Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients

Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients*

F.Y. Kuo[†], Ch. Schwab and I.H. Sloan[†]

[†]School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia

QUASI-MONTE CARLO FINITE ELEMENT METHODS FOR A CLASS OF ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS WITH RANDOM COEFFICIENTS[*]

FRANCES Y. KUO[†§¶], CHRISTOPH SCHWAB[‡‖], AND IAN H. SLOAN[†¶]

Abstract. In this paper quasi-Monte Carlo (QMC) methods are applied to a class of elliptic partial differential equations (PDEs) with random coefficients, where the random coefficient is parametrized by a countably infinite number of terms in a Karhunen-Loève expansion. Models of this kind appear frequently in numerical models of physical systems, and in uncertainty quantification. The method uses a QMC method to estimate expected values of linear functionals of the exact or approximate solution of the PDE, with the expected value considered as an infinite dimensional integral in the parameter space corresponding to the randomness induced by the random coefficient. The analysis exploits the regularity with respect to both the physical variables (the variables in the physical domain) and the parametric variables (the parameters corresponding to randomness). As is common for the analysis of QMC methods, "weights", describing the varying difficulty of different subsets of the variables, are needed in the analysis in order to make sure that the infinite dimensional integration problem is tractable. It turns out that the weights arising from the present analysis are of a non-standard kind, being of neither product nor order-dependent form, but instead a hybrid of the two — we refer to these as "product and order-dependent weights", or "POD weights" in short. Nevertheless these POD weights are of a simple enough form to permit a component-by-component construction of a randomly shifted lattice rule that has optimal convergence properties for the given weighted space setting. If the terms in the expansion for the random coefficient have an appropriate decay property, and if we choose (POD) weights that minimize a certain upper bound on the error, then the solution of the PDE belongs to the joint function space needed for the analysis, and the QMC error (in the sense of a root-mean-square error averaged over shifts) is of order $\mathcal{O}(N^{-1+\delta})$ for arbitrary $\delta > 0$, where N denotes the number of sampling points in the parameter space. Moreover, for convergence rates less than 1, the conditions under which various convergence rates are achieved are exactly those found in a recent study by Cohen, De Vore and Schwab of the same model by best N-term approximations. We analyze the impact of a finite element (FE) discretization on the overall efficiency of the scheme, in terms of accuracy versus overall cost, with results that are comparable to those of the best N-term approximation.

Key words. quasi-Monte Carlo methods, infinite dimensional integration, elliptic partial differential equations with random coefficients, Karhunen-Loève expansion, finite element methods

AMS subject classifications. 65D30, 65D32, 65N30

1. Introduction. In this paper we analyze theoretically the application of quasi-Monte Carlo (QMC) methods combined with finite element (FE) methods to a class of elliptic partial differential equations (PDEs) with random coefficients, where the random coefficient is parametrized by a countably infinite number of parameters. There are a number of applications where such parametric diffusion problems need to be solved: let us only mention diffusion in random heterogeneous media where the

random coefficient is written in a separated expansion such as the Karhunen-Loève expansion (see, e.g. [27, 33] and the references there).

Specifically, we consider elliptic problems of the form

$$-\nabla \cdot (a(\boldsymbol{x}, \boldsymbol{y}) \, \nabla u(\boldsymbol{x}, \boldsymbol{y})) \; = \; f(\boldsymbol{x}) \quad \text{in} \quad D \subset \mathbb{R}^d \,, \quad u(\boldsymbol{x}, \boldsymbol{y}) \; = \; 0 \quad \text{on} \quad \partial D, \qquad (1.1)$$

for $D \subset \mathbb{R}^d$ a bounded domain with a Lipschitz boundary ∂D. The spatial dimension d is 1, 2 or 3. In (1.1), the gradients are understood to be with respect to the physical variable $\boldsymbol{x}$ which belongs to D, and the parameter vector $\boldsymbol{y} = (y_j)_{j \geq 1}$ consists of a countable number of parameters y_j which are assumed to be i.i.d. uniformly distributed, with

$$\boldsymbol{y} \in (-\tfrac{1}{2}, \tfrac{1}{2})^{\mathbb{N}} \; =: \; U \,.$$

The parameter $\boldsymbol{y}$ is then distributed on U with probability measure μ, where

$$\mu(\mathrm{d}\boldsymbol{y}) \; = \; \bigotimes_{j \geq 1} \mathrm{d}y_j \; = \; \mathrm{d}\boldsymbol{y}$$

is the uniform probability measure on U.

The parametric diffusion coefficient $a(\boldsymbol{x}, \boldsymbol{y})$ in (1.1) is assumed to depend linearly on the parameters y_j. Thus we assume

$$a(\boldsymbol{x}, \boldsymbol{y}) \; = \; \bar{a}(\boldsymbol{x}) + \sum_{j \geq 1} y_j \, \psi_j(\boldsymbol{x}) \,, \qquad \boldsymbol{x} \in D \,, \quad \boldsymbol{y} \in U \,. \qquad (1.2)$$

The mean of the field $a(\boldsymbol{x}, \cdot)$ is $\bar{a}(\boldsymbol{x})$, and the covariance is given by

$$\mathbb{E}\left[(a(\boldsymbol{x}, \cdot) - \bar{a}(\boldsymbol{x})) \, (a(\boldsymbol{x}', \cdot) - \bar{a}(\boldsymbol{x}'))\right] \; = \; \int_U (a(\boldsymbol{x}, \boldsymbol{y}) - \bar{a}(\boldsymbol{x})) \, (a(\boldsymbol{x}', \boldsymbol{y}) - \bar{a}(\boldsymbol{x}')) \, \mathrm{d}\boldsymbol{y}$$

$$= \; \sum_{j \geq 1} \psi_j(\boldsymbol{x}) \, \psi_j(\boldsymbol{x}') \,,$$

where by the integral over U we mean

$$\int_U F(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \; := \; \lim_{s \to \infty} \int_{(-\frac{1}{2}, \frac{1}{2})^s} F(y_1, \ldots, y_s, 0, 0, \ldots) \, \mathrm{d}y_1 \cdots \mathrm{d}y_s \,.$$

In order to ensure that the coefficient $a(\boldsymbol{x}, \boldsymbol{y})$ is well-defined for all parameter vectors $\boldsymbol{y} \in U$, we assume that

$$\bar{a} \in L^\infty(D) \,, \qquad \sum_{j \geq 1} \|\psi_j\|_{L^\infty(D)} \; < \; \infty \,, \qquad (1.3)$$

and also that the mean coefficient $\bar{a}$ and the infinite sum in (1.3) are such that for some positive numbers $a_{\min}$ and $a_{\max}$ we have

$$0 \; < \; a_{\min} \leq a(\boldsymbol{x}, \boldsymbol{y}) \leq a_{\max} \,, \qquad \boldsymbol{x} \in D \,, \quad \boldsymbol{y} \in U \,. \qquad (1.4)$$

Later we shall impose further smoothness assumptions on $\bar{a}$ and ψ_j as required.

Our aim in this paper is the efficient computation by QMC integration of expected values of continuous linear functionals of the solution, or the FE approximation of the

solution, of (1.1). Suppose the linear functional is $G(\cdot) : H_0^1(D) \to \mathbb{R}$. Then we are interested in computing expected values of

$$F(\boldsymbol{y}) := G(u(\cdot, \boldsymbol{y})), \qquad \boldsymbol{y} \in U, \tag{1.5}$$

or later the same functional applied to the FE solution $u_h \in H_0^1(D)$. The expected value of F is an integral of the functional $G(\cdot)$ of the parametric solution:

$$\int_U F(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} = \int_U G(u(\cdot, \boldsymbol{y})) \, \mathrm{d}\boldsymbol{y}.$$

We note that this involves integration of the parametric solution over an infinite dimensional domain of integration. We also observe that for the parametric boundary value problem (1.1) under consideration here, to evaluate F at any single point $\boldsymbol{y} \in U$ (as is typically needed in connection with quadrature formulas) requires the solution of a boundary value problem for $u(\boldsymbol{x}, \boldsymbol{y})$, which is more expensive than typical function evaluation, and introduces, through numerical solution of the PDE, an additional layer of discretization error.

The model studied here is exactly the same as in the recent paper [5], in which the method used was a best N-term Galerkin approximation. It will be a central concern to compare rates of convergence, and conditions on ψ_j under which the rates are achieved, with the results of [5]. In both [5] and the present paper, the summability of the fluctuation coefficients ψ_j plays an important role. Accordingly, we will make the assumption, stronger than (1.3), that there exists $0 < p \le 1$ such that

$$\sum_{j \ge 1} \|\psi_j\|_{L^\infty(D)}^p < \infty. \tag{1.6}$$

Notice that this condition implies decay of the fluctuation coefficients ψ_j, with stronger decay required as the value of p becomes smaller. In both [5] and the present paper, the rate of convergence $\mathcal{O}(N^{-1+\delta})$ occurs if (1.6) is satisfied with $p = 2/3$, which is the case if, e.g. $\|\psi_j\|_{L^\infty(D)} \le c\,j^{-3/2-\epsilon}$ for some $\epsilon > 0$. In this paper N denotes the number of QMC points. For values of p between $2/3$ and 1, the rate of convergence in both cases is $\mathcal{O}(N^{-(1/p-1/2)})$.

As usual when working with QMC integration, it turns out to be necessary to introduce numerical parameters $\gamma_{\mathfrak{u}}$, known as "weights", to describe the relative importance of the subset of variables with indices in the finite subset $\mathfrak{u} \subset \mathbb{N}$. The choice of weights $\gamma_{\mathfrak{u}}$ is a delicate question, because the weights must be chosen to satisfy two competing objectives: on the one hand, to ensure that the "worst case error" (see (2.3) ahead) is finite, and on the other hand that the solution u has finite norm in the corresponding function space. In the present work, we choose the weights to minimize a certain upper bound on the product of the worst case error and the norm. The weights that we need in order to make the analysis hold turn out to be of a non-standard kind — they are of neither "product" nor "order-dependent" form (see, e.g., [31, 10, 32] for more on this terminology) but instead are a hybrid of the two. Specifically, the weight $\gamma_{\mathfrak{u}}$ associated with the subset of variables $\{y_j : j \in \mathfrak{u}\}$ is of the form

$$\gamma_{\mathfrak{u}} = \Gamma_{|\mathfrak{u}|} \prod_{j \in \mathfrak{u}} \gamma_j, \tag{1.7}$$

where $|\mathfrak{u}|$ denotes the cardinality (or the "order") of $\mathfrak{u}$. The weights are therefore determined by a specific choice of the two sequences $\Gamma_0 = \Gamma_1 = 1, \Gamma_2, \Gamma_3, \ldots$ and

$\gamma_1, \gamma_2, \gamma_3, \ldots$. We shall refer to this form of weights as "product and order-dependent weights", or "POD weights" for short. (See (6.5) or (6.7) ahead for our precise choice of weights.)

Recall that the purpose of the present paper is to analyze the accuracy and complexity of QMC methods in connection with approximate solution of (1.1) by FE methods. To guarantee that the FE solutions converge, we will make the assumption, stronger than (1.3), that

$$\bar{a} \in W^{1,\infty}(D) , \qquad \sum_{j \geq 1} \|\psi_j\|_{W^{1,\infty}(D)} < \infty , \tag{1.8}$$

where $\|v\|_{W^{1,\infty}(D)} = \max\{\|v\|_{L^\infty(D)}, \|\nabla v\|_{L^\infty(D)}\}$. In practice the infinite sum in (1.2) must be truncated to a finite sum of, say, s terms. For this truncation to make sense, we will assume additionally that the ψ_j are ordered so that $\|\psi_j\|_{L^\infty(D)}$ is non-increasing:

$$\|\psi_1\|_{L^\infty(D)} \geq \|\psi_2\|_{L^\infty(D)} \geq \|\psi_3\|_{L^\infty(D)} \geq \cdots . \tag{1.9}$$

The overall error for our QMC-FE approximation is then a sum of three terms: a truncation error, a QMC error, and a FE error. We will obtain results for estimating the three errors and finally combine them to arrive at an overall error bound.

The outline of the paper is as follows. In §2 we review QMC integration, first for finite dimensional integration, and then for the recent topic of integration with an infinite number of variables. In §3 we introduce the necessary function spaces, derive a weak formulation of the parametric equation (1.1), and establish the existence and uniqueness of solutions. In §4 we investigate the regularity of the parametric solutions with respect to the the spatial variable $\boldsymbol{x}$, as required for the analysis of the FE approximation, and also the regularity with respect to the parametric variable $\boldsymbol{y}$, as required for the analysis of the QMC integration. The next three sections, §5, §6, and §7 are devoted to estimating the truncation error, the QMC error for the exact solution of PDE, and the FE error, respectively. In particular, in §6 we establish suitable weights to ensure that the exact solution of the PDE belongs to the particular infinite dimensional space, and find convergence rate with respect to the number of QMC points N, under the condition (1.6). Then in §8 we combine the three error estimates to obtain an overall QMC-FE error bound. Finally in §9 we give some concluding remarks.

2. Quasi-Monte Carlo Integration in Weighted Spaces.

2.1. QMC integration in the finite dimensional setting. In this subsection we consider QMC integration when the dimension (or the number of integration variables), denoted here by s, is assumed to be finite and fixed. Here the domain of integration is taken to be the s-dimensional unit cube $[-\frac{1}{2}, \frac{1}{2}]^s$ centered at the origin. This is different from the usual QMC convention where the unit cube is $[0,1]^s$. However, all existing QMC results can be applied to $[-\frac{1}{2}, \frac{1}{2}]^s$ by making a trivial translation.

In this subsection we consider integrals of the form

$$I_s(F) := \int_{[-\frac{1}{2}, \frac{1}{2}]^s} F(\boldsymbol{y}) \, d\boldsymbol{y} .$$

In our later applications F will be of the form (1.5), but for the present it is general and depends only on s variables. An N-point QMC approximation to this integral is

an equal-weight rule of the form

$$Q_{s,N}(F) := \frac{1}{N} \sum_{i=1}^{N} F(\boldsymbol{y}^{(i)}),$$

with carefully chosen points $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(N)} \in [-\frac{1}{2}, \frac{1}{2}]^s$. For classical results on QMC methods, see, e.g. [22, 29].

We shall assume that our integrand F belongs to a *weighted* and *anchored* Sobolev space $\mathcal{W}_{s,\boldsymbol{\gamma}}$ which is a Hilbert space containing functions defined over the unit cube $[-\frac{1}{2}, \frac{1}{2}]^s$, with square integrable mixed first derivatives. More precisely, the norm is given by

$$\|F\|_{\mathcal{W}_{s,\boldsymbol{\gamma}}} := \left(\sum_{\mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{-1} \int_{[-\frac{1}{2}, \frac{1}{2}]^{|\mathfrak{u}|}} \left| \frac{\partial^{|\mathfrak{u}|} F}{\partial \boldsymbol{y}_{\mathfrak{u}}}(\boldsymbol{y}_{\mathfrak{u}}; 0) \right|^2 \mathrm{d}\boldsymbol{y}_{\mathfrak{u}} \right)^{1/2}, \qquad (2.1)$$

where $\{1 : s\}$ is a shorthand notation for the set of indices $\{1, 2, \ldots, s\}$, $\frac{\partial^{|\mathfrak{u}|} F}{\partial \boldsymbol{y}_{\mathfrak{u}}}$ denotes the mixed first derivative with respect to the variables y_j with $j \in \mathfrak{u}$, and $(\boldsymbol{y}_{\mathfrak{u}}; 0)$ denotes the vector whose jth component is y_j if $j \in \mathfrak{u}$ and 0 if $j \notin \mathfrak{u}$.

Weighted spaces were first introduced by Sloan and Woźniakowski in [31]. By now there are many variants and generalizations, see e.g. [10, 32]. In (2.1) the "anchor" is $(0, \ldots, 0)$, the center of the unit cube $[\frac{1}{2}, \frac{1}{2}]^s$. (This corresponds to the anchor $(\frac{1}{2}, \ldots, \frac{1}{2})$ in the standard unit cube $[0, 1]^s$.) Traditionally the anchor is often taken at a cube corner, but in the present application to the PDE problem it is more natural, and leads to marginally better results, to place the anchor at the center rather than a corner of the unit cube. Here we consider "general weights" following [32]: there is a weight parameter $\gamma_{\mathfrak{u}} \geq 0$ associated with each group of variables $\boldsymbol{y}_{\mathfrak{u}} = (y_j)_{j \in \mathfrak{u}}$ with indices belonging to the set $\mathfrak{u}$, with the convention that $\gamma_{\emptyset} = 1$. If $\gamma_{\mathfrak{u}} = 0$ then we demand that the corresponding integral of the mixed first derivative is also zero, and we follow the convention $0/0 = 0$. This is a generalization of the more traditional choice of "product weights", see e.g. [31, 10], which assumes that there is one weight parameter γ_j associated with each variable y_j, and the weight associated with the group of variables $\boldsymbol{y}_{\mathfrak{u}}$ is given by the product $\gamma_{\mathfrak{u}} = \prod_{j \in \mathfrak{u}} \gamma_j$. We shall assume, as seems appropriate in our ultimately infinite dimensional setting, that our weights do not depend on the dimension s. (Dependence of the weights on s was allowed in [10, 32].)

Apart from product weights, other forms of weights have been considered in the literature, including "order-dependent weights" and "finite-order weights", see e.g. [32]. As we already discussed in the introduction, the weights that arise from the analysis in this paper are of a special form, which we call "product and order-dependent weights", or "POD weights" for short, see (1.7). This special form of weights has not been considered before, but these POD weights arise naturally in our analysis, and seem to be crucial for achieving the same convergence rate as the best N-term approximation rate obtained in [5].

The norm of I_s as a linear functional on our function space $\mathcal{W}_{s,\boldsymbol{\gamma}}$ is, from [32],

$$\|I_s\| := \sup_{\|F\|_{\mathcal{W}_{s,\boldsymbol{\gamma}}} \leq 1} |I_s(F)| = \left(\sum_{\mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}} \left(\frac{1}{12} \right)^{|\mathfrak{u}|} \right)^{1/2}.$$

We shall assume that we have a sequence of nonnegative weights $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{|\mathfrak{u}|<\infty}$ satisfying

$$\sum_{|\mathfrak{u}|<\infty} \gamma_{\mathfrak{u}} \left(\frac{1}{12}\right)^{|\mathfrak{u}|} < \infty . \tag{2.2}$$

This condition ensures that $\|I_s\|$ is bounded independently of s, which in turn ensures that the integration problem in the later infinite dimensional setting is well defined.

Many recent papers analyzed the *worst case error* of a QMC rule (or a family of QMC rules) over all functions in the unit ball of $\mathcal{W}_{s,\boldsymbol{\gamma}}$, i.e.,

$$e^{\mathrm{wor}}(Q_{s,N}; \mathcal{W}_{s,\boldsymbol{\gamma}}) := \sup_{\|F\|_{\mathcal{W}_{s,\boldsymbol{\gamma}}} \leq 1} |I_s(F) - Q_{s,N}(F)| . \tag{2.3}$$

There is an explicit expression for $e^{\mathrm{wor}}(Q_{s,N}; \mathcal{W}_{s,\boldsymbol{\gamma}})$ which allows it to be analyzed in theory and computed in practice. Various upper bounds for $e^{\mathrm{wor}}(Q_{s,N}; \mathcal{W}_{s,\boldsymbol{\gamma}})$ have been obtained for different families of QMC rules; some are non-constructive, while some are semi- or fully constructive. Of particular interest are bounds of the form $e^{\mathrm{wor}}(Q_{s,N}; \mathcal{W}_{s,\boldsymbol{\gamma}}) \leq C\, N^{-r}$ with r close to 1, which is optimal in $\mathcal{W}_{s,\boldsymbol{\gamma}}$, and with C independent of the dimension s, which can hold if certain conditions on the weights $\boldsymbol{\gamma}$ are satisfied. Note that due to linearity of the functionals $I_s(\cdot)$ and $Q_{s,N}(\cdot)$, we have

$$|I_s(F) - Q_{s,N}(F)| \leq e^{\mathrm{wor}}(Q_{s,N}; \mathcal{W}_{s,\boldsymbol{\gamma}}) \|F\|_{\mathcal{W}_{s,\boldsymbol{\gamma}}} \qquad \text{for all} \quad F \in \mathcal{W}_{s,\boldsymbol{\gamma}} . \tag{2.4}$$

In this paper we will focus on a family of QMC rules known as "shifted rank-1 lattice rules", because this is the family of QMC rules for which we are presently able to obtain comparable results to those in [5]. Shifted rank-1 lattice rules are QMC rules with quadrature points given by the simple formula

$$\boldsymbol{y}^{(i)} = \mathrm{frac}\left(\frac{i\boldsymbol{z}}{N} + \boldsymbol{\Delta}\right) - \left(\tfrac{1}{2}, \ldots, \tfrac{1}{2}\right) , \qquad i = 1, \ldots, N ,$$

where $\boldsymbol{z} \in \mathbb{Z}^s$ is known as the *generating vector*, $\boldsymbol{\Delta} \in [0,1]^s$ is the *shift*, and $\mathrm{frac}(\cdot)$ means to take the fractional part of each component in the vector. The subtraction by the vector $\left(\tfrac{1}{2}, \ldots, \tfrac{1}{2}\right)$ takes care of the translation from the usual unit cube $[0,1]^s$ to the unit cube $[-\tfrac{1}{2}, \tfrac{1}{2}]^s$ considered in this paper.

We present a relevant lattice rule convergence result in the theorem below. Here we consider shifted lattice rules with *random shifts*. In this case, the quality of the rules is determined by the choice of the generating vector $\boldsymbol{z}$. Assuming that this deterministic vector $\boldsymbol{z}$ is chosen and fixed, we denote the corresponding shifted lattice rule with shift $\boldsymbol{\Delta}$ by $Q_{s,N}(F; \boldsymbol{\Delta})$. In the following, $\zeta(x) = \sum_{k=1}^{\infty} k^{-x}$ denotes the Riemann zeta function.

THEOREM 2.1. *Let $s, N \in \mathbb{N}$ be given, with N a prime number, and assume $F \in \mathcal{W}_{s,\boldsymbol{\gamma}}$ for a particular choice of weights $\boldsymbol{\gamma}$. Then a randomly shifted lattice rule can be constructed using a component-by-component algorithm such that the root-mean-square error satisfies, for all $\lambda \in (1/2, 1]$,*

$$\sqrt{\mathbb{E}\left[|I_s(F) - Q_{s,N}(F; \cdot)|^2\right]}$$

$$\leq \left(\sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}} + \frac{1}{12^{\lambda}}\right)^{|\mathfrak{u}|}\right)^{1/(2\lambda)} (N-1)^{-1/(2\lambda)} \|F\|_{\mathcal{W}_{s,\boldsymbol{\gamma}}} , \tag{2.5}$$

where $\mathbb{E}[\cdot]$ *denotes the expectation with respect to the random shift which is uniformly distributed over* $[0,1]^s$.

This result with general weights and prime N is proved in [32, Theorem 3(A)]. The form of the error bound stated in (2.5) is perhaps not immediately apparent in [32]. (We need to take $a_j = 1/2$ and $m_j = 1/12$ in [32].) The necessary manipulations to convert into the present form will be described in the survey paper [18]. This result can be generalized to cover composite values of N at a cost of increasing the constant; more discussion will be given in [18].

The corresponding results for product weights are in earlier papers, and we briefly summarize them now. The component-by-component (CBC) algorithm for the construction of randomly shifted lattice rules is introduced in [30]. The generating vector $\boldsymbol{z}$ is chosen one component at a time, while minimizing a "shift-averaged" worst case error expression. The convergence result is proved in [17, 7]. The fast implementation using FFT is due to [24, 25]. The greedy nature of the algorithm means that the same error bound holds when the dimension s is replaced by any number smaller than s, and it also means that the lattice rule can be extended to higher dimensions at any time. The original algorithm requires a fixed value of N as input, and changing N means that the lattice rule has to be constructed anew. Modified algorithms for obtaining lattice rules that are extensible in N are given in [6, 9].

There are analogous results for "digitally-shifted polynomial lattice rules", which is another family of QMC rules, and there are also QMC rules with higher order convergence, see [8] and the references there. The analysis in [34] is based on the so-called Niederreiter and Sobol′ sequences which are low-discrepancy sequences that can be generated explicitly, and which are extensible in both s and N. The lattice rules in [16, 28] are constructed using a different error criterion in a non-Hilbert space setting. We do not consider these QMC rules in this paper because they do not presently give results that are as good as randomly shifted lattice rules. (We remark that the scaling of $\gamma_{\mathfrak{u}}$ in our definition (2.1) is consistent with [31, 34]. However, all results in [16, 28, 8] will be consistent with the notation in the present paper upon the substitution $\gamma_{\mathfrak{u}} \mapsto \gamma_{\mathfrak{u}}^{1/2}$.)

The CBC construction cost is exponential in s for general weights, or exponential in the order of finite-order weights. Fast CBC construction using FFT is possible at a cost of $\mathcal{O}(s\,N\ln N)$ operations for product weights or order-dependent weights, see [6]. It turns out that an extension of the fast CBC construction from order-dependent weights to POD weights (1.7) is straightforward; this extension will be discussed in [18]. For these special cases, extensible lattice sequences can be constructed at a cost of $\mathcal{O}(s\,N(\ln N)^2)$ operations, see [6].

2.2. QMC integration in the infinite dimensional setting. Here we follow the infinite dimensional setting in [20], but with the anchor at the center of the unit cube rather than at a corner. For F a function depending on infinitely many variables $\boldsymbol{y} = (y_1, y_2, \ldots)$, the integral of interest takes the form

$$I(F) := \lim_{s\to\infty} I_s(F), \qquad I_s(F) := \int_{[-\frac{1}{2},\frac{1}{2}]^s} F(y_1, \ldots, y_s, 0, 0, \ldots)\,\mathrm{d}y_1 \cdots \mathrm{d}y_s, \quad (2.6)$$

and an s-dimensional N-point QMC approximation to $I(F)$ is given by

$$Q_{s,N}(F) := \frac{1}{N}\sum_{i=1}^{N} F(y_1^{(i)}, \ldots, y_s^{(i)}, 0, 0, \ldots). \qquad (2.7)$$

We assume that F belongs to the weighted Sobolev space $\mathcal{W}_{\boldsymbol{\gamma}}$, which is the infinite dimensional version of $\mathcal{W}_{s,\boldsymbol{\gamma}}$, with the norm (2.1) replaced by

$$\|F\|_{\mathcal{W}_{\boldsymbol{\gamma}}} := \left(\sum_{|\mathfrak{u}| < \infty} \gamma_{\mathfrak{u}}^{-1} \int_{[-\frac{1}{2}, \frac{1}{2}]^{|\mathfrak{u}|}} \left| \frac{\partial^{|\mathfrak{u}|} F}{\partial \boldsymbol{y}_{\mathfrak{u}}} (\boldsymbol{y}_{\mathfrak{u}}; 0) \right|^2 \mathrm{d}\boldsymbol{y}_{\mathfrak{u}} \right)^{1/2} , \qquad (2.8)$$

where the sum is now over all subsets $\mathfrak{u} \subset \mathbb{N}$ with finite cardinality. This definition is consistent with (2.1) in the sense that, for a function that depends only on the first s variables, its norm in $\mathcal{W}_{\boldsymbol{\gamma}}$ is the same as its norm in $\mathcal{W}_{s,\boldsymbol{\gamma}}$. Moreover, for a function F that depends on infinitely many variables, if we define $F_s(y_1, \ldots, y_s) := F(y_1, \ldots, y_s, 0, 0, \cdots)$ by anchoring the components beyond dimension s at 0, then we have $\|F_s\|_{\mathcal{W}_{s,\boldsymbol{\gamma}}} = \|F_s\|_{\mathcal{W}_{\boldsymbol{\gamma}}} \leq \|F\|_{\mathcal{W}_{\boldsymbol{\gamma}}}$.

The integration problem in the infinite dimensional setting is well defined because

$$\|I\| := \sup_{\|F\|_{\mathcal{W}_{\boldsymbol{\gamma}}} \leq 1} |I(F)| = \left(\sum_{|\mathfrak{u}| < \infty} \gamma_{\mathfrak{u}} \left(\frac{1}{12} \right)^{|\mathfrak{u}|} \right)^{1/2} ,$$

which is finite due to Assumption (2.2).

We now restate Theorem 2.1 for the infinite dimensional setting. Note that the error bound (2.9) in the theorem is for the s-dimensional integral $I_s(F)$ rather than $I(F)$. The truncation error $I(F) - I_s(F)$ is still to be estimated separately. Observe, however, that the bound in (2.9) is independent of s.

THEOREM 2.2. *Suppose $F \in \mathcal{W}_{\boldsymbol{\gamma}}$ for a particular choice of weights $\boldsymbol{\gamma}$. Then for $s, N \in \mathbb{N}$ with N a prime number, a randomly shifted lattice rule can be constructed using a component-by-component algorithm such that the root-mean-square error for approximating the s-dimensional integral $I_s(F)$ satisfies, for all $\lambda \in (1/2, 1]$,*

$$\sqrt{\mathbb{E}\left[|I_s(F) - Q_{s,N}(F; \cdot)|^2\right]}$$

$$\leq \left(\sum_{|\mathfrak{u}| < \infty} \gamma_{\mathfrak{u}}^{\lambda} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}} + \frac{1}{12^{\lambda}} \right)^{|\mathfrak{u}|} \right)^{1/(2\lambda)} (N-1)^{-1/(2\lambda)} \|F\|_{\mathcal{W}_{\boldsymbol{\gamma}}} , \qquad (2.9)$$

where $\mathbb{E}[\cdot]$ denotes the expectation with respect to the random shift $\boldsymbol{\Delta}$ which is uniformly distributed over $[0, 1]^s$.

Note that the error bound (2.9) is only meaningful when the parameter λ is chosen such that the sum in (2.9) is finite.

3. Parameter-Dependent Variational Formulation.

3.1. Function spaces. We first introduce the function spaces needed in the paper.

Our variational setting of (1.1) is based on the Sobolev space $V = H_0^1(D)$ and its dual space $V^* = H^{-1}(D)$, with the norm in V given by

$$\|v\|_V := \|\nabla v\|_{L^2(D)} .$$

The duality between V^* and V is understood to be with respect to the pivot space $L^2(D)$, which we identify with its own dual. For $f \in V^*$ and for $v \in V$, we understand $\int_D f(\boldsymbol{x}) v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$ as extension by continuity of the $L^2(D)$ inner product to a

duality pairing $\langle f, v \rangle$ between V^* and V. With this convention, $\left| \int_D f(\boldsymbol{x}) \, v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right| \leq$ $\|f\|_{V^*} \|v\|_V$.

In addition to the function space V, we shall consider the function space with additional regularity with respect to $\boldsymbol{x}$,

$$Z := \{ v \in V : \Delta v \in L^2(D) \} . \tag{3.1}$$

Then $Z \subset V$ is a closed subspace which, when equipped with the norm

$$\|v\|_Z := \left(\|v\|_{L^2(D)}^2 + \|\Delta v\|_{L^2(D)}^2 \right)^{1/2} , \tag{3.2}$$

is a Hilbert space. By standard elliptic regularity theory (see, e.g. [11]), Z is known to coincide with $H^2_{\mathrm{loc}}(D)$, and for convex domains D we have $Z = H^2(D) \cap H^1_0(D)$.

Furthermore, as already seen in the introduction, we will make use of the following norm

$$\|v\|_{W^{1,\infty}(D)} := \max\{\|v\|_{L^\infty(D)}, \|\nabla v\|_{L^\infty(D)}\} . \tag{3.3}$$

We shall need function spaces defined with respect to the parameter $\boldsymbol{y}$, namely, the weighted Sobolev spaces $\mathcal{W}_{\boldsymbol{\gamma}}$ defined already in §2. Spaces which are defined in terms of both $\boldsymbol{x}$ and $\boldsymbol{y}$ will play an important role. Therefore, we define the weighted spaces $\mathcal{W}_{\boldsymbol{\gamma}}(U; V)$, which are Bochner versions of the weighted spaces $\mathcal{W}_{\boldsymbol{\gamma}}$, with the norm

$$\|u\|_{\mathcal{W}_{\boldsymbol{\gamma}}(U;V)} := \left(\sum_{|\mathfrak{u}| < \infty} \gamma_{\mathfrak{u}}^{-1} \int_{[-\frac{1}{2}, \frac{1}{2}]^{|\mathfrak{u}|}} \left\| \frac{\partial^{|\mathfrak{u}|} u}{\partial \boldsymbol{y}_{\mathfrak{u}}}(\cdot, (\boldsymbol{y}_{\mathfrak{u}}; 0)) \right\|_V^2 \mathrm{d}\boldsymbol{y}_{\mathfrak{u}} \right)^{1/2} . \tag{3.4}$$

Our goal in this paper is to compute

$$\int_U F(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} , \qquad \text{with} \qquad F(\boldsymbol{y}) = G(u(\cdot, \boldsymbol{y})) , \quad G(\cdot) \in V^* . \tag{3.5}$$

Then for $u \in \mathcal{W}_{\boldsymbol{\gamma}}(U; V)$, using (2.8) and

$$\frac{\partial^{|\mathfrak{u}|} F}{\partial \boldsymbol{y}_{\mathfrak{u}}}(\boldsymbol{y}) = G\left(\frac{\partial^{|\mathfrak{u}|} u}{\partial \boldsymbol{y}_{\mathfrak{u}}}(\cdot, \boldsymbol{y}) \right) ,$$

we have

$$\|F\|_{\mathcal{W}_{\boldsymbol{\gamma}}} \leq \|G(\cdot)\|_{V^*} \|u\|_{\mathcal{W}_{\boldsymbol{\gamma}}(U;V)} < \infty . \tag{3.6}$$

3.2. Parametric weak formulation. We derive the variational formulation of the parametric boundary value problem (1.1) for each value of the parameter $\boldsymbol{y} \in U$, and establish sufficient conditions for the existence and uniqueness of solutions.

For a fixed $\boldsymbol{y} \in U$, we multiply the PDE in (1.1) by a test function $v(\boldsymbol{x})$, and integrate by parts with respect to $\boldsymbol{x}$ using Green's formula and the homogeneous Dirichlet boundary conditions. This results in the following *parameter-dependent weak formulation of the parametric deterministic problem* (1.1): for $\boldsymbol{y} \in U$, find

$$u(\cdot, \boldsymbol{y}) \in V : \int_D a(\boldsymbol{x}, \boldsymbol{y}) \, \nabla u(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_D f(\boldsymbol{x}) \, v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \qquad \forall v \in V . \tag{3.7}$$

The parametric bilinear form $b(\boldsymbol{y}; w, v)$ for $\boldsymbol{y} \in U$ is given by

$$b(\boldsymbol{y}; w, v) := \int_D a(\boldsymbol{x}, \boldsymbol{y}) \, \nabla w(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \qquad \forall v \in V , \qquad (3.8)$$

which by (1.4) is continuous and coercive on $V \times V$, i.e., for all $\boldsymbol{y} \in U$ and all $v, w \in V$ we have

$$b(\boldsymbol{y}; v, v) \geq a_{\min} \|v\|_V^2 \quad \text{and} \quad |b(\boldsymbol{y}; v, w)| \leq a_{\max} \|v\|_V \|w\|_V .$$

We may then use the Lax-Milgram Lemma to infer that for every $f \in V^*$, there exists a unique solution to the parametric weak problem

$$u(\cdot, \boldsymbol{y}) \in V : \quad b(\boldsymbol{y}; u(\cdot, \boldsymbol{y}), v) = \langle f, v \rangle \qquad \forall v \in V ,$$

which satisfies the standard a-priori estimate. In this way we obtain the following theorem.

THEOREM 3.1. *Under Assumptions* (1.3) *and* (1.4), *for every* $f \in V^*$ *and every* $\boldsymbol{y} \in U$, *there exists a unique solution* $u(\cdot, \boldsymbol{y}) \in V$ *of the parametric weak problem* (3.7), *which satisfies*

$$\|u(\cdot, \boldsymbol{y})\|_V \leq \frac{\|f\|_{V^*}}{a_{\min}} .$$

We remark that since the result in the theorem holds parametrically for $\boldsymbol{y} \in U$, the result will hold equally if f becomes random, that is, if $f(\boldsymbol{x})$ in (1.1) is replaced by $f(\boldsymbol{x}, \boldsymbol{y})$, in which case $\|f\|_{V^*}$ in the theorem is replaced by $\|f(\cdot, \boldsymbol{y})\|_{V^*}$.

4. Regularity of the PDE Solution.

4.1. Regularity of $u(\boldsymbol{x}, \boldsymbol{y})$ with respect to $\boldsymbol{x}$. In order to quantify the regularity of the parametric solution u with respect to the variable $\boldsymbol{x}$, which is required for the convergence analysis of the FE discretization in the domain D, we confine ourselves to the regularity required for piecewise linear, continuous Finite Elements (higher order FE discretizations will require correspondingly refined regularity estimates which could be derived analogously). Accordingly, we will often make the stronger assumption that

$$f \in L^2(D) .$$

Standard elliptic regularity then implies that u will have additional regularity as a function of $\boldsymbol{x}$. Specifically, under suitable conditions the solution u will belong to the space Z in (3.1).

To ensure that the solutions of elliptic PDEs with nonconstant coefficients have regularity beyond $H^1(D)$, additional regularity of the coefficients is well-known to be required (see, e.g. [11]). Accordingly, we impose in addition to Assumptions (1.3) and (1.4) also Assumption (1.8). We now show that, under the stronger assumption $f \in L^2(D)$, we can obtain a bound on the Z norm of $u(\cdot, \boldsymbol{y})$ for each value of the parameter $\boldsymbol{y}$.

THEOREM 4.1. *Under Assumptions* (1.3), (1.4) *and* (1.8), *there exists a constant* $C > 0$ (*depending only on D and on the bounds in Assumption* (1.8)) *such that for*

every $f \in L^2(D)$ and every $\boldsymbol{y} \in U$, the solution $u(\cdot, \boldsymbol{y}) \in V$ of the parametric weak problem (3.7) *satisfies*

$$\|u(\cdot, \boldsymbol{y})\|_Z \leq C\|f\|_{L^2(D)} \, . \tag{4.1}$$

Proof. Assumption (1.8) implies for every $\boldsymbol{y} \in U$ that

$$\|a(\cdot, \boldsymbol{y})\|_{W^{1,\infty}(D)} \leq \|\bar{a}\|_{W^{1,\infty}(D)} + \frac{1}{2}\sum_{j \geq 1} \|\psi_j\|_{W^{1,\infty}(D)} < \infty \, ,$$

where $\|\cdot\|_{W^{1,\infty}(D)}$ is defined in (3.3).

We next apply in (1.1) the identity

$$\nabla \cdot (\alpha(\boldsymbol{x})\nabla w(\boldsymbol{x})) = \alpha(\boldsymbol{x})\,\Delta w(\boldsymbol{x}) + \nabla\alpha(\boldsymbol{x}) \cdot \nabla w(\boldsymbol{x}) \, ,$$

which is valid for $\alpha(\boldsymbol{x}) \in W^{1,\infty}(D)$ and for $w \in V$ such that $\Delta w \in L^2(D)$. We find that $u(\cdot, \boldsymbol{y})$ satisfies, for every $\boldsymbol{y} \in U$, the Poisson equation

$$-a(\cdot, \boldsymbol{y})\Delta u(\cdot, \boldsymbol{y}) = \nabla a(\cdot, \boldsymbol{y}) \cdot \nabla u(\cdot, \boldsymbol{y}) + f(\cdot) \quad \text{in} \quad D \, , \quad u(\cdot, \boldsymbol{y})|_{\partial D} = 0 \, .$$

This implies that for every $\boldsymbol{y} \in U$ there holds

$$a_{\min}\|\Delta u(\cdot, \boldsymbol{y})\|_{L^2(D)} \leq \|a(\cdot, \boldsymbol{y})\|_{W^{1,\infty}(D)}\|u(\cdot, \boldsymbol{y})\|_V + \|f\|_{L^2(D)} \, ,$$

and this yields

$$\|u(\cdot, \boldsymbol{y})\|_Z^2 \leq \|u(\cdot, \boldsymbol{y})\|_{L^2(D)}^2 + \frac{1}{a_{\min}^2}\left(\|a(\cdot, \boldsymbol{y})\|_{W^{1,\infty}(D)}\|u(\cdot, \boldsymbol{y})\|_V + \|f\|_{L^2(D)}\right)^2 \, ,$$

and hence

$$\|u(\cdot, \boldsymbol{y})\|_Z^2 \leq \left(\frac{1}{C_P^2} + \frac{2}{a_{\min}^2}\sup_{\boldsymbol{z} \in U}\|a(\cdot, \boldsymbol{z})\|_{W^{1,\infty}(D)}^2\right)\|u(\cdot, \boldsymbol{y})\|_V^2 + \frac{2}{a_{\min}^2}\|f\|_{L^2(D)}^2 \, ,$$

where $C_P = \inf_{v \in V}\|\nabla v\|_{L^2(D)}/\|v\|_{L^2(D)} > 0$ denotes the Poincaré constant. The proof is completed by using Theorem 3.1, together with the imbedding of $L^2(D)$ into V^*. □

4.2. Regularity of $u(\boldsymbol{x}, \boldsymbol{y})$ with respect to $\boldsymbol{y}$. In the ensuing QMC error analysis, we shall require bounds on the mixed first partial derivatives of the parametric solution u, which we will establish in the present subsection. To this end, for arbitrary $j \in \mathbb{N}$, we differentiate the variational form (3.7) of the parametric deterministic problem with respect to y_j, which results with (1.2) in the identity

$$\int_D a(\boldsymbol{x}, \boldsymbol{y})\,\nabla(\partial_{y_j}u(\boldsymbol{x}, \boldsymbol{y})) \cdot \nabla v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_D \psi_j(\boldsymbol{x})\,\nabla u(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0 \quad \forall v \in V \, ,$$

where $\partial_{y_j}u := \partial u/\partial y_j$.

It is clear that we can differentiate again with respect to y_j, or with respect to any other component of $\boldsymbol{y}$, and that indeed we can keep on differentiating, as the problem is completely smooth with respect to $\boldsymbol{y}$. While we need only the mixed first derivatives for the present application, we obtain here all partial derivatives since it is easy to do so. For this purpose we introduce a multi-index notation.

For $\boldsymbol{\nu} = (\nu_j)_{j \geq 1} \in \mathbb{N}_0^{\mathbb{N}}$, where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, we define $|\boldsymbol{\nu}| := \nu_1 + \nu_2 + \cdots$, and we refer to $\boldsymbol{\nu}$ as a "multi-index" and $|\boldsymbol{\nu}|$ as the "length" of $\boldsymbol{\nu}$. By

$$\mathfrak{F} := \{\boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}} \ : \ |\boldsymbol{\nu}| < \infty\}$$

we denote the (countable) set of all "finitely supported" multi-indices (i.e., sequences of nonnegative integers for which only finitely many entries are nonzero). For $\boldsymbol{\nu} \in \mathfrak{F}$ we denote by

$$\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u := \frac{\partial^{|\boldsymbol{\nu}|}}{\partial_{y_1}^{\nu_1} \partial_{y_2}^{\nu_2} \cdots} u$$

the partial derivative of order $\boldsymbol{\nu} \in \mathfrak{F}$ of u with respect to $\boldsymbol{y}$. For a sequence $\boldsymbol{b} = (b_j)_{j \geq 1}$ of real numbers and for $\boldsymbol{\nu} \in \mathfrak{F}$ we write $\boldsymbol{b}^{\boldsymbol{\nu}} = \prod_{j \geq 1} b_j^{\nu_j}$ and $\boldsymbol{\nu}! = \prod_{j \geq 1} \nu_j!$, with the convention that $0^0 = 1$ and $0! = 1$.

THEOREM 4.2. *Under Assumptions* (1.3) *and* (1.4)*, for every* $f \in V^*$*, every* $\boldsymbol{y} \in U$ *and every* $\boldsymbol{\nu} \in \mathfrak{F}$*, the solution* $u(\cdot, \boldsymbol{y})$ *of the parametric weak problem* (3.7) *satisfies*

$$\left\| \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y}) \right\|_V \leq |\boldsymbol{\nu}|! \, \boldsymbol{b}^{\boldsymbol{\nu}} \, \frac{\|f\|_{V^*}}{a_{\min}} \, , \tag{4.2}$$

where the sequence $\boldsymbol{b} = (b_j)_{j \geq 1} \in \ell^1(\mathbb{N})$ *is defined by*

$$b_j := \frac{\|\psi_j\|_{L^\infty(D)}}{a_{\min}} \, , \qquad j \geq 1 \, . \tag{4.3}$$

Proof. For every $v \in V$, $\boldsymbol{y} \in U$ and $\boldsymbol{\nu} \in \mathfrak{F}$ with $|\boldsymbol{\nu}| \neq 0$ we find from (3.7) the recurrence

$$\int_D a(\boldsymbol{x}, \boldsymbol{y}) \, \nabla(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\boldsymbol{x}, \boldsymbol{y})) \cdot \nabla v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$$
$$+ \sum_{j \in \mathrm{supp}(\boldsymbol{\nu})} \nu_j \int_D \psi_j(\boldsymbol{x}) \, \nabla(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu} - \boldsymbol{e}_j} u(\boldsymbol{x}, \boldsymbol{y})) \cdot \nabla v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \ = \ 0 \, , \tag{4.4}$$

where $\boldsymbol{e}_j \in \mathfrak{F}$ denotes the multi-index with entry 1 in position j and zeros elsewhere, and where $\mathrm{supp}(\boldsymbol{\nu}) = \{j \in \mathbb{N} : \nu_j \neq 0\}$ denotes the "support" of $\boldsymbol{\nu}$. We now select in (4.4) the function $v(\boldsymbol{x}) = \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\boldsymbol{x}, \boldsymbol{y}) \in V$. This yields with (1.4) the bound for $|\boldsymbol{\nu}| \neq 0$

$$a_{\min} \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_V^2 \ \leq \ \int_D a(\boldsymbol{x}, \boldsymbol{y}) \, |\nabla(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\boldsymbol{x}, \boldsymbol{y}))|^2 \, \mathrm{d}\boldsymbol{x}$$
$$= - \sum_{j \in \mathrm{supp}(\boldsymbol{\nu})} \nu_j \int_D \psi_j(\boldsymbol{x}) \, \nabla(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu} - \boldsymbol{e}_j} u(\boldsymbol{x}, \boldsymbol{y})) \cdot \nabla(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\boldsymbol{x}, \boldsymbol{y})) \, \mathrm{d}\boldsymbol{x}$$
$$\leq \sum_{j \in \mathrm{supp}(\boldsymbol{\nu})} \nu_j \int_D |\psi_j(\boldsymbol{x})| \, |\nabla(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu} - \boldsymbol{e}_j} u(\boldsymbol{x}, \boldsymbol{y}))| \, |\nabla(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\boldsymbol{x}, \boldsymbol{y}))| \, \mathrm{d}\boldsymbol{x}$$
$$\leq \sum_{j \in \mathrm{supp}(\boldsymbol{\nu})} \nu_j \, \|\psi_j\|_{L^\infty(D)} \, \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu} - \boldsymbol{e}_j} u(\cdot, \boldsymbol{y})\|_V \, \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y}))\|_V \, .$$

Here we used the convention that the absolute value of a vector function denotes its Euclidean norm. The last inequality implies, with b_j as in (4.3), for $|\boldsymbol{\nu}| \neq 0$

$$\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_V \ \leq \ \sum_{j \in \mathrm{supp}(\boldsymbol{\nu})} \nu_j \, b_j \, \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu} - \boldsymbol{e}_j} u(\cdot, \boldsymbol{y})\|_V \, . \tag{4.5}$$

The proof of the bound (4.2) on $\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_V$ will be obtained from (4.5) by induction with respect to $|\boldsymbol{\nu}|$. If $|\boldsymbol{\nu}| = 0$, the assertion is the statement of Theorem 3.1. Assume next that the assertion has been proved for all $\boldsymbol{\nu}' \in \mathfrak{F}$ such that $|\boldsymbol{\nu}'| = n - 1 \geq 0$. Let $\boldsymbol{\nu} \in \mathfrak{F}$ be such that $|\boldsymbol{\nu}| = n \in \mathbb{N}$. Then for a given $j \in \operatorname{supp}(\boldsymbol{\nu})$ we have $\boldsymbol{\nu} - \boldsymbol{e}_j \in \mathfrak{F}$ and $\boldsymbol{\nu}! = \nu_j (\boldsymbol{\nu} - \boldsymbol{e}_j)!$, $|\boldsymbol{\nu} - \boldsymbol{e}_j| = n - 1 = |\boldsymbol{\nu}| - 1 \geq 0$. From (4.5), we obtain with (4.3) the estimate

$$
\begin{aligned}
\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_V &\leq \frac{\|f\|_{V^*}}{a_{\min}} \sum_{j \in \operatorname{supp}(\boldsymbol{\nu})} \nu_j \, b_j \, (|\boldsymbol{\nu}| - 1)! \, \boldsymbol{b}^{\boldsymbol{\nu} - \boldsymbol{e}_j} \\
&= \frac{\|f\|_{V^*}}{a_{\min}} (|\boldsymbol{\nu}| - 1)! \, \boldsymbol{b}^{\boldsymbol{\nu}} \sum_{j \in \operatorname{supp}(\boldsymbol{\nu})} \nu_j = \frac{\|f\|_{V^*}}{a_{\min}} |\boldsymbol{\nu}|! \, \boldsymbol{b}^{\boldsymbol{\nu}} \,,
\end{aligned}
$$

which is (4.2).   □

5. Dimension Truncation. Although in theory the parametric diffusion coefficient (1.2) includes a sum with infinitely many terms, in practice this sum must be truncated. To reduce the truncation error, we assume that the functions ψ_j are ordered so that $\|\psi_j\|_{L^\infty(D)}$ is non-increasing, see (1.9).

Given $s \in \mathbb{N}$ and $\boldsymbol{y} \in U$, we observe that truncating the sum in (1.2) at s terms is the same as setting $y_j = 0$ for $j > s$. We denote by $u(\boldsymbol{x}, (\boldsymbol{y}_{\{1:s\}}; 0))$ the solution of the parametric weak problem (3.7) corresponding to the parametric diffusion coefficient $a(\boldsymbol{x}, (\boldsymbol{y}_{\{1:s\}}; 0))$ in which the sum is truncated at s terms.

THEOREM 5.1. *Under Assumptions* (1.3) *and* (1.4), *for every* $f \in V^*$, *every* $\boldsymbol{y} \in U$ *and every* $s \in \mathbb{N}$, *the solution* $u(\cdot, (\boldsymbol{y}_{\{1:s\}}; 0))$ *of the truncated parametric weak problem* (3.7) *satisfies*

$$
\|u(\cdot, \boldsymbol{y}) - u(\cdot, (\boldsymbol{y}_{\{1:s\}}; 0))\|_V \leq \frac{\|f\|_{V^*}}{2 \, a_{\min}^2} \sum_{j \geq s+1} \|\psi_j\|_{L^\infty(D)} \,.
$$

In addition, if Assumptions (1.6) *and* (1.9) *hold, and when* (1.6) *holds with* $p = 1$ *assume additionally that*

$$
\exists c > 0 \,, \ \eta > 0 \ : \ \|\psi_j\|_{L^\infty(D)} \leq c \, j^{-(1+\eta)} \quad \forall j \geq 1 \,, \tag{5.1}
$$

then

$$
\sum_{j \geq s+1} \|\psi_j\|_{L^\infty(D)} \leq
\begin{cases}
\frac{1}{1/p-1} \left(\sum_{j \geq 1} \|\psi_j\|_{L^\infty(D)}^p \right)^{1/p} s^{-(1/p-1)} & \text{if } p < 1 \,, \\
\frac{c}{\eta} s^{-\eta} & \text{if } p = 1 \,.
\end{cases}
$$

Proof. To simplify the notation in the proof, we write $a^s(\cdot, \boldsymbol{y}) = a(\cdot, (\boldsymbol{y}_{\{1:s\}}; 0))$ and $u^s(\cdot, \boldsymbol{y}) = u(\cdot, (\boldsymbol{y}_{\{1:s\}}; 0))$.

Recall that $u(\cdot, \boldsymbol{y})$ and $u^s(\cdot, \boldsymbol{y})$ are the exact solutions of the variational problems:

$$
\begin{aligned}
\langle a(\cdot, \boldsymbol{y}) \nabla u(\cdot, \boldsymbol{y}), \nabla v \rangle &= \langle f, v \rangle & \forall v \in V \,, \\
\langle a^s(\cdot, \boldsymbol{y}) \nabla u^s(\cdot, \boldsymbol{y}), \nabla v \rangle &= \langle f, v \rangle & \forall v \in V \,.
\end{aligned}
$$

We have from Theorem 3.1 that the solutions exist and are unique. Subtracting the weak formulations, we get

$$
\langle a(\cdot, \boldsymbol{y}) \nabla u(\cdot, \boldsymbol{y}) - a^s(\cdot, \boldsymbol{y}) \nabla u^s(\cdot, \boldsymbol{y}), \nabla v \rangle = 0 \quad \forall v \in V \,,
$$

or equivalently,

$$\langle a(\cdot, \boldsymbol{y})\nabla(u(\cdot, \boldsymbol{y}) - u^s(\cdot, \boldsymbol{y})), \nabla v\rangle = -\langle(a(\cdot, \boldsymbol{y}) - a^s(\cdot, \boldsymbol{y}))\nabla u^s(\cdot, \boldsymbol{y}), \nabla v\rangle \qquad \forall\, v \in V .$$

We interpret this as a weak problem for the error $u(\cdot, \boldsymbol{y}) - u^s(\cdot, \boldsymbol{y})$, with the forcing term $-\nabla \cdot (a(\cdot, \boldsymbol{y}) - a^s(\cdot, \boldsymbol{y}))\nabla u^s(\cdot, \boldsymbol{y}) \in V^*$. This weak solution is unique. Choosing here $v = u(\cdot, \boldsymbol{y}) - u^s(\cdot, \boldsymbol{y}) \in V$, we get with the Cauchy Schwarz inequality

$$a_{\min}\|u(\cdot, \boldsymbol{y}) - u^s(\cdot, \boldsymbol{y})\|_V^2 \leq \|a(\cdot, \boldsymbol{y}) - a^s(\cdot, \boldsymbol{y})\|_{L^\infty(D)}\|u^s(\cdot, \boldsymbol{y})\|_V\|u(\cdot, \boldsymbol{y}) - u^s(\cdot, \boldsymbol{y})\|_V .$$

Canceling one factor $\|u(\cdot, \boldsymbol{y}) - u^s(\cdot, \boldsymbol{y})\|_V$ on both sides, and using the a-priori estimate for $\|u^s(\cdot, \boldsymbol{y})\|_V = \|u(\cdot, (\boldsymbol{y}_{\{1:s\}}; 0))\|_V$ from Theorem 3.1, we find

$$\|u(\cdot, \boldsymbol{y}) - u^s(\cdot, \boldsymbol{y})\|_V \leq \frac{\|f\|_{V^*}}{a_{\min}^2}\|a(\cdot, \boldsymbol{y}) - a^s(\cdot, \boldsymbol{y})\|_{L^\infty(D)} \leq \frac{\|f\|_{V^*}}{2\,a_{\min}^2} \sum_{j \geq s+1} \|\psi_j\|_{L^\infty(D)} .$$

Next, we estimate the tail sum $\sum_{j \geq s+1} \|\psi_j\|_{L^\infty(D)}$. Consider first the case where (1.6) holds for $p < 1$. Then, since $\|\psi_j\|_{L^\infty(D)}$ is non-increasing, we have

$$\|\psi_k\|_{L^\infty(D)}^p \leq \frac{1}{k}\sum_{j=1}^{k} \|\psi_j\|_{L^\infty(D)}^p \leq \frac{1}{k}\sum_{j \geq 1} \|\psi_j\|_{L^\infty(D)}^p ,$$

which yields

$$\sum_{k \geq s+1} \|\psi_k\|_{L^\infty(D)} \leq \left(\sum_{k \geq s+1} k^{-1/p}\right)\left(\sum_{j \geq 1} \|\psi_j\|_{L^\infty(D)}^p\right)^{1/p}$$

$$\leq \frac{1}{1/p - 1}\left(\sum_{j \geq 1} \|\psi_j\|_{L^\infty(D)}^p\right)^{1/p} s^{-(1/p-1)} .$$

If (1.6) holds only for $p = 1$, then we use the additional assumption (5.1) to obtain

$$\sum_{j \geq s+1} \|\psi_j\|_{L^\infty(D)} \leq c \sum_{j \geq s+1} j^{-(1+\eta)} \leq \frac{c}{\eta} s^{-\eta} .$$

This completes the proof. □

6. Quasi-Monte Carlo Integration for Exact Solution of PDE. In this section we at last begin the application of QMC quadrature (2.7) to the infinite dimensional integral (2.6), where the integrand $F(\boldsymbol{y}) = G(u(\cdot, \boldsymbol{y}))$ is a linear functional $G(\cdot)$ of the exact solution $u(\cdot, \boldsymbol{y})$ of the parametric weak problem (3.7). In order to apply the theory developed in §2.2, recall that we need $F \in \mathcal{W}_{\boldsymbol{\gamma}}$, where the norm in $\mathcal{W}_{\boldsymbol{\gamma}}$ is defined by (2.8). In turn this will require that u belongs to the Bochner space $\mathcal{W}_{\boldsymbol{\gamma}}(U; V)$ with the norm defined by (3.4).

A crucial question before we can apply the QMC quadrature is how to choose the weights $\gamma_{\mathfrak{u}}$. This matters not only for the theory, but also because the CBC construction of the randomly shifted lattice rule requires the weights as input. The choice of weights is a delicate question in the present infinite dimensional application: the weights firstly need to be chosen in such a way as to make the infinite sum in the

Bochner space norm of u converge, so that u truly belongs to that space. But there is also a question of choosing weights that give a small worst case error, see (2.3).

In choosing the weights we will be guided by a philosophy apparently first introduced in [21], and subsequently followed in several other papers [10, 35], that one should try to choose weights that minimize not the worst case error, but rather the product of worst case error and norm of F in the error bound (2.4). In our case the norm of F is further bounded by the right hand side of (3.6), thus the strategy we follow in determining the weights in the following theorem is that, for given s and N, we minimize a certain upper bound on the right hand side of

$$|I_s(G(u)) - Q_{s,N}(G(u))| \leq e^{\mathrm{wor}}(Q_{s,N}; \mathcal{W}_{s,\boldsymbol{\gamma}}) \|G(\cdot)\|_{V^*} \|u\|_{\mathcal{W}_{\boldsymbol{\gamma}}(U;V)}.$$

For this minimization argument we need the following simple lemma.

LEMMA 6.1. *Let $n \in \mathbb{N}$, $\lambda > 0$, and $\alpha_i, \beta_i > 0$ for all i. Then the function*

$$g(\gamma_1, \ldots, \gamma_n) = \left(\sum_{i=1}^{n} \gamma_i^{\lambda} \alpha_i \right)^{1/\lambda} \left(\sum_{i=1}^{n} \frac{\beta_i}{\gamma_i} \right)$$

is minimized by taking

$$\gamma_i = c \left(\frac{\beta_i}{\alpha_i} \right)^{1/(1+\lambda)} \qquad for\ any \quad c > 0 .$$

Proof. Differentiating with respect to γ_k gives

$$\frac{\partial}{\partial \gamma_k} g(\gamma_1, \ldots, \gamma_n) = \frac{1}{\lambda} \left(\sum_{i=1}^{n} \gamma_i^{\lambda} \alpha_i \right)^{1/\lambda - 1} \lambda \gamma_k^{\lambda - 1} \alpha_k \left(\sum_{i=1}^{n} \frac{\beta_i}{\gamma_i} \right) - \left(\sum_{i=1}^{n} \gamma_i^{\lambda} \alpha_i \right)^{1/\lambda} \frac{\beta_k}{\gamma_k^2} ,$$

which, upon equating to zero, yields

$$\gamma_k^{1+\lambda} = \frac{\sum_{i=1}^{n} \gamma_i^{\lambda} \alpha_i}{\sum_{i=1}^{n} \beta_i/\gamma_i} \cdot \frac{\beta_k}{\alpha_k} .$$

Note that the scaling of the numbers γ_i is arbitrary in this minimization argument, since $g(c\,\gamma_1, \ldots, c\,\gamma_n) = g(\gamma_1, \ldots, \gamma_n)$ for any $c > 0$. Hence $g(\gamma_1, \ldots, \gamma_n)$ is minimized by choosing $\gamma_i = c\,(\beta_i/\alpha_i)^{1/(1+\lambda)}$ for any $c > 0$. □

We shall also use repeatedly two elementary estimates given below. In the first estimate we use the multi-index notation introduced in §4.2.

LEMMA 6.2. *For all $\alpha_j > 0$ with $\sum_{j \geq 1} \alpha_j < 1$ we have*

$$\sum_{|\mathfrak{u}| < \infty} |\mathfrak{u}|! \prod_{j \in \mathfrak{u}} \alpha_j \leq \sum_{\boldsymbol{\nu} \in \mathfrak{F}} \frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!} \prod_{j \geq 1} \alpha_j^{\nu_j} = \sum_{k=0}^{\infty} \left(\sum_{j \geq 1} \alpha_j \right)^k = \frac{1}{1 - \sum_{j \geq 1} \alpha_j} ,$$

and for all $\beta_j > 0$ with $\sum_{j \geq 1} \beta_j < \infty$ we have

$$\sum_{|\mathfrak{u}| < \infty} \prod_{j \in \mathfrak{u}} \beta_j = \prod_{j \geq 1} (1 + \beta_j) = \exp \left(\sum_{j \geq 1} \log(1 + \beta_j) \right) \leq \exp \left(\sum_{j \geq 1} \beta_j \right) .$$

Proof. The first estimate is obtained by replacing the sum over sets $\mathfrak{u}$ with a sum over multi-indices $\boldsymbol{\nu}$, adding other more general multi-indices to the sum, and then applying the multinomial formula and the geometric series formula. The second estimate makes use of the inequality $\log(1 + x) \leq x$ for all $x > 0$. $\square$

Another issue raised by the infinite dimensional nature of the problem is to choose the value of s and estimate the truncation error $I(G(u)) - I_s(G(u))$. We shall defer the treatment of this issue till §8, although the main part of the required analysis is already discussed in §5.

In the following theorem, Assumption (6.1) is equivalent to Assumption (1.6). Assumption (6.2) is equivalent to the condition that $\sum_{j\geq 1} \frac{1}{2}\|\psi_j\|_{L^\infty(D)} < a_{\min}$, which for the case $p = 1$ puts an additional restriction on the fluctuation of the random coefficients.

THEOREM 6.3. *Under Assumptions* (1.3) *and* (1.4), *and with* $\boldsymbol{b} = (b_j)_{j\geq 1}$ *denoting the sequence defined in* (4.3), *suppose that*

$$\sum_{j\geq 1} b_j^p < \infty \qquad for\ some \quad 0 < p \leq 1 , \tag{6.1}$$

and when $p = 1$ *assume additionally that*

$$\sum_{j\geq 1} b_j < 2 . \tag{6.2}$$

For every $f \in V^*$ *and every* $G(\cdot) \in V^*$, *let* u *denote the the solution of the parametric weak problem* (3.7). *Then for* $s \in \mathbb{N}$, N *a prime number, and weights* $\boldsymbol{\gamma} = (\gamma_\mathfrak{u})$, *a randomly shifted lattice rule with* N *points in* s *dimensions can be constructed by a component-by-component algorithm such that the root-mean-square error for approximating the finite dimensional integral* $I_s(G(u))$ *satisfies, for all* $\lambda \in (1/2, 1]$,

$$\sqrt{\mathbb{E}\left[|I_s(G(u)) - Q_{s,N}(G(u); \cdot)|^2\right]} \leq C_{\boldsymbol{\gamma}}(\lambda)\, N^{-1/(2\lambda)}\, \|f\|_{V^*}\, \|G(\cdot)\|_{V^*} , \tag{6.3}$$

where $\mathbb{E}[\cdot]$ *denotes the expectation with respect to the random shift which is uniformly distributed over* $[0, 1]^s$, *and*

$$C_{\boldsymbol{\gamma}}(\lambda) := \frac{2^{1/(2\lambda)}}{a_{\min}} \left(\sum_{|\mathfrak{u}|<\infty} \gamma_\mathfrak{u}^\lambda \left[\rho(\lambda)\right]^{|\mathfrak{u}|} \right)^{1/(2\lambda)} \left(\sum_{|\mathfrak{u}|<\infty} \frac{(|\mathfrak{u}|!)^2 \prod_{j\in\mathfrak{u}} b_j^2}{\gamma_\mathfrak{u}} \right)^{1/2} ,$$

with

$$\rho(\lambda) := \frac{2\zeta(2\lambda)}{(2\pi^2)^\lambda} + \frac{1}{12^\lambda} ,$$

but $C_{\boldsymbol{\gamma}}(\lambda)$ *is possibly infinite.*

Let

$$\lambda := \begin{cases} \frac{1}{2-2\delta} & for\ some \quad \delta \in (0, 1/2) & when\ p \in (0, 2/3] , \\ \frac{p}{2-p} & & when\ p \in (2/3, 1) , \\ 1 & & when\ p = 1 . \end{cases} \tag{6.4}$$

Then the choice of weights

$$\gamma_\mathfrak{u} = \gamma_\mathfrak{u}^* := \left(|\mathfrak{u}|! \prod_{j\in\mathfrak{u}} \frac{b_j}{\sqrt{\rho(\lambda)}} \right)^{2/(1+\lambda)} \tag{6.5}$$

minimizes $C_{\boldsymbol{\gamma}}(\lambda)$, *and leads to*

$$C_{\boldsymbol{\gamma}}(\lambda) \;<\; \infty \qquad and \qquad \|u\|_{\mathcal{W}_{\boldsymbol{\gamma}}(U;V)} \;<\; \infty \,. \tag{6.6}$$

In particular, $C_{\boldsymbol{\gamma}}(\frac{1}{2-2\delta}) \to \infty$ *as* $\delta \to 0$, *and* $C_{\boldsymbol{\gamma}}(\frac{p}{2-p}) \to \infty$ *as* $p \to (2/3)^{+}$. *The error is of order*

$$\begin{cases} N^{-(1-\delta)} & when \; p \in (0, 2/3] \,, \\ N^{-(1/p-1/2)} & when \; p \in (2/3, 1) \,, \\ N^{-1/2} & when \; p = 1 \,. \end{cases}$$

If, instead of (6.5), *we define the weights by*

$$\gamma_{\mathfrak{u}} \;:=\; \left(|\mathfrak{u}|! \prod_{j\in\mathfrak{u}} (2\,b_j) \right)^{2-p} , \tag{6.7}$$

then $C_{\boldsymbol{\gamma}}(\lambda)$ *is no longer minimized, but* (6.6) *still holds provided that* $\delta < p/2$ *when* $p \in (0, 2/3]$. *Moreover, the rate of convergence remains the same.*

  Proof. It follows from Theorem 4.2 and the definition of the $\mathcal{W}_{\boldsymbol{\gamma}}(U;V)$ norm (3.4) that

$$\|u\|_{\mathcal{W}_{\boldsymbol{\gamma}}(U;V)} \;\leq\; \frac{\|f\|_{V^*}}{a_{\min}} \left(\sum_{|\mathfrak{u}|<\infty} \frac{(|\mathfrak{u}|!)^2 \prod_{j\in\mathfrak{u}} b_j^2}{\gamma_{\mathfrak{u}}} \right)^{1/2} .$$

This together with Theorem 2.2 and (3.6) yield the error bound (6.3). We stress that at this point $\|u\|_{\mathcal{W}_{\boldsymbol{\gamma}}(U;V)}$ and/or $C_{\boldsymbol{\gamma}}(\lambda)$ may or may not be finite.

  The aim of this proof is to choose the weights $\gamma_{\mathfrak{u}}$ such that $C_{\boldsymbol{\gamma}}(\lambda)$ is finite (and hence so is $\|u\|_{\mathcal{W}_{\boldsymbol{\gamma}}(U;V)}$), and that $C_{\boldsymbol{\gamma}}(\lambda)$ is as small as possible. In the course of our derivation below we shall choose the value of λ according to the value of p, but until then λ and p are independent.

  Clearly $C_{\boldsymbol{\gamma}}(\lambda)$ will be bounded if and only if

$$g_{\lambda,t}(\boldsymbol{\gamma}) \;:=\; \left(\sum_{|\mathfrak{u}|\leq t} \gamma_{\mathfrak{u}}^{\lambda} \, [\rho(\lambda)]^{|\mathfrak{u}|} \right)^{1/\lambda} \left(\sum_{|\mathfrak{u}|\leq t} \frac{(|\mathfrak{u}|!)^2 \prod_{j\in\mathfrak{u}} b_j^2}{\gamma_{\mathfrak{u}}} \right)$$

is bounded independently of t. From Lemma 6.1 we learn that $g_{\lambda,t}(\boldsymbol{\gamma})$ is minimized by choosing $\gamma_{\mathfrak{u}}$ as in (6.5) for $|\mathfrak{u}| \leq t$. This scaling of weights is consistent with the convention that $\gamma_{\emptyset} = 1$. Now since this argument is valid for arbitrarily large but finite t, we may choose $\gamma_{\mathfrak{u}}$ as in (6.5) for all $|\mathfrak{u}| < \infty$.

  Next we demonstrate that $C_{\boldsymbol{\gamma}}(\lambda) < \infty$ for the weights given by (6.5). We have

$$\sum_{|\mathfrak{u}|<\infty} (\gamma_{\mathfrak{u}}^*)^{\lambda} \, [\rho(\lambda)]^{|\mathfrak{u}|} \;=\; \sum_{|\mathfrak{u}|<\infty} \frac{(|\mathfrak{u}|!)^2 \prod_{j\in\mathfrak{u}} b_j^2}{\gamma_{\mathfrak{u}}^*}$$

$$=\; \sum_{|\mathfrak{u}|<\infty} (|\mathfrak{u}|!)^{2\lambda/(1+\lambda)} \prod_{j\in\mathfrak{u}} \left(b_j^{2\lambda} \rho(\lambda) \right)^{1/(1+\lambda)} \;=:\; A_{\lambda} \,,$$

and thus $C_{\boldsymbol{\gamma}^*}(\lambda) = 2^{1/(2\lambda)} A_{\lambda}^{1/(2\lambda)+1/2}/a_{\min}$.

For $\lambda \in (1/2, 1)$, we have $2\lambda/(1 + \lambda) < 1$ and we further estimate A_λ as follows: we multiply and divide by $\prod_{j \in \mathfrak{u}} \alpha_j^{2\lambda/(1+\lambda)}$, with $\alpha_j > 0$ to be specified later, and we apply Hölder's inequality with conjugate exponents $(1 + \lambda)/(2\lambda)$ and $(1 + \lambda)/(1 - \lambda)$, to obtain

$$
A_\lambda = \sum_{|\mathfrak{u}|<\infty} (|\mathfrak{u}|!)^{2\lambda/(1+\lambda)} \prod_{j \in \mathfrak{u}} \alpha_j^{2\lambda/(1+\lambda)} \prod_{j \in \mathfrak{u}} \left(\frac{b_j^{2\lambda} \rho(\lambda)}{\alpha_j^{2\lambda}} \right)^{1/(1+\lambda)}
$$

$$
\leq \left(\sum_{|\mathfrak{u}|<\infty} |\mathfrak{u}|! \prod_{j \in \mathfrak{u}} \alpha_j \right)^{2\lambda/(1+\lambda)} \left(\sum_{|\mathfrak{u}|<\infty} \prod_{j \in \mathfrak{u}} \left(\frac{b_j^{2\lambda} \rho(\lambda)}{\alpha_j^{2\lambda}} \right)^{1/(1-\lambda)} \right)^{(1-\lambda)/(1+\lambda)}
$$

$$
\leq \left(\frac{1}{1 - \sum_{j \geq 1} \alpha_j} \right)^{2\lambda/(1+\lambda)} \exp \left(\frac{1 - \lambda}{1 + \lambda} [\rho(\lambda)]^{1/(1-\lambda)} \sum_{j \geq 1} \left(\frac{b_j}{\alpha_j} \right)^{2\lambda/(1-\lambda)} \right)
$$

which holds and A_λ is finite, see Lemma 6.2, provided that

$$
\sum_{j \geq 1} \alpha_j < 1 \qquad \text{and} \qquad \sum_{j \geq 1} \left(\frac{b_j}{\alpha_j} \right)^{2\lambda/(1-\lambda)} < \infty . \tag{6.8}
$$

We now choose

$$
\alpha_j := \frac{b_j^p}{\theta} \qquad \text{for some} \quad \theta > \sum_{j \geq 1} b_j^p . \tag{6.9}
$$

Then the first sum in (6.8) is less than 1 due to Assumption (6.1). Noting that (6.1) implies that $\sum_{j \geq 1} b_j^{p'} < \infty$ for all $p' \geq p$, we conclude that the second sum in (6.8) converges for

$$
\frac{2\lambda}{1 - \lambda}(1 - p) \geq p \qquad \Longleftrightarrow \qquad p \leq \frac{2\lambda}{1 + \lambda} \qquad \Longleftrightarrow \qquad \lambda \geq \frac{p}{2 - p} .
$$

Since λ must be strictly between $1/2$ and 1 for the argument above, when $p \in (0, 2/3]$ we choose $\lambda = 1/(2 - 2\delta)$ for some $\delta \in (0, 1/2)$, and when $p \in (2/3, 1)$ we set $\lambda = p/(2 - p)$.

For the case $p = 1$ we take $\lambda = 1$, and we have $\rho(1) = 1/4$. Then using Lemma 6.2 and Assumption (6.2) we obtain

$$
A_1 = \sum_{|\mathfrak{u}|<\infty} |\mathfrak{u}|! \prod_{j \in \mathfrak{u}} \left(\frac{b_j}{2} \right) \leq \frac{1}{1 - \sum_{j \geq 1} (b_j/2)} < \infty .
$$

Finally we show that $C_{\boldsymbol{\gamma}}(\lambda) < \infty$ for the weights given by (6.7). For the case $p = 1$ and $\lambda = 1$, the weights (6.5) and (6.7) are the same, so we need to consider only the cases $p \in (0, 2/3]$ and $p \in (2/3, 1)$. To simplify the presentation below we introduce $p' := \lambda(2 - p)$. Then, with λ given by (6.4), with the additional restriction that $\delta < p/2$, it is easy to verify that $p' = p$ for $p \in (2/3, 1)$ and $p < p' < 1$ for $p \in (0, 2/3]$. In both cases we have

$$
\left(\frac{a_{\min} C_{\boldsymbol{\gamma}}(\lambda)}{2^{1/(2\lambda)}} \right)^2 = \left(\sum_{|\mathfrak{u}|<\infty} (|\mathfrak{u}|!)^{p'} \prod_{j \in \mathfrak{u}} \left((2b_j)^{p'} \rho(\lambda) \right) \right)^{\frac{2-p}{p'}} \left(\sum_{|\mathfrak{u}|<\infty} (|\mathfrak{u}|!)^p \prod_{j \in \mathfrak{u}} \frac{b_j^p}{2^{2-p}} \right) .
$$

For the first sum, we multiply and divide by $\prod_{j \in \mathfrak{u}} \alpha_j^{p'}$, with $\alpha_j > 0$ to be specified later, and we apply Hölder's inequality with conjugate exponents $1/p'$ and $1/(1-p')$. For the second sum, we multiply and divide by $\prod_{j \in \mathfrak{u}} \alpha_j^p$, with the same α_j, and we apply Hölder's inequality with conjugate exponents $1/p$ and $1/(1-p)$. We obtain

$$
\left(\frac{a_{\min} C_\gamma(\lambda)}{2^{1/(2\lambda)}} \right)^2 \le \left(\sum_{|\mathfrak{u}| < \infty} |\mathfrak{u}|! \prod_{j \in \mathfrak{u}} \alpha_j \right)^{2-p} \left(\sum_{|\mathfrak{u}| < \infty} \prod_{j \in \mathfrak{u}} \left(\frac{(2b_j)^{p'} \rho(\lambda)}{\alpha_j^{p'}} \right)^{\frac{1}{1-p'}} \right)^{\frac{(1-p')(2-p)}{p'}}
$$
$$
\times \left(\sum_{|\mathfrak{u}| < \infty} |\mathfrak{u}|! \prod_{j \in \mathfrak{u}} \alpha_j \right)^{p} \left(\sum_{|\mathfrak{u}| < \infty} \prod_{j \in \mathfrak{u}} \left(\frac{b_j^p}{2^{2-p} \alpha_j^p} \right)^{\frac{1}{1-p}} \right)^{1-p} ,
$$

which is finite as above provided that

$$
\sum_{j \ge 1} \alpha_j < 1 , \qquad \sum_{j \ge 1} \left(\frac{b_j}{\alpha_j} \right)^{\frac{p'}{1-p'}} < \infty , \qquad \text{and} \qquad \sum_{j \ge 1} \left(\frac{b_j}{\alpha_j} \right)^{\frac{p}{1-p}} < \infty ,
$$

and this can be achieved by choosing α_j as in (6.9), since $(1-p)p'/(1-p') \ge p$. This completes the proof. $\square$

A small motivation for considering the weights (6.7) instead of the optimal weights (6.5) is that there is no parameter δ to be specified in (6.7) for the case $p \in (0, 2/3]$.

A similar error analysis (but without minimizing the error bound) can also be performed for other types of QMC integration rules. If we use product weights with randomly shifted lattice rules (or digitally shifted polynomial lattice rules, see e.g. [8]), we can also obtain the $\mathcal{O}(N^{-1+\delta})$ convergence rate but now we need the stronger condition that $p \le 1/2$, rather than $p \le 2/3$. For the lattice rules of [16, 28] we get $\mathcal{O}(N^{-1+\delta})$ convergence, but again with $p \le 1/2$. Niederreiter and Sobol′ sequences give $\mathcal{O}(N^{-1+\delta})$ convergence, see [34], but they require an even stronger condition that $p \le 1/3$. Our best results are therefore for randomly shifted lattice rules with POD weights (6.5): they provide the convergence rate $\mathcal{O}(N^{-1+\delta})$ under the weakest summability requirements on $\|\psi_j\|_{L^\infty(D)}$; specifically, when the summability exponent p is between $2/3$ and 1, the QMC rates correspond exactly to the bounds for (nonlinear) best N-term approximation results obtained in [5].

7. Finite Element Discretization. In what follows, we impose the additional assumption

$$
D \subset \mathbb{R}^d \text{ is a convex and bounded polyhedron with plane faces .} \tag{7.1}
$$

This assumption is made to simplify the exposition of the FE method, and could be substantially relaxed. For example, on domains D with curved boundaries, standard results on FE analysis as in e.g. [3] will imply corresponding results.

Let us denote by $\{V_h\}_h$ a one-parameter family of subspaces $V_h \subset V$ of dimensions $M_h < \infty$. We think of the spaces V_h as spaces of continuous, piecewise linear finite elements on a sequence of regular, simplicial meshes in D obtained from an initial, regular triangulation of D by recursive, uniform bisection of simplices. Then it is well known that for functions $v \in Z$ there holds $v \in V \cap H^2(D)$, and that there exists a constant $C > 0$ such that, as $h \to 0$,

$$
\inf_{v_h \in V_h} \|v - v_h\|_V \le C h \|v\|_Z , \tag{7.2}
$$

where the norm $\|\cdot\|_Z$ is defined in (3.2).

For any $\boldsymbol{y} \in U$, we define the *parametric FE approximation* $u_h(\cdot, \boldsymbol{y})$ as the FE solution of the parametric deterministic problem: for every $f \in V^*$ and every $\boldsymbol{y} \in U$, find

$$u_h(\cdot, \boldsymbol{y}) \in V_h : \quad b(\boldsymbol{y}; u_h(\cdot, \boldsymbol{y}), v_h) = \langle f, v_h \rangle \qquad \forall v_h \in V_h . \tag{7.3}$$

Here, $b(\boldsymbol{y}; \cdot, \cdot)$ denotes the parametric bilinear form (3.8). In particular the FE approximation (7.3) is defined *pointwise* with respect to the parameter $\boldsymbol{y} \in U$, so that the application of a QMC rule to the FE approximation is well defined.

THEOREM 7.1. *Under Assumptions* (1.3), (1.4), (1.8), *and* (7.1), *for every* $f \in V^*$ *and every* $\boldsymbol{y} \in U$, *the FE approximations* $u_h(\cdot, \boldsymbol{y})$ *are stable in the sense that*

$$\|u_h(\cdot, \boldsymbol{y})\|_V \leq \frac{\|f\|_{V^*}}{a_{\min}} . \tag{7.4}$$

Moreover, for every $f \in L^2(D)$, *as* $h \to 0$, *there holds the asymptotic convergence estimate*

$$\|u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y})\|_V \leq C\, h\, \|f\|_{L^2(D)} \simeq C\, M_h^{-1/d}\, \|f\|_{L^2(D)} , \tag{7.5}$$

where the constant $C > 0$ *is independent of* h.

Proof. The estimate (7.4) follows from Theorem 3.1 on noting the conformity $V_h \subset V$. Assumption (1.4) implies that the FE approximations are quasioptimal uniformly with respect to the parameter vector $\boldsymbol{y} \in U$: from (3.7) and (7.3) we find

$$\|u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y})\|_V \leq \frac{a_{\max}}{a_{\min}} \inf_{v_h \in V_h} \|u(\cdot, \boldsymbol{y}) - v_h\|_V .$$

Next we assume $f \in L^2(D)$, and apply the approximation property (7.2) and the regularity estimate (4.1) to obtain

$$\|u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y})\|_V \leq C\, h\, \|u(\cdot, \boldsymbol{y})\|_Z \leq C\, h\, \|f\|_{L^2(D)} ,$$

which proves (7.5).   □

Since we are interested in estimating the error in approximating functionals (3.5), we will also impose a regularity assumption on the functional $G(\cdot)$:

$$G(\cdot) \in L^2(D) .$$

Moreover, since in the expression (3.5) only a bounded linear functional $G(\cdot)$ of u rather than u itself enters, the discretization error of $G(u)$ is of interest as well. It is well known that $|G(u(\cdot, \boldsymbol{y})) - G(u_h(\cdot, \boldsymbol{y}))|$ can converge faster than $\|u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y})\|_V$.

THEOREM 7.2. *Under Assumptions* (1.3), (1.4), (1.8), *and* (7.1), *for every* $f \in L^2(D)$, *every* $G(\cdot) \in L^2(D)$, *and every* $\boldsymbol{y} \in U$, *as* $h \to 0$, *the FE approximations* $G(u_h(\cdot, \boldsymbol{y}))$ *satisfy the asymptotic convergence estimate*

$$|G(u(\cdot, \boldsymbol{y})) - G(u_h(\cdot, \boldsymbol{y}))| \leq C\, h^2\, \|f\|_{L^2(D)}\, \|G(\cdot)\|_{L^2(D)}$$
$$\simeq C\, M_h^{-2/d}\, \|f\|_{L^2(D)}\, \|G(\cdot)\|_{L^2(D)} , \tag{7.6}$$

where the constant $C > 0$ *is independent of* $\boldsymbol{y} \in U$.

Proof. The error bound (7.6) follows from a Aubin-Nitsche duality argument together with the regularity assumption $G(\cdot) \in L^2(D)$ and the a-priori estimate (4.1)

on u: for $G(\cdot) \in L^2(D)$ and any $\boldsymbol{y} \in U$, we define $v_G(\cdot, \boldsymbol{y}) \in V$ as the unique solution of the adjoint problem

$$v_G(\cdot, \boldsymbol{y}) \in V : \quad b(\boldsymbol{y}; w, v_G(\cdot, \boldsymbol{y})) = G(w) \qquad \forall w \in V . \tag{7.7}$$

Problem (7.7) admits a unique solution and, on account of the symmetry $b(\boldsymbol{y}; w, v) = b(\boldsymbol{y}; v, w)$ for all $v, w \in V$, we also have

$$b(\boldsymbol{y}; v_G(\cdot, \boldsymbol{y}), w) = G(w) \qquad \forall w \in V .$$

As this problem is of exactly the same type as the original parametric problem (3.7), we have for the representer $v_G(\cdot, \boldsymbol{y})$ analogous regularity assertions as for $u(\cdot, \boldsymbol{y})$. In particular, there exists a constant $C > 0$ which is independent of $\boldsymbol{y}$ such that

$$\|v_G(\cdot, \boldsymbol{y})\|_Z \leq C \|G(\cdot)\|_{L^2(D)} . \tag{7.8}$$

Using (7.7) and the Galerkin orthogonality of the FE discretization, we may write, for every $\boldsymbol{y} \in U$ and every $v_h \in V_h$,

$$
\begin{aligned}
|G(u(\cdot, \boldsymbol{y})) - G(u_h(\cdot, \boldsymbol{y}))| &= |G(u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y}))| \\
&= |b(\boldsymbol{y}; u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y}), v_G(\cdot, \boldsymbol{y}))| \\
&= |b(\boldsymbol{y}; u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y}), v_G(\cdot, \boldsymbol{y}) - v_h)| \\
&\leq C \|u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y})\|_V \|v_G(\cdot, \boldsymbol{y}) - v_h\|_V .
\end{aligned}
$$

Finally we apply (7.2), (7.4), and (7.8) to obtain

$$|G(u(\cdot, \boldsymbol{y})) - G(u_h(\cdot, \boldsymbol{y}))| \leq C h^2 \|f\|_{L^2(D)} \|v_G(\cdot, \boldsymbol{y})\|_Z \leq C h^2 \|f\|_{L^2(D)} \|G(\cdot)\|_{L^2(D)} ,$$

which completes the proof. $\square$

8. Combined Quasi-Monte Carlo Finite Element Error Analysis. We now present the error analysis of the combined QME FE approximation for the integral (3.5). To obtain a computable approximation of (3.5), we approximate the infinite dimensional integral using a randomly shifted lattice rule with N points in s dimensions. A realization for a draw of the shift $\boldsymbol{\Delta}$ will be denoted by $Q_{s,N}(\cdot; \boldsymbol{\Delta})$. Moreover, for each evaluation of the integrand, we replace the exact solution $u(\cdot, \boldsymbol{y})$ of the parametric weak problem (3.7) by its FE approximation $u_h(\cdot, \boldsymbol{y}) \in V_h \subset V$ from a FE space V_h of dimension $M_h < \infty$.

Thus we may express the overall error as a sum of a *dimension truncation error* (which is implicit when a finite dimensional QMC method is used for an infinite dimensional integral), a *QMC quadrature error*, and a *FE discretization error*:

$$
\begin{aligned}
&I(G(u)) - Q_{s,N}(G(u_h); \boldsymbol{\Delta}) \\
&= (I - I_s)(G(u)) + (I_s(G(u)) - Q_{s,N}(G(u); \boldsymbol{\Delta})) + Q_{s,N}(G(u - u_h); \boldsymbol{\Delta}) .
\end{aligned}
$$

The mean-square error with respect to the random shift can then be bounded by

$$
\begin{aligned}
\mathbb{E}\left[|I(G(u)) - Q_{s,N}(G(u_h); \cdot)|^2\right] \leq{}& 3\,|(I - I_s)(G(u))|^2 \\
&+ 3\,\mathbb{E}\left[|I_s(G(u)) - Q_{s,N}(G(u); \cdot)|^2\right] \\
&+ 3\,\mathbb{E}\left[|Q_{s,N}(G(u - u_h); \cdot)|^2\right] . \tag{8.1}
\end{aligned}
$$

The QMC error, i.e., the second term in (8.1), is already analyzed in Theorem 6.3. For the truncation error, i.e., the first term in (8.1), we use the estimate

$$
\begin{aligned}
|(I - I_s)(G(u))| &= \left| \int_U G(u(\cdot, \boldsymbol{y}) - u(\cdot, (\boldsymbol{y}_{\{1:s\}}; 0))) \, \mathrm{d}\boldsymbol{y} \right| \\
&\leq \sup_{\boldsymbol{y} \in U} |G(u(\cdot, \boldsymbol{y}) - u(\cdot, (\boldsymbol{y}_{\{1:s\}}; 0)))| \\
&\leq \|G(\cdot)\|_{V^*} \sup_{\boldsymbol{y} \in U} \|u(\cdot, \boldsymbol{y}) - u(\cdot, (\boldsymbol{y}_{\{1:s\}}; 0)\|_V \ ,
\end{aligned}
$$

and then apply Theorem 5.1. Finally, for the FE error, i.e., the third term in (8.1), we apply the property that the QMC quadrature weights $1/N$ are positive and sum to 1, to obtain

$$
\mathbb{E}\left[|Q_{s,N}(G(u - u_h); \cdot)|^2\right] \leq \sup_{\boldsymbol{y} \in U} |G(u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y}))|^2 \ ,
$$

and then apply Theorem 7.2. The combined error estimate is summarized in Theorem 8.1 below.

THEOREM 8.1. *Under the same assumptions and definitions as in Theorems 5.1, 6.3 and 7.2, if we approximate the integral over U by the randomly shifted lattice rule from Theorem 6.3 with N points in s dimensions, and for each shifted lattice point we solve the approximate elliptic problem* (7.3) *by one common FE discretization in the domain D with M_h degrees of freedom with linear cost $\mathcal{O}(M_h)$ (e.g. by Multigrid Methods), then we have the root-mean-square error bound*

$$
\begin{aligned}
&\sqrt{\mathbb{E}\left[|I(G(u)) - Q_{s,N}(G(u_h); \cdot)|^2\right]} \\
&\leq C\left(\kappa(s, N) \|f\|_{V^*} \|G(\cdot)\|_{V^*} + M_h^{-2/d} \|f\|_{L^2(D)} \|G(\cdot)\|_{L^2(D)}\right)
\end{aligned}
$$

where

$$
\kappa(s, N) = \begin{cases}
s^{-(1/p-1)} + N^{-(1-\delta)} & \text{when } p \in (0, 2/3) \ , \\
s^{-1/2} + N^{-(1-\delta)} & \text{when } p = 2/3 \ , \\
s^{-(1/p-1)} + N^{-(1/p-1/2)} & \text{when } p \in (2/3, 1) \ , \\
s^{-\eta} + N^{-1/2} & \text{when } p = 1 \ .
\end{cases}
$$

If we choose s, N, and M_h to balance the terms so that the root-mean-square error is of order $\mathcal{O}(\varepsilon)$ for some $\varepsilon > 0$, then the overall cost of this QMC-FE approximation is of order

$$
\mathcal{O}(s \, N \, M_h) = \begin{cases}
\mathcal{O}\left(\varepsilon^{-(p/(1-p) + 1/(1-\delta) + d/2)}\right) & \text{when } p \in (0, 2/3) \ , \\
\mathcal{O}\left(\varepsilon^{-(2 + 1/(1-\delta) + d/2)}\right) & \text{when } p = 2/3 \ , \\
\mathcal{O}\left(\varepsilon^{-(p/(1-p) + 2p/(2-p) + d/2)}\right) & \text{when } p \in (2/3, 1) \ , \\
\mathcal{O}\left(\varepsilon^{-(1/\eta + 2 + d/2)}\right) & \text{when } p = 1 \ .
\end{cases}
$$

Note that we have assumed the strong regularity assumptions of Theorem 7.2, namely, $f \in L^2(D)$, $G(\cdot) \in L^2(D)$, and Assumption (1.8) holds. If instead of $G(\cdot) \in L^2(D)$ we have only that $G(\cdot) \in V^*$, then we can use (7.5) from Theorem 7.1 to obtain a FE convergence rate of order $\mathcal{O}(M_h^{-1/d})$ instead of $\mathcal{O}(M_h^{-2/d})$.

9. Concluding Remarks. In this paper we have focused our attention on obtaining a good convergence rate with respect to the number of QMC points N, and have put aside the issue of reducing the overall computational cost. The cost can be dramatically reduced if we use instead multi-level and/or changing dimension algorithms which have been analyzed in a number of recent papers, see e.g. [20, 23, 15, 12, 26] for infinite dimensional integration, and e.g. [4, 1, 2] for applications in PDEs. This is the theme of our subsequent paper [19], which requires a non-trivial generalization of the results in this paper. It also includes an extension to higher order moments (k-point correlation functions) of the solution of PDE.

We have shown in Theorem 4.2 that the solution of the PDE (1.1) is very smooth with respect to the parametric variable $\boldsymbol{y}$, but in this paper we have only considered QMC methods for integrands with mixed first derivatives. An obvious extension of the present work is to consider instead higher order QMC methods, see e.g. [8, Chapter 15], and this will be analyzed in our future work.

We considered the parametric, second order elliptic equation (1.1) with inhomogeneous, *isotropic* diffusion coefficient $a(\boldsymbol{x}, \boldsymbol{y})$. This was done for ease of notation and of exposition only; completely analogous results can be developed for *anisotropic* diffusion tensors $\{A_{ij}(\boldsymbol{x}, \boldsymbol{y}) : i, j = 1, ..., d\}$ with affine dependence on $\boldsymbol{y}$ in place of $a(\boldsymbol{x}, \boldsymbol{y})$, as well as for more general elliptic and parabolic partial differential operators. Details will be presented in [18].

The random coefficient $a(\boldsymbol{x}, \boldsymbol{y})$ in this paper is assumed to depend in an affine manner on the parametric variable $\boldsymbol{y}$ from the bounded parameter domain $U = (-\frac{1}{2}, \frac{1}{2})^{\mathbb{N}}$ with uniform distribution. This is often referred to as the "uniform case". It is of high practical interest to consider instead the "log-normal case" where it is assumed that the logarithm of $a(\boldsymbol{x}, \boldsymbol{y})$ depends on $\boldsymbol{y}$ which belongs to the unbounded domain $\mathbb{R}^{\mathbb{N}}$ with normal distribution. This is the topic of [13], which requires a combination of innovative QMC and FE analyses. A successful computational study without theoretical analysis on the application of QMC methods to the log-normal case was already reported in the recent paper [14].

REFERENCES

[1] A. BARTH, CH. SCHWAB, AND N. ZOLLINGER, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, to appear in Numer. Math.

[2] J. CHARRIER, R. SCHEICHL, AND A. L. TECKENTRUP, *Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods*, Bath Institute For Complex Systems, Preprint 2/11 (2011).

[3] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, Elsevier, Amsterdam 1978.

[4] K. A. CLIFFE, M. B. GILES, R. SCHEICHL, AND A. L. TECKENTRUP, *Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients*, to appear in Computing and Visualization in Science.

[5] A. COHEN, R. DE VORE AND CH. SCHWAB, *Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs*, Found. Comp. Math., 10 (2010), pp. 615–646.

[6] R. COOLS, F. Y. KUO, AND D. NUYENS, *Constructing embedded lattice rules for multivariate integration*, SIAM J. Sci. Comput., 28 (2006), pp. 2162–2188.

[7] J. DICK, *On the convergence rate of the component-by-component construction of good lattice rules*, J. Complexity, 20 (2004), pp. 493–522.

[8] J. DICK AND F. PILLICHSHAMMER, *Digital Nets and Sequences*, Cambridge University Press, 2010.

[9] J. DICK, F. PILLICHSHAMMER, AND B. J. WATERHOUSE, *The construction of good extensible rank-1 lattices*, Math. Comp., 77 (2008), pp. 2345–2374.

[10] J. DICK, I. H. SLOAN, X. WANG, AND H. WOŹNIAKOWSKI, *Liberating the weights*, J. Complexity, 20 (2004), pp. 593–623.

[11] D. GILBARG AND N. S. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, New York, 2nd Ed., 2001.

[12] M. GNEWUCH, *Infinite-dimensional integration on weighted Hilbert spaces*, Math. Comp., to appear.

[13] I. G. GRAHAM, F. Y. KUO, J. NICHOLS, R. SCHEICHL, CH. SCHWAB, AND I. H. SLOAN, *QMC-FE methods for PDEs with log-normal random coefficients* (in preparation).

[14] I. G. GRAHAM, F. Y. KUO, D. NUYENS, R. SCHEICHL, AND I. H. SLOAN, *Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications*, J. Comput. Phy., 230 (2011), pp. 3668–3694.

[15] F.J. HICKERNELL, T. MÜLLER-GRONBACH, B. NIU, AND K. RITTER, *Deterministic multi-level algorithms for infinite-dimensional integration on $\mathbb{R}^{\mathbb{N}}$*, J. Complexity, 27 (2011), pp. 331–351.

[16] S. JOE, *Construction of good rank-1 lattice rules based on the weighted star discrepancy*, in Monte Carlo and Quasi-Monte Carlo Methods 2004 (H. Niederreiter and D. Talay, eds.), Springer Verlag, pp. 181–196, 2006.

[17] F. Y. KUO, *Component-by-component constructions achieve the optimal rate of convergence*, J. Complexity, 19 (2003), pp 301–320.

[18] F. Y. KUO, CH. SCHWAB, AND I. H. SLOAN, *Quasi-Monte Carlo methods for very high dimensional integration — the standard weighted-space setting and beyond* (in preparation).

[19] ———, *Multi-level QMC-FE methods for PDEs with random coefficients* (in preparation).

[20] F. Y. KUO, I. H. SLOAN, G. W. WASILKOWSKI, AND H. WOŹNIAKOWSKI, *Liberating the dimension*, J. Complexity, 26 (2010), pp. 422–454.

[21] G. LARCHER, G. LEOBACHER, AND K. SCHEICHER, *On the tractability of the Brownian bridge algorithm*, J. Complexity, 19 (2004), pp. 511–528.

[22] H. NIEDERREITER, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.

[23] B. NIU, F.J. HICKERNELL, T. MÜLLER-GRONBACH, AND K. RITTER, *Multi-level Monte Carlo algorithms for infinite-dimensional integration on $\mathbb{R}^{\mathbb{N}}$*, J. Complexity, 26 (2010), pp. 229–254.

[24] D. NUYENS AND R. COOLS, *Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces*, Math. Comp., 75 (2006), pp. 903–920.

[25] ———, *Fast component-by-component construction of rank-1 lattice rules with a non-prime number of points*, J. Complexity, 22 (2006), pp. 4–28.

[26] L. PLASKOTA AND G. W. WASILKOWSKI, *Tractability of infinite-dimensional integration in the worst case and randomized settings*, J. Complexity, appeared online.

[27] CH. SCHWAB AND R. A. TODOR, *Karhunen-Loève approximation of random fields by generalized fast multipole methods*, J. Comput. Phy., 217 (2006), pp. 100–122.

[28] V. SINESCU AND S. JOE, *Good lattice rules based on the general weighted star discrepancy*, Math. Comp., 76 (2007), pp. 989–1004.

[29] I. H. SLOAN AND S. JOE, *Lattice Methods for Multiple Integration*, Oxford University Press, Oxford, 1994.

[30] I. H. SLOAN, F. Y. KUO, AND S. JOE, *Constructing randomly shifted lattice rules in weighted Sobolev spaces*, SIAM J. Numer. Anal., 40 (2002), pp. 1650–1665.

[31] I. H. SLOAN AND H. WOŹNIAKOWSKI, *When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals?*, J. Complexity, 14 (1998), pp. 1–33.

[32] I. H. Sloan, X. Wang, and H. Woźniakowski, *Finite-order weights imply tractability of multivariate integration*, J. Complexity, 20 (2004), pp. 46–74.

[33] R. A. TODOR AND CH. SCHWAB, *Convergence rates of sparse chaos approximations of elliptic problems with stochastic coefficients*, IMA J. Numer. Anal., 27 (2007), pp. 232–261.

[34] X. WANG, *Strong tractability of multivariate integration using quasi-Monte Carlo algorithms*, Math. Comp., 72 (2002), pp. 823–838.

[35] X. WANG AND I. H. SLOAN, *Brownian bridge and principal component analysis: towards removing the curse of dimensionality*, IMA J. Numer. Anal., 27 (2007), pp. 631–654.

Research Reports