} B_i$ with $B_0 = 1$. For $|\alpha| = 0$, this is simply obtained by estimate (5.8) and (3.1). Let now the induction hypothesis hold for all $\beta < \alpha$. Inserting this hypothesis into (8.4) yields

$$\begin{aligned} & \left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \\ & \lesssim 2^{-\ell} \frac{a_{\max}(\mathbf{y})}{\sqrt{a_{\min}(\mathbf{y})}} \tilde{\kappa}(\mathbf{y}) \|f\|_{L^2(D)} \left(|\alpha|! \left(\frac{2\tilde{\gamma}}{\log 2} \right)^{\alpha} + 2\tilde{\gamma}^{\alpha} \sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} 3^{|\alpha-\beta|} B_{|\alpha-\beta|} \right). \end{aligned}$$

With

$$\sum_{\mathbf{0} \neq \beta \leq \alpha} \binom{\alpha}{\beta} 3^{|\alpha-\beta|} B_{|\alpha-\beta|} \leq 3^{|\alpha|-1} \sum_{k=0}^{|\alpha|-1} B_k \sum_{\substack{\beta < \alpha \\ |\beta|=k}} \binom{\alpha}{\beta} = 3^{|\alpha|-1} \sum_{k=0}^{|\alpha|-1} B_k \binom{|\alpha|}{k} = 3^{|\alpha|-1} B_{|\alpha|}$$

and the estimate $B_k \leq k!(\log 2)^{-k}$, cf. [4], we get

$$\left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \lesssim 2^{-\ell} \frac{a_{\max}(\mathbf{y})}{\sqrt{a_{\min}(\mathbf{y})}} \tilde{\kappa}(\mathbf{y}) \|f\|_{L^2(D)} |\alpha|! \left(\frac{\tilde{\gamma}}{\log 2} \right)^{\alpha} (2^{|\alpha|} + 2 \cdot 3^{|\alpha|-1}).$$

Moreover, since $2^{|\alpha|} + 2 \cdot 3^{|\alpha|-1} \leq 3^{|\alpha|}$ for $|\alpha| > 2$, $a_{\max}(\mathbf{y}) \leq \tilde{a}_{\max}(\mathbf{y})$ and $a_{\min}(\mathbf{y}) \geq \tilde{a}_{\min}(\mathbf{y})$, we conclude

$$\left\| \sqrt{a(\mathbf{y})} \nabla \partial_{\mathbf{y}}^{\alpha} \delta_{\ell}(\mathbf{y}) \right\|_{L^2(D)} \lesssim 2^{-\ell} \frac{(\tilde{a}_{\max}(\mathbf{y}))^2}{(\tilde{a}_{\min}(\mathbf{y}))^{3/2}} |\alpha|! \left(\frac{3\tilde{\gamma}}{\log 2} \right)^{\alpha} \|f\|_{L^2(D)},$$

which implies finally the assertion. ■

For $p > 2$, there holds the following L^p -extension of estimate (8.2)

$$(8.5) \quad \left\| \partial_{\mathbf{y}}^{\alpha} \theta_{\ell}(\mathbf{y}) \right\|_{W^{1,p}} \lesssim \tilde{\kappa}(\mathbf{y})^2 |\alpha|! \left(\frac{3C(p, D)\tilde{\gamma}}{\log 2} \right)^{\alpha} \|f\|_{L^p(D)},$$

which can be proven if the load f belongs to $L^p(D)$. It is applied to get the following result also for moments of order $p > 2$.

Lemma 8.2. *The derivatives of the difference $\mathbf{u}_{\ell}^2 - \mathbf{u}_{\ell-1}^2$ satisfy the estimate*

$$(8.6) \quad \left\| \partial_{\mathbf{y}}^{\alpha} (\mathbf{u}_{\ell}^2 - \mathbf{u}_{\ell-1}^2)(\mathbf{y}) \right\|_{W^{1,1}(D)} \lesssim 2^{-\ell} |\alpha|! \left(\frac{6\tilde{\gamma}}{\log 2} \right)^{\alpha} \tilde{\kappa}(\mathbf{y})^3 \|f\|_{L^2(D)}^2.$$

Moreover, for the p -th powers of two successive Galerkin projections \mathbf{u}_{ℓ}^p and $\mathbf{u}_{\ell-1}^p$, there holds the estimate

$$(8.7) \quad \left\| \partial_{\mathbf{y}}^{\alpha} (\mathbf{u}_{\ell}^p - \mathbf{u}_{\ell-1}^p)(\mathbf{y}) \right\|_{W^{1,1}(D)} \lesssim 2^{-\ell} |\alpha|! \left(\frac{3pC(p, D)\tilde{\gamma}}{\log 2} \right)^{\alpha} \tilde{\kappa}(\mathbf{y})^{3/2p+1/2} \|f\|_{L^p(D)}^p$$

provided that $f \in L^p(D)$ for $p > 2$.

Proof. For the sake of convenience, we demonstrate the proof only for the case of $p = 2$. The case of a general p can be treated in a similar way with the help of Faà di Bruno's formula, the $L^p(D)$ -extension of Proposition 5.8 and the m -dimensional Hölder inequality. It holds

$$\begin{aligned} \|\partial_{\mathbf{y}}^{\alpha}(\mathbf{u}_{\ell}^2 - \mathbf{u}_{\ell-1}^2)(\mathbf{y})\|_{W^{1,1}(D)} &\leq \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \|\partial_{\mathbf{y}}^{\beta} \theta_{\ell}(\mathbf{y}) \partial_{\mathbf{y}}^{\alpha-\beta}(\mathbf{u}_{\ell} + \mathbf{u}_{\ell-1})(\mathbf{y})\|_{W^{1,1}(D)} \\ &\lesssim \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \|\partial_{\mathbf{y}}^{\alpha} \theta_{\ell}(\mathbf{y})\|_{H^1(D)} \|\partial_{\mathbf{y}}^{\alpha-\beta}(\mathbf{u}_{\ell} + \mathbf{u}_{\ell-1})(\mathbf{y})\|_{H^1(D)}. \end{aligned}$$

Using the estimate (8.2), the fact that the Galerkin projection $\mathbf{u}_{\ell}(\mathbf{y})$ has the same regularity with respect to the parametric variable as the solution itself, and Lemma 5.4, we obtain

$$\begin{aligned} \|\partial_{\mathbf{y}}^{\alpha}(\mathbf{u}_{\ell}^2 - \mathbf{u}_{\ell-1}^2)(\mathbf{y})\|_{W^{1,1}(D)} &\lesssim \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} 2^{-\ell} |\beta|! \tilde{\kappa}(\mathbf{y})^2 \left(\frac{3\tilde{\gamma}}{\log 2}\right)^{\beta} (\alpha - \beta)! \left(\frac{2\gamma}{\log 2}\right)^{\alpha-\beta} \kappa(\mathbf{y}) \|f\|_{L^2(D)}^2 \\ &\lesssim 2^{-\ell} |\alpha|! \left(\frac{3\tilde{\gamma}}{\log 2}\right)^{\alpha} \tilde{\kappa}(\mathbf{y})^3 \|f\|_{L^2(D)}^2 \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \\ &= 2^{-\ell} |\alpha|! \left(\frac{3\tilde{\gamma}}{\log 2}\right)^{\alpha} \tilde{\kappa}(\mathbf{y})^3 \|f\|_{L^2(D)}^2 2^{|\alpha|}, \end{aligned}$$

which yields the assertion for $p = 2$. ■

By employing Lemma 8.1 and Lemma 8.2, we derive now the crucial estimate which will be used in the forecoming Theorem 8.4.

Lemma 8.3. *In case of MLQMC and MLGQ, the estimate*

$$(8.8) \quad \|(\mathbf{I} - \mathbf{Q}_{j-\ell})(\mathbf{u}_{\ell}^p - \mathbf{u}_{\ell-1}^p)(\mathbf{y})\|_X \lesssim \varepsilon_{j-\ell} 2^{-\ell} \|f\|_{L^{p+e}(D)}^p$$

holds for all $p \geq 1$, where $X = H^1(D)$, $e = 1$ if $p = 1$ and $X = W^{1,1}(D)$, $e = 0$ if $p \geq 2$, respectively.

Proof. The results of Lemma 8.1 and Lemma 8.2 imply

$$(8.9) \quad \|\partial_{\mathbf{y}}^{\alpha} \theta_{\ell}\|_{C_{\sigma}^0(\mathbb{R}^m; H^1(D))} \lesssim 2^{-\ell} |\alpha|! \left(\frac{3\tilde{\gamma}}{\log 2}\right)^{\alpha} \|f\|_{L^2(D)}$$

for all $\alpha \geq \mathbf{0}$, provided that the weight σ from Definition 5.2 satisfies $\beta \geq 4\tilde{\gamma}$, and

$$(8.10) \quad \|\partial_{\mathbf{y}}^{\alpha}(\mathbf{u}_{\ell}^2 - \mathbf{u}_{\ell-1}^2)(\mathbf{y})\|_{C_{\sigma}^0(\mathbb{R}^m; W^{1,1}(D))} \lesssim 2^{-\ell} |\alpha|! \left(\frac{6\tilde{\gamma}}{\log 2}\right)^{\alpha} \|f\|_{L^2(D)}^2$$

for all $\alpha \geq \mathbf{0}$, provided that the weight σ satisfies $\beta \geq 6\tilde{\gamma}$. Moreover, if $f \in L^p(D)$, we derive from (8.7) that

$$(8.11) \quad \|\partial_{\mathbf{y}}^{\alpha}(\mathbf{u}_{\ell}^p - \mathbf{u}_{\ell-1}^p)(\mathbf{y})\|_{C_{\sigma}^0(\mathbb{R}^m; W^{1,1}(D))} \lesssim 2^{-\ell} |\alpha|! \left(\frac{3pC(p, D)\tilde{\gamma}}{\log 2}\right)^{\alpha} \|f\|_{L^p(D)}^p$$

for all $\boldsymbol{\alpha} \geq \mathbf{0}$, provided that the weight $\boldsymbol{\sigma}$ satisfies $\boldsymbol{\beta} \geq (3p+1)\tilde{\gamma}$. In the analysis of Section 6, the dependency of the error estimate on the behaviour of the derivatives arises in Lemma 6.3. In view of the proof of this lemma, it holds for $\tau_k < \frac{\log 2}{3\tilde{\gamma}_k}$ that

$$\sup_{z \in \Sigma(\tau_k)} \boldsymbol{\sigma}(\operatorname{Re}(z)) \|\theta_\ell(z)\|_{C_{\boldsymbol{\sigma}_k^*}^0(\mathbb{R}^{m-1}; X)} \lesssim 2^{-\ell} \|f\|_{L^2(D)}.$$

Analogously, we get for the p -th moments with $\tau_k < \frac{\log 2}{p3\tilde{\gamma}_k}$ that

$$\sup_{z \in \Sigma(\tau_k)} \boldsymbol{\sigma}(\operatorname{Re}(z)) \|(\mathbf{u}_\ell^p - \mathbf{u}_{\ell-1}^p)(z)\|_{C_{\boldsymbol{\sigma}_k^*}^0(\mathbb{R}^{m-1}; X)} \lesssim 2^{-\ell} \|f\|_{L^p(D)}^p.$$

We thus obtain the estimate (8.8) in case of the Gaussian quadrature when the number of quadrature points of $\mathbf{Q}_{j-\ell}$ in (6.5) is chosen according to $\varepsilon_{j-\ell}$. In the analysis of the quasi-Monte Carlo quadrature, we have to replace the estimate (7.5) by (8.9) for the mean and the estimate (7.6) by (8.10) and (8.11) to obtain the additional factor $2^{-\ell}$ in Theorem 7.1. Choosing the number of quadrature points of $\mathbf{Q}_{j-\ell}$ according to the accuracy $\varepsilon_{j-\ell}$, yields then (8.8). ■

With the previous lemmata at hand, we are finally able to bound the error of the multilevel quadrature methods under consideration. Again, to show this result also for moments of order $p > 2$, we need the L^p -extension of estimate (5.8) for $j = 0$:

$$(8.12) \quad \|u(\mathbf{y})\|_{W^{2,p}(D)} \lesssim \tilde{\kappa}(\mathbf{y}) \|f\|_{L^p(D)}.$$

This estimate holds if the load f belongs to $L^p(D)$.

Theorem 8.4. *Let $\{\mathbf{Q}_\ell\}$ be a sequence of quadrature rules which satisfy (4.2) and let $u \in C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m, H_0^1(D))$ be the solution to (2.6) which satisfies (3.1) and (3.2). Then, in case of MLQMC and MLGQ, the errors of the multilevel estimators (4.3) and (4.4) are bounded by*

$$(8.13) \quad \left\| \mathcal{M}_u^p - \sum_{\ell=0}^j \mathbf{Q}_{j-\ell}(\mathbf{u}_\ell^p - \mathbf{u}_{\ell-1}^p) \right\|_X \lesssim 2^{-j} j \|f\|_{L^{p+e}(D)}^p$$

where $e = 1$, $X = H^1(D)$ if $p = 1$ and $e = 0$, $X = W^{1,1}(D)$ if $p \geq 2$.

Proof. We shall apply the following multilevel splitting of the error

$$(8.14) \quad \left\| \mathcal{M}_u^p - \sum_{\ell=0}^j \mathbf{Q}_{j-\ell}(\mathbf{u}_\ell^p - \mathbf{u}_{\ell-1}^p) \right\|_X \leq \|\mathcal{M}_u^p - \mathcal{M}_{u_j}^p\|_X + \sum_{\ell=0}^j \|(\mathbf{I} - \mathbf{Q}_{j-\ell})(\mathbf{u}_\ell^p - \mathbf{u}_{\ell-1}^p)\|_X.$$

Employing (4.2) in (8.8), the sum on the right hand side is estimated by

$$\sum_{\ell=0}^j \|(\mathbf{I} - \mathbf{Q}_{j-\ell})(\mathbf{u}_\ell^p - \mathbf{u}_{\ell-1}^p)\|_X \lesssim \sum_{\ell=0}^j 2^{-(j-\ell)} 2^{-\ell} \|f\|_{L^{p+e}(D)}^p \lesssim j 2^{-j} \|f\|_{L^{p+e}(D)}^p.$$

For $p = 1$, due to (3.1) and the continuity of \mathbf{I} on $C_{\boldsymbol{\sigma}}^0(\mathbb{R}^m; H^2(D))$ for $\boldsymbol{\sigma}$ with $\boldsymbol{\beta} \geq 3\tilde{\gamma}$, the first term on the right hand side of (8.14) satisfies

$$\|\mathbb{E}_u - \mathbb{E}_{u_j}\|_{H^1(D)} \lesssim 2^{-j} \sup_{\mathbf{y} \in \mathbb{R}^m} \boldsymbol{\sigma}(\mathbf{y}) \sqrt{\kappa(\mathbf{y})} \|u(\mathbf{y})\|_{H^2(D)} \lesssim 2^{-j} \|f\|_{L^2(D)}.$$

For $p \geq 2$, we use (3.1), the continuity of \mathbf{I} on $C_{\sigma}^0(\mathbb{R}^m; H^2(D))$ for σ with $\beta \geq 4p\tilde{\gamma}$ and the assumption (8.12) to obtain

$$(8.15) \quad \|\mathcal{M}_u^p - \mathcal{M}_{u_j}^p\|_{W^{1,1}(D)} \lesssim 2^{-j} \sup_{\mathbf{y} \in \mathbb{R}^m} \sigma(\mathbf{y}) \kappa(\mathbf{y})^p \|u(\mathbf{y})\|_{W^{2,p}(D)}^p \lesssim 2^{-j} \|f\|_{L^p(D)}^p.$$

Inserting (8.15) into (8.14) yields

$$\left\| \mathcal{M}_u^p - \sum_{\ell=0}^j \mathbf{Q}_{j-\ell}(\mathbf{u}_{\ell}^p - \mathbf{u}_{\ell-1}^p) \right\|_X \lesssim 2^{-j} \|f\|_{L^{p+e}(D)}^p + j2^{-j} \|f\|_{L^{p+e}}^p \lesssim j2^{-j} \|f\|_{L^{p+e}(D)}^p.$$

This completes the proof. \blacksquare

Finally, when applying Monte Carlo quadrature, we have to modify the above arguments since the error needs now to be bounded in the mean square sense.

Theorem 8.5. *Under the assumption of Theorem 8.4, the errors of the multilevel estimators (4.3) and (4.4) are in case of MLMC bounded by*

$$(8.16) \quad \left(\mathbb{E} \left\| \mathcal{M}_u^p - \sum_{\ell=0}^j \mathbf{Q}_{j-\ell}(\mathbf{u}_{\ell}^p - \mathbf{u}_{\ell-1}^p) \right\|_X^2 \right)^{\frac{1}{2}} \lesssim 2^{-j} j \|f\|_{L^{2p}(D)}^p.$$

Proof. It holds

$$\left(\mathbb{E} \left\| \mathcal{M}_u^p - \sum_{\ell=0}^j \mathbf{Q}_{j-\ell}(\mathbf{u}_{\ell}^p - \mathbf{u}_{\ell-1}^p) \right\|_X^2 \right)^{\frac{1}{2}} \leq \left(\mathbb{E} \|\mathcal{M}_u^p - \mathcal{M}_{u_j}^p\|_X^2 \right)^{\frac{1}{2}} + \sum_{\ell=0}^j \left(\mathbb{E} \|(\mathbf{I} - \mathbf{Q}_{j-\ell})(\mathbf{u}_{\ell}^p - \mathbf{u}_{\ell-1}^p)\|_X^2 \right)^{\frac{1}{2}}$$

by the triangle inequality. The first term on the right hand side is just $\|\mathcal{M}_u^p - \mathcal{M}_{u_j}^p\|_X$ which is already estimated in the proof of Theorem 8.4. The second term on the right hand side is estimated in accordance with

$$\left(\mathbb{E} \|(\mathbf{I} - \mathbf{Q}_{j-\ell})(\mathbf{u}_{\ell}^p - \mathbf{u}_{\ell-1}^p)(\mathbf{y})\|_X^2 \right)^{\frac{1}{2}} \lesssim \varepsilon_{j-\ell} \|(\mathbf{u}_{\ell}^p - \mathbf{u}_{\ell-1}^p)(\mathbf{y})\|_{L_{\rho}^2(\mathbb{R}^m; H_0^1(D))} \lesssim \varepsilon_{j-\ell} 2^{-\ell} \|f\|_{L^{2p}(D)}^p.$$

Here, the last inequality is obtained by using estimate (3.1) if $p = 1$. If $p \geq 2$, then the estimate is immediately derived by the generalized Hölder inequality. \blacksquare

9. Complexity. In this section, we give estimates for the complexity of the different multilevel quadrature methods. We assume for each quadrature point in $\mathbf{Q}_{j-\ell}$ that the computational cost for solving the resulting elliptic partial differential equation on level ℓ is $\mathcal{O}(2^{n\ell})$.

The complexity of the multilevel Monte-Carlo quadrature has been considered before in e.g. [3, 11, 12], where the number of quadrature points on each level is chosen in order to get an overall root mean square error of ε . Our analysis differs slightly since we equilibrate the error contribution on each level.

The Monte Carlo quadrature with N_{ℓ} quadrature points has a precision of $\mathcal{O}(N_{\ell}^{-1/2})$. Thus, to reach an accuracy of $2^{-\ell}$, we need $N_{\ell} \approx 2^{2\ell}$ quadrature points. Therefore, the overall cost of MLMC quadrature to achieve the accuracy 2^{-j} are bounded by

$$C_{\text{MLMC}} = \sum_{\ell=0}^j 2^{2(j-\ell)} 2^{n\ell} = 2^{2j} \sum_{\ell=0}^j 2^{(n-2)\ell} = \begin{cases} \mathcal{O}(2^{2j}), & \text{if } n = 1, \\ \mathcal{O}(j2^{2j}), & \text{if } n = 2, \\ \mathcal{O}(2^{nj}), & \text{if } n > 2. \end{cases}$$

For the quasi-Monte Carlo quadrature, we need $N_\ell \approx 2^\ell \ell^m$ quadrature points to ensure the accuracy $2^{-\ell}$. Indeed, there holds

$$N_\ell^{-1} (\log N_\ell)^m = 2^{-\ell} \ell^{-m} \log \left(2^\ell \ell^m \right)^m \leq 2^{-\ell} \ell^{-m} (\ell + m \log \ell)^m \leq C_1(m) 2^{-\ell}.$$

This leads to the overall cost of MLQMC method

$$C_{\text{MLQMC}} = \sum_{\ell=0}^j 2^{j-\ell} (\log(2^{j-\ell}))^m 2^{n\ell} \leq \sum_{\ell=0}^j 2^{j-\ell} (j-\ell)^m 2^{n\ell} = \begin{cases} \mathcal{O}(2^j j^{m+1}), & \text{if } n = 1, \\ \mathcal{O}(2^{nj}), & \text{if } n \geq 2. \end{cases}$$

In case of the Gaussian quadrature, we have that $N_\ell \approx \left[\frac{1}{2} ((\ell \log 2)^2 + 1) \right]^m \leq C_2(m) \ell^{2m}$ according to (6.5). This yields

$$C_{\text{MLGQ}} \leq C_2(m) \sum_{\ell=0}^j (j-\ell)^{2m} 2^{n\ell} \leq C_3(m) 2^{nj}.$$

quadrature points in total for MLGQ.

The cost complexities of the multilevel quadrature methods under consideration are summarized in Table 9.1.

| spatial dimension | $n = 1$ | $n = 2$ | $n = 3$ |
|-------------------|---------------|------------|----------|
| MLMC | 2^{2j} | $2^{2j} j$ | 2^{3j} |
| MLQMC | $2^j j^{m+1}$ | 2^{2j} | 2^{3j} |
| MLGQ | 2^j | 2^{2j} | 2^{3j} |

Table 9.1

Cost complexities of the different multilevel quadrature methods to get the accuracy $j2^{-j}$.

Remark 9.1. *The constants hidden in the computational cost estimates depend in case of MLQMC and MLGQ exponentially on the dimension m and thus explode for large m . For high dimensions m , the anisotropy induced by the weight $\tilde{\gamma}$ has to be taken into account to avoid this exponential dependence.*

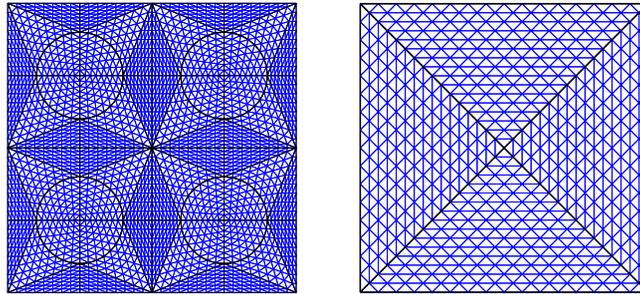


Figure 10.1. *Computational domains with inscribed coarse grids.*

10. Numerical results. In this section, we present numerical examples to validate and quantify the presented methods by studying the convergence for the mean and for the second moment of the solution to (2.1). To that end, we consider two different settings. In the first case, the diffusion coefficient is represented exactly by a Karhunen-Loève expansion of finite rank. In the second case, the diffusion coefficient is described by a Gaussian correlation function. Hence, the Karhunen-Loève expansion has to be appropriately truncated, where the truncation rank m tends to ∞ as the overall accuracy increases. The domain of the spatial variable is always the unit square. For the approximation of the Karhunen-Loève expansion, we employ the pivoted Cholesky decomposition as described in [22, 25] together with a piecewise constant finite element discretization of the two-point correlation.

Since we use linear finite elements, the representation on different grids for the solution's second moment is performed via a quadratic prolongation. The reference solution is not analytically known. Thus, we compute it numerically by a quasi-Monte Carlo method on a finer spatial grid with 1 048 576 finite elements and 10^6 samples based on the Halton sequence, cf. [21], as described in Section 7. All computations have been carried out on a computing server consisting of four nodes* with up to 64 processes.

10.1. An example with a covariance of finite rank. In our first numerical example, we focus on the covariance function $\text{Cov}_b(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^4 \mathbb{1}_{B_i}(\mathbf{x}) \mathbb{1}_{B_i}(\mathbf{y})$, where the sets B_1, \dots, B_4 are discs of diameter 0.3 equispaced in $D = (0, 1)^2$. A visualization of the associated triangulation can be seen in the left picture of Figure 10.1. We consider $f \equiv 1$ as load vector. Figure 10.2 shows the solution's mean (on the left) and the solution's variance (on the right).

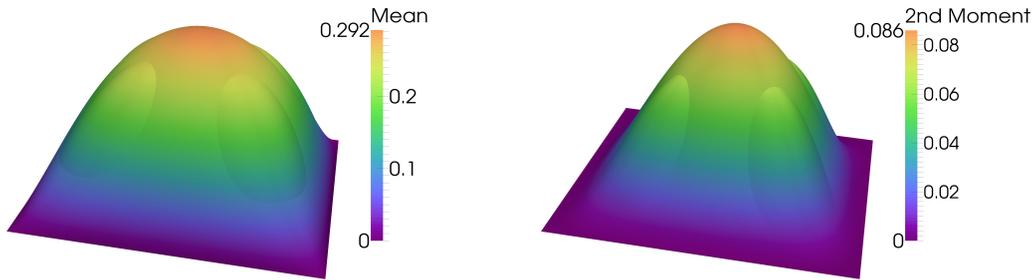


Figure 10.2. Solution's mean (left) and uncentered second moment (right) in case of the first example.

For the multilevel methods, we choose the respective number of samples as follows. We set $N_\ell = 10 \cdot 4^{j-\ell}$ for MLMC and $N_\ell = 10 \cdot 2^{(j-\ell)/(1-\delta)}$ for MLQMC where $\delta = 0, 0.25, 0.5$. As quadrature method in the quasi-Monte Carlo case, we employ the quasi Monte-Carlo method based on the Halton sequence with respect to the auxiliary density function $\sqrt{\rho(y)}/\bar{\rho}$ as presented in Section 7. The number of samples for MLGQ is controlled by the quadrature orders determined from (6.5). In order to guarantee the overall precision of ε , we set the accuracy of the Gaussian quadrature in each stochastic dimension to $\varepsilon/4$. On the right hand side of Figure 10.3, we visualize the chosen quadrature orders for MLGQ on each level and in each stochastic dimension for the computations on level 6. A comparison of the number of samples

*Each node consists of two quad-core Intel(R) Xeon(R) X5550 CPUs with a clock rate of 2.67GHz (hyper-threading enabled) and 48GB of main memory.

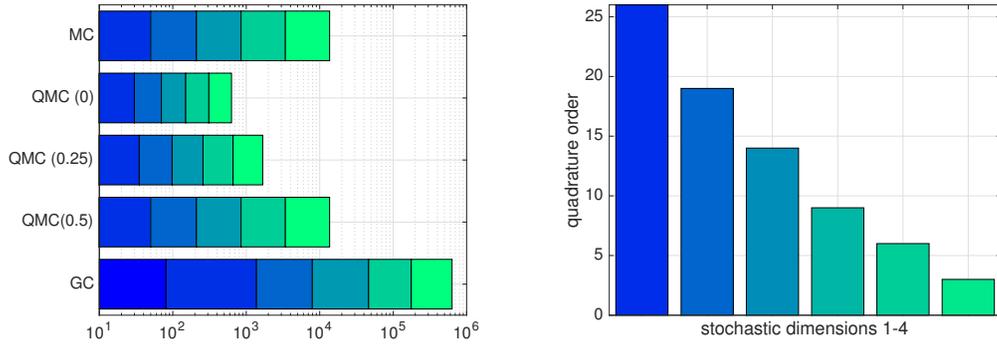


Figure 10.3. Number of samples for the different quadrature methods for the first example (left) and related orders of the Gaussian quadrature (right).

for each of the methods is depicted on the left hand side of Figure 10.3. Here, the different colors refer to the samples on each particular level sorted in decreasing order from left to right, i.e. $\ell = 6, 5, \dots, 1$. Note that, for MLMC, we averaged five realizations of the multilevel estimator in order to approximate the root mean square error. This fact is not taken into account in the visualization in Figure 10.3, i.e. the number of samples refer to one realization of the multilevel estimator. The error plots found in Figure 10.4 indicate that the three meth-

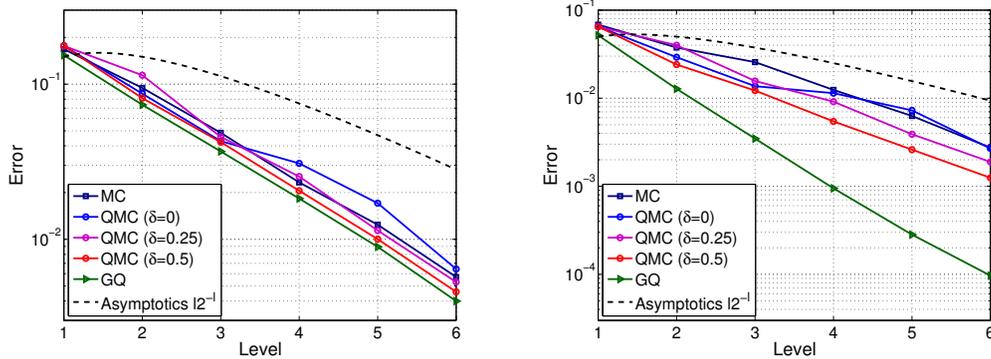


Figure 10.4. Error of the mean (left) and of the second moment (right) in case of the first example.

ods, i.e. MLGQ, MLMC, and MLQMC, provide the theoretic order of convergence of $j2^{-j}$ for the approximate mean with respect to the H^1 -norm (picture on the left) and for the approximate second moment with respect to the $W^{1,1}$ -norm (picture on the right). Indeed, it seems that the logarithmic term j which stems from the summation of the multilevel error does not occur here. Moreover, the choice of $\delta = 0$, i.e. linearly increasing number of samples, seems sufficient in MLQMC to maintain the optimal rate of convergence. Since MLGQ provides a higher convergence rate for the second moment, it seems that the finite element error for the computation of the second moment converges with a higher rate than expected in Lemma 3.1. In addition, this suggests that MLGQ overestimates the quadrature error.

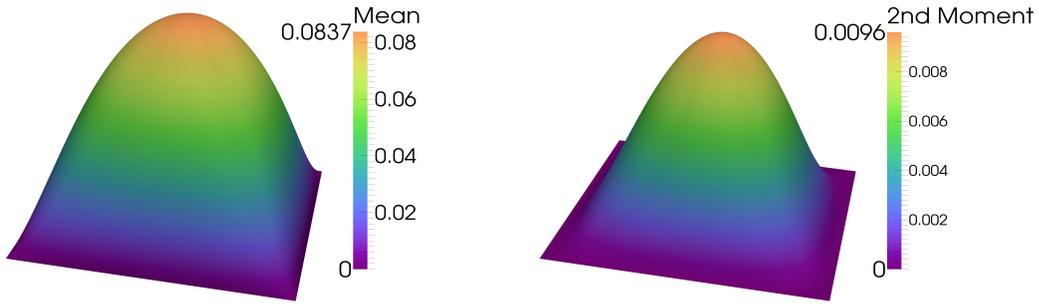


Figure 10.5. Solution's mean (left) and solution's uncentered second moment in case of the second example.

10.2. An example with a covariance of infinite rank. For our second example, we consider the covariance function $\text{Cov}_b(\mathbf{x}, \mathbf{y}) = 0.5 \exp(-2\|\mathbf{x} - \mathbf{y}\|_2^2)$. The computational domain is again the unit square $D = (0, 1)^2$ which is triangulated as seen in the right picture of Figure 10.1. The load vector is again $f \equiv 1$. The Karhunen-Loève expansion for the reference solution is approximated with a trace error of 10^{-12} , cf. [25], which results in $m = 72$ terms. Figure 10.5 shows the solution's mean (picture on the left) and its variance (picture on the right). Notice that the QMC approach with auxiliary density is here no longer feasible since m tends to infinity, cf. Section 7. Therefore, we apply the QMC approach without auxiliary density. According to [24], convergence can then be shown if the sequence $\{\gamma_k\}_k$ decays sufficiently fast. Nevertheless, we emphasize that, in case of a finite and relatively small stochastic dimensionality like in the previous example, the QMC approach with auxiliary density performs better.

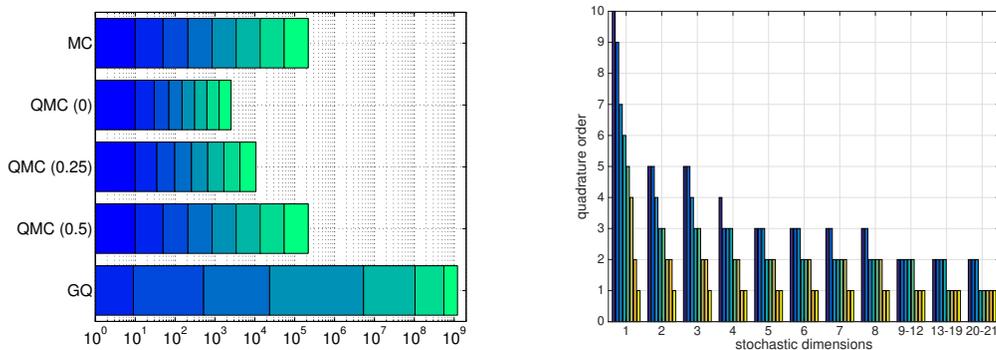


Figure 10.6. Number of samples for the different quadrature methods for the second example (left) and related orders of the Gaussian quadrature (right).

In order to preserve the approximation order of 2^{-j} , the Karhunen-Loève expansion is truncated after $m = 21$ terms on the finest level of computation $j = 8$. The number of samples for MLMC and MLQMC is chosen as in the previous example. For MLGQ, to account for the anisotropies in the integrand, we choose in this example $N_k \approx \log(\varepsilon)/\log(\tau_k)$, which is in accordance with the related formula from [2] for the case of a bounded parameter

| | method | MC | QMC($\delta = 0.5$) | QMC($\delta = 0.25$) | QMC($\delta = 0$) | GQ |
|-----------|--------------|--------|-----------------------|------------------------|---------------------|-----------|
| Example 1 | single-level | 51200 | 10240 | 1015 | 320 | 456976 |
| | multilevel | 300 | 60 | 25 | 20 | 2371 |
| | speed-up | 171 | 171 | 41 | 16 | 193 |
| Example 2 | single-level | 819200 | 163840 | 6450 | 1280 | 663552000 |
| | multilevel | 400 | 80 | 26 | 20 | 267749 |
| | speed-up | 2048 | 2048 | 248 | 64 | 2478 |

Table 11.1

Cost of the multilevel quadrature methods and the related single-level methods in terms of fine-grid samples.

domain. Although the results of Section 8 imply that we have to choose τ_k with respect to $3\tilde{\gamma}_k/\log 2$, we set for reasons of computability $\tau_k = \log 2/\gamma_k$ which already seems to provide a sufficiently small quadrature error. The related orders of the Gaussian quadrature in each single dimension and on each level are illustrated in the right plot of Figure 10.6. The number

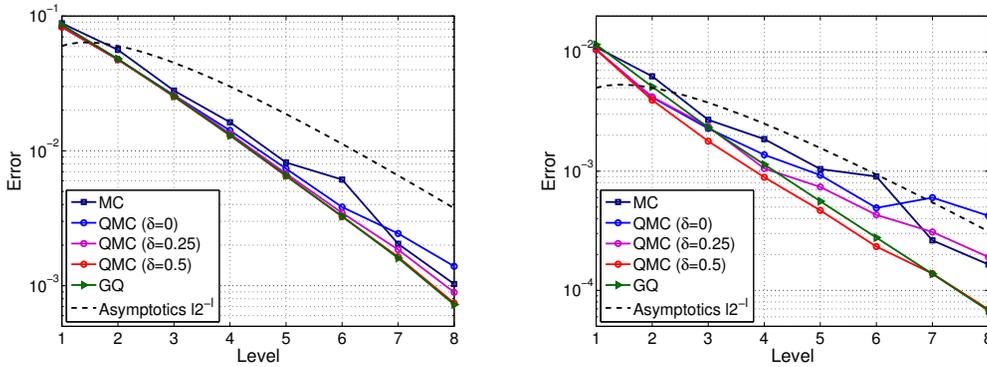


Figure 10.7. Error of the mean (left) and of the second moment (right) in case of the second example.

of samples obtained is shown in the left plot of Figure 10.6. The related computational errors are depicted in Figure 10.7. We achieve similar results as in the previous example. As before, it seems that the logarithmic factor j does not occur in the computation of the mean. In case of the second moment, only MLQMC with $\delta = 0.5$ and MLGQ provide a linear convergence rate 2^{-j} . The convergence rate of MLQMC with $\delta = 0.25$ and MLMC is rather $j2^{-j}$. For MLQMC with $\delta = 0$, the convergence rate even stagnates on level 6, 7 and 8. This suggests that all methods except for MLQMC with $\delta = 0$ yield the expected convergence rate and that the quadrature settings of MLQMC with $\delta = 0.5$ and MLGQ slightly overestimate the quadrature error. Nevertheless, a higher order convergence of MLGQ for the second moment, as in the previous example, is not obtained here.

11. Conclusion. In this article, we have presented the analysis for approximating the moments of the solution to partial differential equations with log-normally distributed diffusion coefficient by multilevel quadrature techniques. Our theoretical results as well as our numerical findings corroborate that any given single-level quadrature method can significantly be accelerated by multilevel techniques. To quantify this gain of the multilevel quadrature, we

list in Table 11.1 the total cost of the different quadrature methods for $j = 6$ levels in the first example and for $j = 8$ levels in the second example. The total cost of the multilevel Monte Carlo method, includes now that we average five runs. The cost for the multilevel methods is measured in terms of samples on the finest grid. We determine this number by summing up the number of samples on all level ℓ scaled by the respective work $2^{-n(j-\ell)}$. The cost are found in the row entitled by “single-level” and “multilevel” for the single-level and multilevel quadrature, respectively. In addition, Table 11.1 contains the ratios between these cost in the row entitled by “speed-up”. As can be seen, we obtain a remarkable acceleration for each particular quadrature method up to a factor 2478 in case of the Gaussian quadrature for the second example.

REFERENCES

- [1] R. A. ADAMS, *Sobolev Spaces*, Academic Press, New York, 1975.
- [2] I. BABUŠKA, F. NOBILE, AND R. TEMPONE, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM J. Numer. Anal., 45 (2007), pp. 1005–1034.
- [3] A. BARTH, C. SCHWAB, AND N. ZOLLINGER, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, Numer. Math., 119 (2011), pp. 123–161.
- [4] J. BECK, R. TEMPONE, F. NOBILE, AND L. TAMELLINI, *On the optimal polynomial approximation of stochastic pdes by Galerkin and collocation methods*, Math. Models Methods Appl. Sci., 22 (2012), p. 1250023.
- [5] M. BIERI, *A sparse composite collocation finite element method for elliptic sPDEs*, SIAM J. Numer. Anal., 49 (2011), pp. 2277–2301.
- [6] M. BIERI, R. ANDREEV, AND C. SCHWAB, *Sparse tensor discretization of elliptic sPDEs*, SIAM J. Sci. Comput., 31 (2010), pp. 4281–4304.
- [7] S. C. BRENNER AND L. R. SCOTT, *The mathematical theory of finite element methods*, Springer, Berlin, 3rd ed., 2008.
- [8] H.-J. BUNGARTZ AND M. GRIEBEL, *Sparse grids*, Acta Numer., 13 (2004), pp. 147–269.
- [9] R. CAFLISCH, *Monte Carlo and quasi-Monte Carlo methods*, Acta Numer., 7 (1998), pp. 1–49.
- [10] J. CHARRIER, *Strong and weak error estimates for elliptic partial differential equations with random coefficients*, SIAM J. Numer. Anal., 50 (2012), pp. 216–246.
- [11] J. CHARRIER, R. SCHEICHL, AND A. TECKENTRUP, *Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods*, SIAM J. Numer. Anal., 51 (2013), pp. 322–352.
- [12] K. A. CLIFFE, M.B. GILES, R. SCHEICHL, AND A.L. TECKENTRUP, *Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients*, Comput. Vis. Sci., 14 (2011), pp. 3–15.
- [13] P. FRAUENFELDER, C. SCHWAB, AND R. TODOR, *Finite elements for elliptic problems with stochastic coefficients*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 205–228.
- [14] R. GHANEM AND P. SPANOS, *Stochastic finite elements: a spectral approach*, Springer, New York, 1991.
- [15] M. GILES, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.
- [16] M. GILES AND B. WATERHOUSE, *Multilevel quasi-Monte Carlo path simulation*, Radon Series Comp. Appl. Math., 8 (2009), pp. 1–18.
- [17] C. J. GITTELSON, *Stochastic Galerkin discretization of the log-normal isotropic diffusion problem.*, Math. Models Methods Appl. Sci., 20 (2010), pp. 237–263.
- [18] C. J. GITTELSON, *Convergence rates of multilevel and sparse tensor approximations for a random elliptic PDE*, SIAM J. Numer. Anal., 51 (2013), pp. 2426–2447.
- [19] I. GRAHAM, F. KUO, J. NICHOLS, R. SCHEICHL, C. SCHWAB, AND I. SLOAN, *Quasi-Monte Carlo finite element methods for elliptic PDEs with log-normal random coefficient*, Numer. Math., 131 (2015), pp. 329–368.
- [20] M. GRIEBEL AND H. HARBRECHT, *On the construction of sparse tensor product spaces*, Math. Comput., 82 (2013), pp. 975–994.

- [21] J. H. HALTON, *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals*, Numer. Math., 2 (1960), pp. 84–90.
- [22] H. HARBRECHT, M. PETERS, AND R. SCHNEIDER, *On the low-rank approximation by the pivoted Cholesky decomposition*, Appl. Numer. Math., 62 (2012), pp. 28–440.
- [23] H. HARBRECHT, M. PETERS, AND M. SIEBENMORGEN, *On multilevel quadrature for elliptic stochastic partial differential equations*, Sparse Grids and Applications, 88 (2013), pp. 161–179.
- [24] ———, *On the quasi-Monte Carlo method with Halton points for elliptic PDEs with log-normal diffusion*, Preprint 2013-28 Mathematisches Institut Universität Basel, (2013). (to appear in Math. Comput.).
- [25] ———, *Efficient approximation of random fields for numerical applications*, Numer. Linear Algebra Appl., 22 (2015), pp. 596–617.
- [26] S. HEINRICH, *The multilevel method of dependent tests*, in Advances in stochastic simulation methods (St. Petersburg, 1998), Stat. Ind. Technol., Birkhäuser Boston, Boston, MA, 2000, pp. 47–61.
- [27] ———, *Multilevel Monte Carlo methods*, in Lect. Notes in Large Scale Scientific Computing, London, 2001, Springer, pp. 58–67.
- [28] F. J. HICKERNELL, I. H. SLOAN, AND G. W. WASILKOWSKI, *On tractability of weighted integration over bounded and unbounded regions in reals.*, Math. Comput., 73 (2004).
- [29] V. A. HOANG AND C. SCHWAB, *N-term Wiener chaos approximation rate for elliptic PDEs with lognormal Gaussian random inputs*, Math. Models Methods Appl. Sci., 4 (2014), p. 797826.
- [30] F. Y. KUO, C. SCHWAB, AND I. H. SLOAN, *Multi-level quasi-Monte Carlo finite element methods for a class of elliptic PDEs with random coefficients*, Found. Comput. Math., 15 (2015), pp. 411–449.
- [31] M. LOÈVE, *Probability theory. I+II*, vol. 45 of Graduate Texts in Mathematics, Springer, New York, 4th ed., 1977.
- [32] B. MORO, *The full Monte*, Risk Magazine, 8 (1995), pp. 57–58.
- [33] H. NIEDERREITER, *Random number generation and quasi-Monte Carlo methods*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.
- [34] F. NOBILE, R. TEMPONE, AND C. G. WEBSTER, *An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2411–2442.
- [35] E. NOVAK AND H. WOZNIAKOWSKI, *Tractability of Multivariate Problems: Linear Information*, EMS Tracts in Mathematics, European Mathematical Society, 2008.
- [36] A. B. OWEN, *Halton sequences avoid the origin*, SIAM Rev., 48 (2006), pp. 487–503.
- [37] C. SCHWAB AND C. J. GITTELSON, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Numer., 20 (2011), pp. 291–467.
- [38] C. G. SIMADER, *On Dirichlet’s Boundary Value Problem: LP-Theory based on a Generalization of Gårding’s Inequality*, Lecture Notes in Mathematics, Springer, Berlin-Heidelberg-New York, 1972.
- [39] I. H. SLOAN AND H. WOZNIAKOWSKI, *When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?*, J. Complexity, 14 (1997), pp. 1–33.
- [40] A. TECKENTRUP, R. SCHEICHL, M.B. GILES, AND E. ULLMANN, *Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients*, Numer. Math., 125 (2013), pp. 569–600.
- [41] J. V. USPENSKY, *On the convergence of quadrature formulas related to an infinite interval*, Trans. Amer. Math. Soc., 30 (1928), pp. 542–559.
- [42] C. ZENGER, *Sparse grids*, in Parallel Algorithms for Partial Differential Equations, W. Hackbusch, ed., vol. 31 of Notes on Numerical Fluid Mechanics, Braunschweig/Wiesbaden, 1991, Vieweg, pp. 241–251.