QMC Galerkin discretization of parametric operator equations

Eidgenössische Technische Hochschule Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

QMC Galerkin discretization of parametric operator equations

Ch. Schwab

Research Report No. 2012-18
July 2012

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

QMC Galerkin Discretization of Parametric Operator Equations

Christoph Schwab, SAM, ETH Zürich

Abstract
We extend recent results from [KSS1] of QMC quadrature and Finite Element discretization for parametric, scalar second order elliptic partial differential equations to general QMC-Galerkin discretizations of parametric operator equations, which depend on possibly countably many parameters. Such problems typically arise in the numerical solution of differential and integral equations with random field inputs. The present setting covers general second order elliptic equations which are possibly indefinite (Helmholtz equation), or which are given in saddle point variational form (such as mixed formulations). The also cover nonsymmetric variational formulations which appear in space-time Galerkin discretizations of parabolic problems or countably parametric nonlinear initial value problems [HaSc11].
Research supported in part by the European Research Council (ERC) under Grant AdG 247277

1 Introduction

The efficient numerical computation of statistical quantities for solutions of partial differential and of integral equations with random inputs is a key task in uncertainty quantification in engineering and in the sciences. The quantity of interest being expressed as a mathematical expectation, the efficient computation of these quantities involves two basic steps: i) approximate (numerical) solution of the operator equation, and ii) numerical integration. In the present note, we outline a general strategy towards these two aims which is based on i) Galerkin discretization and ii) Quasi Monte-Carlo (QMC) integration, following [KSS1].

Contrary to Monte-Carlo methods which require uniformly distributed samples of random input functions, QMC (and other) quadrature methods require the *introduction of coordinates of integration* prior to numerical quadrature. In the context of

Seminar für Angewandte Mathematik, ETH Züriche-mail: schwab@math.ethz.ch

random field inputs with nondegenerate covariance operators, a *countable number of coordinates* is required to describe the random input data, e.g. by a Karhunen-Loève expansion. Therefore, in the present note, we consider in particular that the operator equation contains not only a finite number of random input parameters, but rather depends on *random field inputs*, i.e. it contains random functions of space and, in evolution problems, of time which describe uncertainty in the problem under consideration. We survey and generalize recent results on the convergence of QMC Finite Element discretizations of linear, scalar, second order elliptic partial differential equations in divergence form obtained recently in [KSS1, GKN3S].

2 Parametric operator equations

We generalize the results of [CDS1] and study well-posedness, regularity and polynomial approximation of solutions for a family of abstract parametric saddle point problems. Particular attention is paid to the case of *countably many parameters*. The abstract results in the present section cover a wide range of operator equations: among them are (stationary and time-dependent) diffusion in random media [CDS1], wave propagation [HoSc12], optimal control problems for uncertain systems [KS11] and of interest.

2.1 Abstract Parametric Saddle Point Problems

Throughout, we denote by $\mathscr{X}$ and $\mathscr{Y}$ two reflexive Banach spaces over $\mathbb{R}$ (all results will hold with the obvious modifications also for spaces over $\mathbb{C}$) with (topological) duals $\mathscr{X}'$ and $\mathscr{Y}'$, respectively. By $\mathscr{L}(\mathscr{X},\mathscr{Y}')$, we denote the set of bounded linear operators $A : \mathscr{X} \to \mathscr{Y}'$. The Riesz representation theorem associates each $A \in \mathscr{L}(\mathscr{X},\mathscr{Y}')$ in a one-to-one correspondence a bilinear form $b(\cdot,\cdot) : \mathscr{X} \times \mathscr{Y} \to \mathbb{R}$ by means of

$$b(v,w) = \langle w, Av\rangle_{\mathscr{Y}\times\mathscr{Y}'} \quad \text{for all } v \in \mathscr{X}, w \in \mathscr{Y} \,. \tag{1}$$

Here and in what follows, we indicate spaces in duality pairings $\langle\cdot,\cdot\rangle$ by subscripts.

We shall be interested in the solution of linear operator equations $Au = f$ and make use of the following solvability result which is a straightforward consequence of the closed graph theorem, see, e.g., [BF].

Proposition 1 *A bounded, linear operator $A \in \mathscr{L}(\mathscr{X},\mathscr{Y}')$ is boundedly invertible* if and only if *its bilinear form satisfies* inf-sup conditions*: ex. $\gamma > 0$ s.t.*

$$\inf_{0\neq v\in\mathscr{X}} \sup_{0\neq w\in\mathscr{Y}} \frac{\mathbf{b}(v,w)}{\|v\|_{\mathscr{X}}\|w\|_{\mathscr{Y}}} \geq \gamma\,, \quad \inf_{0\neq w\in\mathscr{Y}} \sup_{0\neq v\in\mathscr{X}} \frac{\mathbf{b}(v,w)}{\|v\|_{\mathscr{X}}\|w\|_{\mathscr{Y}}} \geq \gamma\,. \tag{2}$$

If (2) *holds then for every $f \in \mathscr{Y}'$ the operator equation*

$$\textit{find } u \in \mathscr{X} : \quad \mathbf{b}(u,v) = \langle f, v\rangle_{\mathscr{Y}'\times\mathscr{Y}} \quad \forall v \in \mathscr{Y} \tag{3}$$

admits a unique solution $u \in \mathscr{X}$ *and there holds* $\|u\|_{\mathscr{X}} = \|A^{-1}f\|_{\mathscr{X}} \leq \gamma^{-1}\|f\|_{\mathscr{Y}'}$.

2.2 Parametric operator families

We shall be interested in QMC quadratures applied to solutions of *parametric families of operators A*. From partial differential equations with random field input (see, e.g. [ScGJGActa11]), we consider, in particular, operator families which depend on infinitely many parameters (obtained, for example, by Karhunen-Loève expansion of random input functions). To this end, we denote by $y := (y_j)_{j\geq 1} \in \mathscr{U}$ the possibly (for random field inpus with nondegenerate covariance kernels) countable set of parameters. We assume the parameters to take values in $\mathscr{U} \subseteq \mathbb{R}^{\mathbb{N}}$. In particular, each realization of y is a sequence of real numbers. Two main cases arise in practice: first, the *"uniform case"*: the parameter domain $\mathscr{U} = [-1/2,1/2]^{\mathbb{N}}$ and, second, the *"lognormal case"*: the parameter domain $\mathscr{U} \subset \mathbb{R}^{\mathbb{N}}$. In both cases, we account for randomness in inputs by equipping these parameter domains with countable product probability measures (thereby stipulating *mathematical independence* of the random coordinates y_j). Specifically, in the uniform case

$$\rho(dy) = \bigotimes_{j\geq 1} \frac{dy_j}{2}\,, \quad y \in [-1/2,1/2]^{\mathbb{N}}\,. \tag{4}$$

and in the lognormal case with the Gaussian measure

$$\gamma(dy) = \bigotimes_{j\geq 1} \gamma_1(dy_j)\,, \quad y \in \mathbb{R}^{\mathbb{N}}\,. \tag{5}$$

By $\mathbb{N}_0^{\mathbb{N}}$ we denote the set of all sequences of nonnegative integers, and by $\mathfrak{F} = \{\nu \in \mathbb{N}_0^{\mathbb{N}} : |\nu| < \infty\}$ the set of "finitely supported" such sequences, i.e., sequences of nonnegative integers which have only a finite number of nonzero entries. For $\nu \in \mathfrak{F}$, we denote by $\mathfrak{n} \subset \mathbb{N}$ the set of coordinates j such that $\nu_j \neq 0$, *with* j *repeated* $\nu_j \geq 1$ *many times*. Analogously, $\mathfrak{m} \subset \mathbb{N}$ denotes the supporting coordinate set for $\mu \in \mathfrak{F}$.

We consider *parametric* families of continuous, linear operators which we denote as $A(y) \in \mathscr{L}(\mathscr{X},\mathscr{Y}')$. We now make precise the dependence of $A(y)$ on the parameter sequence y which is required for our regularity and approximation results.

Assumption 1 *The parametric operator family* $\{A(y) \in \mathscr{L}(\mathscr{X},\mathscr{Y}') : y \in \mathscr{U}\}$ *is a* regular p-analytic operator family *for some* $0 < p \leq 1$*, i.e.,*

(i) $A(y) \in \mathscr{L}(\mathscr{X},\mathscr{Y}')$ *is boundedly invertible for every* $y \in \mathscr{U}$ *with uniformly bounded inverses* $A(y)^{-1} \in \mathscr{L}(\mathscr{Y}',\mathscr{X})$*, i.e., there exists* $C_0 > 0$ *such that*

$$\sup_{y\in\mathscr{U}} \|A(y)^{-1}\|_{\mathscr{L}(\mathscr{Y}',\mathscr{X})} \leq C_0 \tag{6}$$

and

(ii)for any fixed $y \in [-1/2,1/2]^{\mathbb{N}}$, the operators $A(y)$ are analytic with respect to each y_j such that there exists a nonnegative sequence $b = (b_j)_{j\geq 1} \in \ell^p(\mathbb{N})$ such that

$$\forall \nu \in \mathfrak{F}\backslash\{0\}: \quad \sup_{y\in\mathscr{U}} \left\| (A(0))^{-1}(\partial_y^\nu A(y)) \right\|_{\mathscr{L}(\mathscr{X},\mathscr{X})} \leq C_0 b^\nu \,. \tag{7}$$

Here $\partial_y^\nu A(y) := \partial_{y_1}^{\nu_1}\partial_{y_2}^{\nu_2}\cdots A(y)$; the notation b^ν signifies the (finite due to $\nu\in\mathfrak{F}$) product $b_1^{\nu_1}b_2^{\nu_2}\ldots$ where we use the convention $0^0 := 1$.

We verify the abstract assumptions in the particular setting of *affine parameter dependence*; this case arises, for example, in diffusion problems where the diffusion coefficients are given in terms of a Karhunen-Loève expansion (see, e.g. [ST06] for such Karhunen-Loève expansions and their numerical analysis, in the context of elliptic PDEs with random coefficients). Then, there exists a family $\{A_j\}_{j\geq 0} \subset \mathscr{L}(\mathscr{X},\mathscr{Y}')$ such that $A(y)$ can be written in the form

$$\forall y \in \mathscr{U}: \quad A(y) = A_0 + \sum_{j\geq 1} y_j A_j \,. \tag{8}$$

We shall refer to $A_0 = A(0)$ as "nominal"operator, and to the operators A_j, $j \geq 1$ as "fluctuation" operators. In order for the sum in (8) to converge, we impose the following assumptions on the sequence $\{A_j\}_{j\geq 0} \subset \mathscr{L}(\mathscr{X},\mathscr{Y}')$. In doing so, we associate with the operator A_j the bilinear forms $\mathbf{b}_j(\cdot,\cdot) : \mathscr{X}\times\mathscr{Y} \to \mathbb{R}$ via

$$\forall v \in \mathscr{X},\ w \in \mathscr{Y}: \quad \mathbf{b}_j(v,w) = {}_{\mathscr{Y}}\langle w, A_j v\rangle_{\mathscr{Y}'} \,, \quad j = 0,1,2....$$

Assumption 2 *The family $\{A_j\}_{j\geq 0}$ in* (8) *satisfies the following conditions:*

1. *The "nominal" or "mean field" operator $A_0 \in \mathscr{L}(\mathscr{X},\mathscr{Y}')$ is boundedly invertible, i.e. (cf. Proposition 1) there exists $\gamma_0 > 0$ such that*

$$\inf_{0\neq v\in\mathscr{X}} \sup_{0\neq w\in\mathscr{Y}} \frac{\mathbf{b}_0(v,w)}{\|v\|_{\mathscr{X}}\|w\|_{\mathscr{Y}}} \geq \gamma_0 \,, \ \inf_{0\neq w\in\mathscr{Y}} \sup_{0\neq v\in\mathscr{X}} \frac{\mathbf{b}_0(v,w)}{\|v\|_{\mathscr{X}}\|w\|_{\mathscr{Y}}} \geq \gamma_0 \,. \tag{A1}$$

2. *The "fluctuation" operators $\{A_j\}_{j\geq 1}$ are small with respect to A_0 in the following sense: there exists a constant $0 < \kappa < 2$ such that for γ_0 as in* (**A1**) *holds*

$$\sum_{j\geq 1} b_j \leq \kappa < 2 \,, \quad \textit{where} \quad b_j := \|A_0^{-1}A_j\|_{\mathscr{L}(\mathscr{X},\mathscr{Y}')} \,, \quad j = 1,2,\ldots \,. \tag{A2}$$

Condition (**A2**) (and, hence, Assumption 2) is sufficient for the bounded invertibility of $A(y)$, uniformly w.r. to the parameter vector $y \in \mathscr{U} = [-1/2,1/2]^{\mathbb{N}}$.

Theorem 2. *Under Assumption 2, for every realization $y \in \mathscr{U} = [-1/2,1/2]^{\mathbb{N}}$ of the parameter vector, the parametric operator $A(y)$ is boundedly invertible. Specifically, for the bilinear form $\mathbf{b}(y;\cdot,\cdot) : \mathscr{X}\times\mathscr{Y} \to \mathbb{R}$ associated with $A(y) \in \mathscr{L}(\mathscr{X},\mathscr{Y}')$ via*

$$\mathbf{b}(y;w,v) := {}_{\mathscr{Y}}\langle v, A(y)w\rangle_{\mathscr{Y}'} \tag{9}$$

there hold the uniform (w.r. to $y \in \mathscr{U}$) inf-sup conditions (2) *with* $\gamma = (1-\kappa/2)\gamma_0$,

$$\forall y \in \mathscr{U}: \quad \inf_{0\neq v\in\mathscr{X}} \sup_{0\neq w\in\mathscr{Y}} \frac{\mathbf{b}(y;v,w)}{\|v\|_{\mathscr{X}}\|w\|_{\mathscr{Y}}} \geq \gamma\,, \ \inf_{0\neq w\in\mathscr{Y}} \sup_{0\neq v\in\mathscr{X}} \frac{\mathbf{b}(y;v,w)}{\|v\|_{\mathscr{X}}\|w\|_{\mathscr{Y}}} \geq \gamma\,. \quad (10)$$

In particular, for every $f \in \mathscr{Y}'$ and for every $y \in \mathscr{U}$, the parametric operator equation

$$\text{find } u(y) \in \mathscr{X}: \quad \mathbf{b}(y;u(y),v) = \langle f, v\rangle_{\mathscr{Y}\times\mathscr{Y}'} \quad \forall v \in \mathscr{Y} \quad (11)$$

admits a unique solution $u(y)$ which satisfies the a-priori estimate

$$\sup_{y\in\mathscr{U}} \|u(y)\|_{\mathscr{X}} \leq C\|f\|_{\mathscr{Y}'}\,. \quad (12)$$

Proof. We use Proposition 1, which gives necessary and sufficient conditions for bounded invertibility; also, $1/\gamma$ is a bound for the inverse. By Assumption 2, the nominal part A_0 of $A(y)$ in (8) is boundedly invertible, and we write for every $y \in \mathscr{U}$: $A(y) = A_0\left(I + \sum_{j\geq 1} y_j A_0^{-1} A_j\right)$. We see that $A(y)$ is boundedly invertible iff the Neumann Series in the second factor is. Since $|y_j| \leq 1/2$, a sufficient condition for this is (**A2**) which implies, with Proposition 1, the assertion with $\gamma = \gamma_0(1-\kappa/2)$. □

From the preceding considerations, the following is readily verified.

Corollary 1. *The affine parametric operator family* (8) *satisfies Assumption 1 with*

$$C_0 = \frac{1}{(1-\kappa/2)\gamma_0} \quad \text{and} \quad b_j := \|A_0^{-1}A_j\|_{\mathscr{L}(\mathscr{X},\mathscr{Y}')}\,, \quad \text{for all } j \geq 1\,.$$

2.3 Analytic parameter dependence of solutions

The dependence of the solution $u(y)$ of the parametric, variational problem (11) on the parameter vector y is analytic, with precise bounds on the growth of the partial derivatives. The following bounds of the parametric solution's dependence on the parameter vector y will, as in [KSS1], allow us to prove dimension independent rates of convergence of QMC quadratures.

Theorem 3. *Under Assumption 1, for every $f \in \mathscr{Y}'$ and for every $y \in \mathscr{U}$, the unique solution $u(y) \in \mathscr{X}$ of the parametric operator equation*

$$A(y)\,u(y) = f \qquad \text{in } \mathscr{Y}' \quad (13)$$

depends analytically on the parameters, and the partial derivatives of the parametric solution family $u(y)$ satisfy the bounds

$$\sup_{y\in\mathscr{U}} \|(\partial_y^{\nu} u)(y)\|_{\mathscr{X}} \leq C_0|\nu|!\,\tilde{b}^{\nu}\|f\|_{\mathscr{Y}'} \quad \text{for all } \nu \in \mathfrak{F}, \quad (14)$$

where $0! := 1$ *and the where sequence* $\tilde{b} = (\tilde{b}_j)_{j\geq 1} \in \ell^p(\mathbb{N})$ *is defined by* $\tilde{b}_j = b_j/\ln 2$ *for all* $j \in \mathbb{N}$*, with* b_j *as in* (**A2**).

For a proof, we refer, for example, to [CDS1, KS11].

2.4 Spatial regularity of solutions

For the convergence rate analysis of Galerkin discretizations of the parametric operator equation (13), we will require regularity of the parametric solution $u(y)$. To this end, we assume given *scales of smoothness spaces* $\{\mathscr{X}_t\}_{\geq 0}$ and $\{\mathscr{Y}'_t\}_{t\geq 0}$, with

$$\mathscr{X} = \mathscr{X}_0 \supset \mathscr{X}_1 \supset \mathscr{X}_2 \supset \dots , \qquad \mathscr{Y}' = \mathscr{Y}'_0 \supset \mathscr{Y}'_1 \supset \mathscr{Y}'_2 \supset \dots . \tag{15}$$

The scales $\{\mathscr{X}_t\}_{t\geq 0}$ and $\{\mathscr{Y}'_t\}_{t\geq 0}$ are assumed to be defined also for noninteger values of the smoothness parameter $t \geq 0$ by interpolation.

Instances of smoothness scales (15) in the context of the diffusion problem considered in [CDS1, KSS1] are, in a *convex domain* D, the choices $\mathscr{X} = H^1_0(D)$, $\mathscr{X}_1 = (H^2 \cap H^1_0)(D)$, $\mathscr{Y}' = H^{-1}(D)$, $\mathscr{Y}'_1 = L^2(D)$. In a nonconvex polygon (or polyhedron), analogous smoothness scales are available, but involve Sobolev spaces with weights.

In the ensuing convergence analysis of Galerkin discretizations of (13), we will assume that the data regularity $f \in \mathscr{Y}'_t$ for some $t > 0$ implies that

$$\forall y \in \mathscr{U} : \quad u(y) = A(y)^{-1} f \in \mathscr{X}_t \, . \tag{16}$$

Such regularity is available for a wide range of parametric differential equations (see [ScGJGActa11, HaSc11, KS11] and the references there).

2.5 Galerkin Discretization

As the inverse $A(y)^{-1}$ is not available explicitly, we will have to compute, for given QMC quadrature points $y \in \mathscr{U}$, an approximate inverse. We consider the case when it is obtained by *Galerkin discretization*: given two one-parameter families $\{\mathscr{X}^h\}_{h>0} \subset \mathscr{X}$ and $\{\mathscr{Y}^h\}_{h>0} \subset \mathscr{Y}$ of subspaces of equal, finite dimension N_h, which are dense in $\mathscr{X}$ resp. in $\mathscr{Y}$, i.e.

$$\forall u \in \mathscr{X} : \quad \limsup_{h\to 0} \inf_{0\neq u^h \in \mathscr{X}^h} \|u - u^h\|_{\mathscr{X}} = 0 \, . \tag{17}$$

We will also assume the *approximation property*:

$$\forall 0 < t \leq \bar{t} : \ \exists C_t > 0 : \ \forall u \in \mathscr{X}_t \ \forall 0 < h \leq h_0 : \quad \inf_{w^h \in \mathscr{X}^h} \|u - w^h\|_{\mathscr{X}} \leq C_t h^t \|u\|_{\mathscr{X}_t} \, . \tag{18}$$

The maximum amount of smoothness in the scale $\mathscr{X}_t$, denoted by $\bar{t}$, depends of the problem class under consideration and on the Sobolev scale: e.g. for elliptic problems in polygonal domains, it is well known that choosing for $\mathscr{X}_t$ the usual Sobolev spaces will allow (16) with t only in a rather small interval $0 < t \leq \bar{t}$, whereas choosing $\mathscr{X}_t$ as *weighted Sobolev spaces* will allow rather large values of $\bar{t}$.

Proposition 4 *Assume that the subspace sequences* $\{\mathscr{X}^h\}_{h>0} \subset \mathscr{X}$ *and* $\{\mathscr{Y}^h\}_{h>0} \subset \mathscr{Y}$ *are stable, i.e. that there exists* $\bar{\gamma} > 0$ *and* $h_0 > 0$ *such that for every* $0 < h \leq h_0$, *there hold the* uniform (w.r. to $y \in \mathscr{U}$) discrete inf-sup conditions

$$\forall y \in \mathscr{U} : \quad \inf_{0\neq v^h \in \mathscr{X}^h} \sup_{0\neq w^h \in \mathscr{Y}^h} \frac{\mathbf{b}(y; v^h, w^h)}{\|v^h\|_{\mathscr{X}} \|w^h\|_{\mathscr{Y}}} \geq \bar{\gamma} > 0 \tag{19}$$

and

$$\forall y \in \mathscr{U} : \quad \inf_{0\neq w^h \in \mathscr{Y}^h} \sup_{0\neq v^h \in \mathscr{X}^h} \frac{\mathbf{b}(y; v^h, w^h)}{\|v^h\|_{\mathscr{X}} \|w^h\|_{\mathscr{Y}}} \geq \bar{\gamma} > 0 \,. \tag{20}$$

Then, for every $0 < h \leq h_0$, *and for every* $y \in \mathscr{U}$, *the Galerkin approximations* $u^h \in \mathscr{X}^h$, *given by*

$$\textit{find } u^h(y) \in \mathscr{X}^h : \quad \mathbf{b}(y; u^h(y), v^h) = \langle f, v^h \rangle_{\mathscr{Y}\times\mathscr{Y}'} \quad \forall v^h \in \mathscr{Y}^h \tag{21}$$

admits a unique solution $u^h(y)$ *which satisfies the a-priori estimate*

$$\sup_{y\in\mathscr{U}} \|u^h(y)\|_{\mathscr{X}} \leq \bar{\gamma}^{-1} \|f\|_{\mathscr{Y}'} \,. \tag{22}$$

Morever, there exists a constant $C > 0$ *such that for all* $y \in \mathscr{U}$ *holds* quasioptimality

$$\|u(y) - u^h(y)\|_{\mathscr{X}} \leq C\bar{\gamma}^{-1} \inf_{0\neq w^h \in \mathscr{X}^h} \|u(y) - w^h\|_{\mathscr{X}} \,. \tag{23}$$

3 QMC Integration

For a given bounded, linear functional $G(\cdot) : \mathscr{X} \to \mathbb{R}$, we are interested in computing expected values of

$$F(y) := G(u(\cdot, y)) \,, \qquad y \in \mathscr{U} \,, \tag{24}$$

(respectively of its parametric Galerkin approximation $u^h(y) \in \mathscr{X}_h \subset \mathscr{X}$ defined in (21)). The expected value of F is an integral of the functional $G(\cdot)$ of the parametric solution:

$$\int_{\mathscr{U}} F(y)\,\mathrm{d}y = \int_{\mathscr{U}} G(u(\cdot,y))\,\mathrm{d}y = G\left(\int_{\mathscr{U}} u(\cdot,y)\,\mathrm{d}y\right) \,.$$

The issue is thus the numerical evaluation of Bochner integrals of $\mathscr{X}$-valued functions over the infinite dimensional domain of integration $\mathscr{U}$. We also observe that for

the parametric operator equation (13), to evaluate F at a single QMC point $y \in \mathscr{U}$ requires the approximate (Galerkin) solution of one operator equation for $u(\cdot, y) \in \mathscr{X}$. This is more expensive than mere function evaluation, and introduces an additional *Galerkin discretization error*.

Here, we follow the analysis of [KSS1], where scalar diffusion problems were considered. We prove that all results in [KSS1] have extensions to the present, more general, setting with analogous proofs.

In [CDS1] and the present paper, the summability of the fluctuation operators A_j, $j \geq 1$, plays an important role for proving dimension-independent convegence rates of approximations of the parametric solution maps. Accordingly, we will make the assumption, stronger than Assumption (**A2**) that there exists $0 < p < 1$ such that

$$\sum_{j \geq 1} \|A_j\|^p_{\mathscr{L}(\mathscr{X}, \mathscr{Y}')} < \infty \,. \tag{**A3**}$$

Notice that this condition is, by (**A1**), equivalent to $(b_j)_{j \geq 1} \in \ell^p(\mathbb{N})$, and implies decay of the fluctuation coefficients A_j, with stronger decay as the value of p becomes smaller. In both [CDS1, KSS1] and the present paper, the rate of convergence $\mathscr{O}(N^{-1+\delta})$ is attained if (**A3**) is satisfied with $p = 2/3$. Here and throughout what follows, N denotes the number of QMC points. For values of p between $2/3$ and 1, the rate of convergence in both cases is $\mathscr{O}(N^{-(1/p-1/2)})$.

Recall that the purpose of the present paper is to analyze the accuracy and complexity of QMC methods in connection with the Galerkin approximation (21) of (11). To obtain convergence rates, we strengthen Assumption (**A2**) to the requirement

$$\sup_{y \in \mathscr{U}} \|A(y)^{-1}\|_{\mathscr{L}(\mathscr{Y}_t', \mathscr{X}_t)} < \infty \,, \quad 0 \leq t \leq 1 \,. \tag{**A4**}$$

For application of QMC quadrature rules, the infinite sum in (8) must be truncated to a finite sum of, say, s terms. Below, the parameter s shall be referred to as "QMC-truncation dimension". For this truncation to make sense, we will assume additionally that the A_j are enumerated so that the bounds b_j in (**A2**) are decreasing:

$$b_1 \geq b_2 \geq \ldots \geq b_j \geq \cdots \,. \tag{**A5**}$$

The overall error for the QMC-Galerkin approximation is then a sum of three terms: a *truncation error*, a *QMC error*, and the *Galerkin discretization error*. We bound the three errors and finally combine them to arrive at an overall QMC-Galerkin error bound.

3.1 Finite dimensional setting

In this subsection we review QMC integration when the truncation dimension (ie. the number of integration variables), denoted by s, is assumed to be finite and fixed. The domain of integration is taken to be the s-dimensional unit cube $[-\frac{1}{2}, \frac{1}{2}]^s$ cen-

tered at the origin (note that this is different from the usual QMC convention where the unit cube is $[0,1]^s$). Application of existing QMC results to this setting may require a coordinate translation. We thus consider integrals of the form

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

In our later applications F will be of the form (24), 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

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

with carefully chosen points $y^{(1)},\ldots,y^{(N)} \in [-\frac{1}{2},\frac{1}{2}]^s$. For classical results on QMC methods, see, e.g. [Nie92, SJ94].

We shall assume that our integrand F belongs to a *weighted* and *anchored* Sobolev space $\mathscr{W}_{s,\gamma}$ 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. The norm is given by

$$\|F\|_{\mathscr{W}_{s,\gamma}} := \left(\sum_{\mathfrak{u}\subseteq\{1:s\}} \gamma_{\mathfrak{u}}^{-1} \int_{[-\frac{1}{2},\frac{1}{2}]^{|\mathfrak{u}|}} \left|\frac{\partial^{|\mathfrak{u}|}F}{\partial y_{\mathfrak{u}}}(y_{\mathfrak{u}};0)\right|^2 \mathrm{d}y_{\mathfrak{u}}\right)^{1/2}, \tag{25}$$

where $\{1:s\}$ is a shorthand notation for the set of indices $\{1,2,\ldots,s\}$, $\frac{\partial^{|\mathfrak{u}|}F}{\partial y_{\mathfrak{u}}}$ denotes the mixed first derivative with respect to the variables y_j with $j\in\mathfrak{u}$, and $(y_{\mathfrak{u}};0)$ denotes the vector whose jth component is y_j if $j\in\mathfrak{u}$ and 0 if $j\notin\mathfrak{u}$. We also will require $u\in\mathscr{W}_\gamma(U;\mathscr{X})$ which is defined as the Bochner space of $\mathscr{X}$-valued functions for which the norm (25) (with the $\|\circ\|_{\mathscr{X}}$ in place of the absolute value) is finite.

The weighted spaces $\mathscr{W}_{s,\gamma}$ were first introduced by Sloan and Woźniakowski in [SW98]. By now there are many variants and generalizations, see e.g. [DSWW04, SWW04]. In (25) 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 [SWW04]: there is a weight parameter $\gamma_{\mathfrak{u}}\geq 0$ associated with each group of variables $y_{\mathfrak{u}} = (y_j)_{j\in\mathfrak{u}}$ with indices belonging to the set $\mathfrak{u}$, with the convention that $\gamma_\emptyset = 1$. Following [KSS1], we consider so-called *product and order dependent ("POD" for short) weights*, which are given by

$$\gamma_{\mathfrak{u}} = \Gamma_{|\mathfrak{u}|}\prod_{j\in\mathfrak{u}}\gamma_j > 0\,. \tag{26}$$

Here $|\mathfrak{u}|$ denotes the cardinality (or the "order") of $\mathfrak{u}$. The weights are therefore determined by a specific choice of the sequences $\Gamma_0 = \Gamma_1 = 1, \Gamma_2, \Gamma_3, \ldots$ and $\gamma_1, \gamma_2, \gamma_3, \ldots$. (See (39) ahead for the precise choice of weights.)

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 $\mathscr{W}_{s,\gamma}$, i.e.,

$$e^{\mathrm{wor}}(Q_{s,N}; \mathscr{W}_{s,\gamma}) := \sup_{\|F\|_{\mathscr{W}_{s,\gamma}} \le 1} |I_s(F) - Q_{s,N}(F)| \,. \tag{27}$$

There is an explicit expression for $e^{\mathrm{wor}}(Q_{s,N}; \mathscr{W}_{s,\gamma})$ which allows it to be analyzed in theory and computed in practice. Various upper bounds for $e^{\mathrm{wor}}(Q_{s,N}; \mathscr{W}_{s,\gamma})$ 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 $e^{\mathrm{wor}}(Q_{s,N}; \mathscr{W}_{s,\gamma}) \le C N^{-r}$ with r close to 1, which is optimal in $\mathscr{W}_{s,\gamma}$, and with C independent of the dimension s, which can hold if certain conditions on the weights γ are satisfied. Note that due to linearity of the functionals $I_s(\cdot)$ and $Q_{s,N}(\cdot)$, we have

$$|I_s(F) - Q_{s,N}(F)| \le e^{\mathrm{wor}}(Q_{s,N}; \mathscr{W}_{s,\gamma}) \, \|F\|_{\mathscr{W}_{s,\gamma}} \qquad \text{for all} \quad F \in \mathscr{W}_{s,\gamma}. \tag{28}$$

We focus on a family of QMC rules known as "shifted rank-1 lattice rules". These are QMC rules with quadrature points given by

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

where $z \in \mathbb{Z}^s$ is known as the *generating vector*, $\Delta \in [0,1]^s$ is the *shift*, and $\mathrm{frac}(\cdot)$ means to take the fractional part of each component in the vector. The subtraction by the vector $(\frac{1}{2}, \ldots, \frac{1}{2})$ takes care of the translation from $[0,1]^s$ to $[-\frac{1}{2}, \frac{1}{2}]^s$ considered here. We consider randomly shifted lattice rules. In this case, the quality of the rules is determined by the choice of the generating vector z and we denote the corresponding shifted lattice rule with shift Δ by $Q_{s,N}(\Delta; F)$.

Theorem 5. *([KSS2, Theorem 5]) Let $s, N \in \mathbb{N}$ be given, and assume $F \in \mathscr{W}_{s,\gamma}$ for a particular choice of weights γ. 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 $\lambda \in (1/2, 1]$,*

$$\begin{aligned} &\sqrt{\mathbb{E}\left[|I_s(F) - Q_{s,N}(\cdot; F)|^2\right]} \\ &\le \left(\sum_{\emptyset \ne \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}} + \frac{1}{12^{\lambda}} \right)^{|\mathfrak{u}|} \right)^{1/(2\lambda)} [\varphi(N)]^{-1/(2\lambda)} \, \|F\|_{\mathscr{W}_{s,\gamma}} \,, \end{aligned} \tag{29}$$

where $\mathbb{E}[\cdot]$ denotes the expectation with respect to the random shift which is uniformly distributed over $[0,1]^s$. Here, $\zeta(x)$ denotes the Riemann zeta function, and $\varphi(N)$ the Euler totient function.

The result with general weights but restricted to prime N first obtained in [SWW04, Theorem 3(A)]. 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 [SKJ02]. The generating vector z is chosen one component at a time, while minimizing a "shift-averaged" worst case error expression. The convergence result is proved in [K03, D04]. The fast implementation using FFT is due to R. Cools etal. [NC06a]. The error bound remains valid when the truncation dimension s is replaced by any smaller number, and it also means that the corresponding lattice rule can be extended to higher dimensions at any time. Algorithms for obtaining lattice rules that are extensible in N were given in Cools, Kuo and Nuyens [CKN06]. There are analogous results for "digitally-shifted polynomial lattice rules", which is another family of QMC rules, and QMC rules with higher order convergence, see [DiPi10] and the references there. 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 $\mathcal{O}(sN\ln N)$ operations for product weights or order-dependent weights, see [CKN06]. Extension of the fast CBC construction from order-dependent weights to POD weights (26) is straightforward; this extension is given in [KSS2]. For these cases, extensible lattice sequences can be constructed at a cost of $\mathcal{O}(sN(\ln N)^2)$ operations, see [CKN06].

3.2 Infinite dimensional setting

QMC integration in (countably) infinitely many variables was considered in [KSWW10b]. Here, as in [KSS1], we proceed analogously, but with the anchor at the center of the unit cube rather than at a corner of it. For F a function depending on countably many variables $y = (y_1, y_2, \ldots)$, the integral of interest takes the form

$$I(F) := \lim_{s\to\infty} I_s(F)\,, \qquad I_s(F) := \int_{[-\frac{1}{2},\frac{1}{2}]^s} F(y_1,\ldots,y_s,0,0,\ldots)\,\mathrm{d}y_1\cdots\mathrm{d}y_s\,, \quad (30)$$

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

$$Q_{s,N}(F) := \frac{1}{N}\sum_{i=1}^{N} F(y_1^{(i)},\ldots,y_s^{(i)},0,0,\ldots)\,. \quad (31)$$

We assume that F belongs to the weighted Sobolev space $\mathscr{W}_\gamma$, which is the infinite dimensional version of $\mathscr{W}_{s,\gamma}$, with the norm (25) replaced by

$$\|F\|_{\mathscr{W}_\gamma} := \left(\sum_{|\mathfrak{u}|<\infty} \gamma_{\mathfrak{u}}^{-1} \int_{[-\frac{1}{2},\frac{1}{2}]^{|\mathfrak{u}|}} \left|\frac{\partial^{|\mathfrak{u}|}F}{\partial y_{\mathfrak{u}}}(y_{\mathfrak{u}};0)\right|^2 \mathrm{d}y_{\mathfrak{u}}\right)^{1/2}, \quad (32)$$

where the sum is now over all subsets $\mathfrak{u} \subset \mathbb{N}$ with finite cardinality. This definition is consistent with (25) in the sense that, for a function that depends only on the first s variables, its norm in $\mathscr{W}_\gamma$ is the same as its norm in $\mathscr{W}_{s,\gamma}$. Moreover, for a function F that depends on infinitely many variables, if we define $F_s(y_1,\dots,y_s) := F(y_1,\dots,y_s,0,0,\cdots)$ by anchoring the components beyond dimension s at 0, then $\|F_s\|_{\mathscr{W}_{s,\gamma}} = \|F_s\|_{\mathscr{W}_\gamma} \le \|F\|_{\mathscr{W}_\gamma}$.

Although in theory the parametric operator (8) includes a sum with infinitely many terms, for the application of QMC qudrature rules this sum must be truncated. To reduce the resulting *dimension truncation error*, we recall that we assumed the integration coordinates y_j to be ordered so that b_j in (**A2**), is non-increasing, see Assumption (**A5**).

Given $s \in \mathbb{N}$ and $y \in U$, we observe that truncating the sum in (8) at s terms amounts to setting $y_j = 0$ for $j > s$. We thus denote by $u^s(x,y) := u(x,(y_{\{1:s\}};0))$ the solution of the parametric weak problem (11) corresponding to the parametric operator $A((y_{\{1:s\}};0))$ in which the sum (8) is truncated at s terms. Then Theorem 2 holds when $u(\cdot,y)$ is replaced by $u^s(\cdot,y)$.

Theorem 6. *Under Assumptions* (**A2**)*, for every $f \in \mathscr{Y}'$, for every $y \in \mathscr{X}'$ and for every $s \in \mathbb{N}$, the solution $u^s(\cdot,y) = u(\cdot,(y_{\{1:s\}};0))$ of the s-term truncated parametric weak problem* (11) *satisfies, with b_j as defined in* (**A2**)*,*

$$\|u(\cdot,y) - u^s(\cdot,y)\|_{\mathscr{X}} \le C\gamma^{-1}\|f\|_{\mathscr{Y}'} \sum_{j\ge s+1} b_j \tag{33}$$

for some constant $C > 0$ independent of f. Moreover, for every $G(\cdot) \in \mathscr{X}'$, we have

$$|I(G(u)) - I_s(G(u))| \le \tilde{C}\gamma^{-1}\|f\|_{\mathscr{Y}'}\|G(\cdot)\|_{\mathscr{X}'}\left(\sum_{j\ge s+1} b_j\right)^2 \tag{34}$$

for some constant $\tilde{C} > 0$ independent of f and $G(\cdot)$. In addition, if Assumptions (**A3**) *and* (**A5**) *hold, then*

$$\sum_{j\ge s+1} b_j \le \min\left(\frac{1}{1/p-1},1\right)\left(\sum_{j\ge 1} b_j^p\right)^{1/p} s^{-(1/p-1)}\,.$$

The proof is the same as that of Theorem 5.1 in [KSS1].

4 Quasi-Monte Carlo Integration for Exact Solution of PDE

We apply QMC quadrature (31) to the infinite dimensional integral (30), where the integrand $F(y) = G(u(\cdot,y))$ is a continuous, linear functional $G(\cdot)$ of the solution $u(\cdot,y)$ of the parametric weak problem (11). In order to apply the theory developed in Section 3.2, recall that we need $F \in \mathscr{W}_\gamma$, where the norm in $\mathscr{W}_\gamma$ is defined by (32).

To determine good choices of QMC weights, we minimize the product of worst case error and the weighted norm $\|F\|_{\mathscr{W}_{s,\gamma}}$ in the error bound (28). This idea is due to [LLS04, DSWW04] and was used in [KSS1] for determining the weights. For given s and N, we minimize a certain upper bound on the right hand side of

$$|I_s(G(u)) - Q_{s,N}(G(u))| \le e^{\mathrm{wor}}(Q_{s,N};\mathscr{W}_{s,\gamma})\,\|G(\cdot)\|_{\mathscr{X}'}\,\|u\|_{\mathscr{W}_\gamma(U;\mathscr{X})}\,.$$

Another issue raised by the infinite dimensional nature of the problem is to choose the value of s and estimate the truncation error $I(G(u)) - I_s(G(u))$, which was estimated in Theorem 6. The following results are generalizations of the results obtained in [KSS1] for the scalar, parametric diffusion problems.

Theorem 7 (Root-mean-square error bound). *Under Assumptions* (**A2**) *and* (10) *let b_j be defined thereFor every $f \in \mathscr{Y}'$ and for every $G(\cdot) \in \mathscr{X}'$, let $u(\cdot,y)$ denote the the solution of the parametric weak problem* (11)*. Then for $s,N \in \mathbb{N}$ and weights $\gamma = (\gamma_{\mathfrak{u}})$, 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 $\lambda \in (1/2,1]$,*

$$\sqrt{\mathbb{E}\,[|I_s(G(u)) - Q_{s,N}(\cdot;G(u))|^2]} \le \frac{C_\gamma(\lambda)}{[\varphi(N)]^{1/(2\lambda)}}\,\frac{\|f\|_{\mathscr{Y}'}\,\|G(\cdot)\|_{\mathscr{X}'}}{\gamma}\,, \tag{35}$$

where $\mathbb{E}[\cdot]$ denotes the expectation with respect to the random shift Δ (uniformly distributed over $[0,1]^s$), and where $C_\gamma(\lambda)$ and $\rho(\lambda)$ are as in [KSS1].

Theorem 7 is proved exactly as Theorem 6.1 in [KSS1].

In the nest theorem which is proved exactly as Theorem 6.1 in [KSS1], Assumption (36) with $p < 1$ is equivalent to Assumption (**A3**). Assumption (37) is equivalent to the condition that $\sum_{j\ge1} \frac{1}{2}\|\psi_j\|_{L^\infty(D)} < a_{\min}$ in [KSS1] which for the case $p = 1$ puts an additional restriction on the fluctuation of the random coefficients.

Theorem 8 (Optimal choice of weights). *Under the same assumptions and definitions as in Theorem 7, for b_j as in* (**A2**) *suppose that* (**A3**) *holds, i.e.*

$$\sum_{j\ge1} b_j^p < \infty \qquad \textit{for some} \quad 0 < p \le 1\,, \tag{36}$$

and when $p = 1$ assume additionally that

$$\sum_{j\ge1} b_j < 2\,. \tag{37}$$

Let

$$\lambda := \begin{cases} \frac{1}{2-2\delta} \quad \textit{for some} \quad \delta \in (0,1/2) & \textit{when } p \in (0,2/3]\,, \\ \frac{p}{2-p} & \textit{when } p \in (2/3,1)\,, \\ 1 & \textit{when } p = 1\,. \end{cases} \tag{38}$$

Then the choice of weights

$$\gamma_{\mathfrak{u}} = \gamma_{\mathfrak{u}}^* := \left(|\mathfrak{u}|! \prod_{j\in\mathfrak{u}} \frac{b_j}{\sqrt{\rho(\lambda)}} \right)^{2/(1+\lambda)} \tag{39}$$

minimizes the constant $C_\gamma(\lambda)$ in the bound (35)*. In particular, $C_\gamma(\lambda) < \infty$ and $\|u\|_{\mathscr{W}_\gamma(U;\mathscr{X})} < \infty$. Moreover, $C_\gamma(\frac{1}{2-2\delta}) \to \infty$ as $\delta \to 0$, and $C_\gamma(\frac{p}{2-p}) \to \infty$ as $p \to (2/3)^+$. Consequently, the root-mean-square error in Theorem 7 is of order $N^{-(1-\delta)}$ when $p \in (0,2/3]$, $N^{-(1/p-1/2)}$ when $p \in (2/3,1)$ and of order $N^{-1/2}$ when $p = 1$ (assuming that the Euler totient function $\varphi[N] \sim N$).*

5 Galerkin Discretization Error

To analyze the impact of approximate inversion of the operator equation by Galerkin discretization, we recall Section 2.5. For any $y \in \mathscr{U}$, the *parametric FE approximation* $u^h(\cdot,y) \in \mathscr{X}^h$ is defined as in (21). Here, $b(y;\cdot,\cdot)$ denotes the parametric bilinear form (9). In particular the FE approximation (21) is defined *pointwise* with respect to the parameter $y \in \mathscr{U}$.

Theorem 9. *Under Assumptions* (**A2**)*,* (10) *and* (16) *for every $f \in \mathscr{Y}'$ and for every $y \in \mathscr{U}$, the approximations $u^h(\cdot,y)$ are stable, i.e.* (22) *holds. For every $f \in \mathscr{Y}_t'$ with $0 < t \le 1$ exists a constant $C > 0$ such that as $h \to 0$ holds*

$$\|u(\cdot,y) - u^h(\cdot,y)\|_{\mathscr{X}} \le C h^t \|f\|_{\mathscr{Y}_t'} . \tag{40}$$

Proof. Since $f \in \mathscr{Y}_t'$ for some $t > 0$ implies with (16) that $u(y) \in \mathscr{X}_t$ and, with the approximation property (23),

$$\|u(\cdot,y) - u^h(\cdot,y)\|_{\mathscr{X}} \le C h^t \|u(\cdot,y)\|_{\mathscr{X}_t}$$

where the constant C is independent h and of y. This proves (40). □

Since we are interested in estimating the error in approximating functionals (30), we will also impose a regularity assumption on the functional $G(\cdot) \in \mathscr{X}'$:

$$\exists\, 0 < t' \le 1 : \quad G(\cdot) \in (\mathscr{X}')_{t'} \tag{41}$$

and the *adjoint regularity*: for t' as in (41), and for every $y \in \mathscr{U}$,

$$w(y) = (A^*(y))^{-1} G \in \mathscr{Y}_{t'} , \quad \sup_{y\in\mathscr{U}} \|w(y)\|_{\mathscr{Y}_{t'}} \le C \|G\|_{\mathscr{X}_{t'}'} . \tag{42}$$

Moreover, since in the expression (30) 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 known that $|G(u(\cdot,y)) - G(u^h(\cdot,y))|$ may converge faster than $\|u(\cdot,y) - u^h(\cdot,y)\|_{\mathscr{X}}$.

Theorem 10. *Under Assumptions* (**A2**)*,* (10)*,* (**A4**)*, and* (16)*,* (42) *for every $f \in \mathscr{Y}_t'$ with $0 < t \le 1$, for every $G(\cdot) \in \mathscr{X}_{t'}'$ with $0 < t' \le 1$ and for every $y \in \mathscr{U}$, as $h \to$*

0, *there exists a constant* $C > 0$ *independent of* $h > 0$ *and of* $y \in U$ *such that the Galerkin approximations* $G(u^h(\cdot,y))$ *satisfy*

$$\left| G(u(\cdot,y)) - G(u^h(\cdot,y)) \right| \leq C h^{\tau} \|f\|_{\mathscr{X}_t'} \|G(\cdot)\|_{\mathscr{X}_{t'}'} , \tag{43}$$

where $0 < \tau := t + t'$ *and where the constant* $C > 0$ *is independent of* $y \in U$.

Proof. The error bound (43) follows from an Aubin-Nitsche duality argument together with the regularity assumptions (16), (41).

6 Combined QMC Galerkin Error Analysis

We conclude with bounds for the combined QME FE approximation of the integral (30). To obtain a computable approximation of (30), 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 Δ will be denoted by $Q_{s,N}(\cdot;\Delta)$. Moreover, for each evaluation of the integrand, we replace the exact solution $u(\cdot,y)$ of the parametric weak problem (11) by its Galerkin approximation $u^h(\cdot,y) \in \mathscr{X}^h \subset \mathscr{X}$ from a FE space $\mathscr{X}^h$ 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*:

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

We bound the mean-square error with respect to the random shift by

$$\begin{aligned} \mathbb{E}\left[|I(G(u)) - Q_{s,N}(G(u^h);\cdot)|^2 \right] &\leq 3\,|(I - I_s)(G(u))|^2 \\ &\quad + 3\,\mathbb{E}\left[|I_s(G(u)) - Q_{s,N}(G(u);\cdot)|^2 \right] + 3\,\mathbb{E}\left[|Q_{s,N}(G(u-u^h);\cdot)|^2 \right] . \end{aligned} \tag{44}$$

The dimension truncation error, i.e., the first term in (44), was provided in Theorem 6. The QMC error, i.e., the second term in (44), is already analyzed in Theorem 7 (together with Theorem 8). Finally, for the Galerkin projection error, i.e., for the third term in (44), we apply the property that the QMC quadrature weights $1/N$ are positive and sum to 1, to obtain

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

and then apply Theorem 10. We summarize the combined error estimate and cost analysis in the following theorem.

Theorem 11. *Under the same assumptions and definitions as in Theorems 6, 7, 8, and 10, if we approximate the integral over* $\mathscr{U}$ *by the randomly shifted lattice rule*

from Theorems 7 and 8 with N points in s dimensions, and for each shifted lattice point we solve the approximate elliptic problem (21) *by Galerkin discretization in the domain D with* one common subspace $\mathscr{X}^h$ *with* $M_h = \dim(\mathscr{X}^h)$ *degrees of freedom and with the approximation property* (18) *with linear cost* $\mathscr{O}(M_h)$ *(for example, by multilevel methods), there holds the root-mean-square error bound*

$$\sqrt{\mathbb{E}\left[|I(G(u)) - Q_{s,N}(\cdot; G(u^h))|^2\right]}$$
$$\leq C\left(\kappa(s,N)\,\|f\|_{\mathscr{Y}'}\,\|G(\cdot)\|_{\mathscr{X}'} + h^\tau\,\|f\|_{\mathscr{Y}'_t}\|G(\cdot)\|_{\mathscr{X}'_{t'}}\right)\,,$$

where $\tau = t + t'$*, and*

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

up to a double logarithmic factor of N when N is not prime due to the Euler totient function $\varphi(N)$ *in* (35)*.*

References

[BF] F. Brezzi and M. Fortin, Mixed and Hybrid Finite Element Methods, Springer Verlag, Berlin, 1991.

[CDS1] A. Cohen, R. DeVore and Ch. Schwab, Convergence rates of best *N*-term Galerkin approximation for a class of elliptic sPDEs, Found. Comput. Math. **10** (6), 2010, 615–646. DOI 10.1007/s10208-010-9072-2

[CKN06] R. COOLS, F. Y. KUO, AND D. NUYENS, *Constructing embedded lattice rules for multivariate integration*, SIAM J. Sci. Comput., 28 (2006), pp. 2162–2188.

[D04] J. DICK, *On the convergence rate of the component-by-component construction of good lattice rules*, J. Complexity, 20 (2004), pp. 493–522.

[DiPi10] J. DICK AND F. PILLICHSHAMMER, *Digital Nets and Sequences*, Cambridge University Press, 2010.

[DSWW04] J. DICK, I. H. SLOAN, X. WANG, AND H. WOŹNIAKOWSKI, *Liberating the weights*, J. Complexity, 20 (2004), pp. 593–623.

[GKN3S] I.G. Graham, F.Y.Kuo, J.A. Nichols, R. Scheichl, Ch. Schwab and I.H. Sloan, Quasi-Monte Carlo Finite Element Methods for elliptic PDEs with log-normal random coefficients (in review).

[HaSc11] M. Hansen and Ch. Schwab, *Analytic regularity and best N-term approximation of high dimensional parametric initial value problems*, Report 2011-64, Seminar for Applied Mathematics, ETH Zürich (in review).

[HoSc12] V.H. Hoang and Ch. Schwab, *Analytic regularity and gpc approximation for parametric and random 2nd order hyperbolic PDEs*, Report 2010-19, Seminar for Applied Mathematics, ETH Zürich (to appear in Analysis and Applications (2011)).

[KS11] A. Kunoth and Ch. Schwab, Analytic Regularity and GPC Approximation for Stochastic Control Problems Constrained by Linear Parametric Elliptic and Parabolic PDEs, Report 2011-54, www.sam.math.ethz.ch/ reports Seminar für Angewandte Mathematik, ETH Zürich (in review).

[KSS1] F. Y. Kuo, Ch. Schwab, and I. H. Sloan, *Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficient* (to appear in SIAM J. Numer. Anal.)

[KSS2] 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*, ANZIAM J. (2012) (to appear).

[KSS3] F. Y. Kuo, Ch. Schwab, and I. H. Sloan, *Multi-Level Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficient* (in preparation).

[KSWWat10] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski and B. J. Waterhouse, *Randomly shifted lattice rules with the optimal rate of convergence for unbounded integrands*, J. Complexity **26**, 135–160 (2010).

[KWWat06] F. Y. Kuo, G. W. Wasilkowski, and B. J. Waterhouse, *Randomly shifted lattice rules for unbounded integrals*, J. Complexity **22**, 630–651 (2006).

[K03] F. Y. Kuo, *Component-by-component constructions achieve the optimal rate of convergence*, J. Complexity, 19 (2003), pp 301–320.

[KSWW10b] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Woźniakowski, *Liberating the dimension*, J. Complexity, 26 (2010), pp. 422–454.

[LLS04] G. Larcher, G. Leobacher, and K. Scheicher, *On the tractability of the Brownian bridge algorithm*, J. Complexity, 19 (2004), pp. 511–528.

[Nie92] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.

[NC06a] 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.

[ScGJGActa11] C. Schwab and C.J. Gittelson, *Sparse tensor discretizations of high-dimensional parametric and stoch astic PDEs*, Acta Numerica **20** (2011), Cambridge University Press.

[ST06] 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.

[SJ94] I. H. Sloan and S. Joe, *Lattice Methods for Multiple Integration*, Oxford University Press, Oxford, 1994.

[SKJ02] 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.

[SW98] I. H. Sloan and H. Woźniakowski, *When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals?*, J. Complexity, 14 (1998), pp. 1–33.

[SWW04] I. H. Sloan, X. Wang, and H. Woźniakowski, *Finite-order weights imply tractability of multivariate integration*, J. Complexity, 20 (2004), pp. 46–74.

[W02] X. Wang, *Strong tractability of multivariate integration using quasi-Monte Carlo algorithms*, Math. Comp., 72 (2002), pp. 823–838.

Research Reports

No.	Authors/Title
12-18	*Ch. Schwab* QMC Galerkin discretization of parametric operator equations
12-17	*N.H. Risebro, Ch. Schwab and F. Weber* Multilevel Monte-Carlo front tracking for random scalar conservation laws
12-16	*R. Andreev and Ch. Tobler* Multilevel preconditioning and low rank tensor iteration for space-time simultaneous discretizations of parabolic PDEs
12-15	*V. Gradinaru and G.A. Hagedorn* A timesplitting for the semiclassical Schrödinger equation
12-14	*Ph. Grohs and G. Kutyniok* Parabolic molecules
12-13	*V. Kazeev, O. Reichmann and Ch. Schwab* Low-rank tensor structure of linear diffusion operators in the TT and QTT formats
12-12	*F. Müller, P. Jenny and D.W. Meyer* Multilevel Monte Carlo for two phase flow and transport in random heterogeneous porous media
12-11	*V. Kazeev, O. Reichmann and Ch. Schwab* *hp*-DG-QTT solution of high-dimensional degenerate diffusion equations
12-10	N.H. Risebro and F. Weber A note on front tracking for the Keyfitz-Kranzer system
12-09	*U. Koley and N.H. Risebro* Convergence of finite difference schemes for symmetric Keyfitz-Kranzer system
12-08	*S. Mishra, Ch. Schwab and J. Šukys* Monte Carlo and multi-level Monte Carlo finite volume methods for uncertainty quantification in nonlinear systems of balance laws
12-07	*A. Hiltebrand and S. Mishra* Entropy stable shock capturing space-time discontinuous Galerkin schemes for systems of conservation laws
12-06	*R. Hiptmair, A. Moiola and I. Perugia* Trefftz discontinuous Galerkin methods for acoustic scattering on locally refined meshes