Uncertainty Quantification and Quasi-Monte Carlo

Sommersemester 2025

Vesa Kaarnioja
vesa.kaarnioja@fu-berlin.de

FU Berlin, FB Mathematik und Informatik

Eighth lecture, June 16, 2025

Recap: Suppose that $f \in H_{s,\boldsymbol{\gamma}}$ for all $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u}\subseteq\{1:s\}}$. The unanchored, weighted Sobolev space $H_{s,\boldsymbol{\gamma}}$ is equipped with the norm

$$\|f\|_{s,\boldsymbol{\gamma}}^2 := \sum_{\mathfrak{u}\subseteq\{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \int_{[0,1]^{|\mathfrak{u}|}} \bigg(\int_{[0,1]^{s-|\mathfrak{u}|}} \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} f(\boldsymbol{y}) \,\mathrm{d}\boldsymbol{y}_{-\mathfrak{u}} \bigg)^2 \,\mathrm{d}\boldsymbol{y}_{\mathfrak{u}}.$$

For any given sequence of weights $\boldsymbol{\gamma}$, we can use the CBC algorithm (*implementational details were considered during the* $7^{\rm th}$ *lecture*) to obtain a generating vector for a randomly shifted rank-1 lattice QMC rule satisfying the error bound

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}|I_s f - Q_{n,s}^{\boldsymbol{\Delta}} f|^2} \leq \bigg(\frac{1}{\varphi(n)} \sum_{\varnothing\neq\mathfrak{u}\subseteq\{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \bigg(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}} \bigg)^{|\mathfrak{u}|} \bigg)^{1/(2\lambda)} \|f\|_{s,\boldsymbol{\gamma}} \quad (1)$$

for all $\lambda \in (1/2, 1]$. We can use the following strategy:

- For a given integrand f, estimate the norm $\|f\|_{s,\boldsymbol{\gamma}}$.
- Find weights $\boldsymbol{\gamma}$ which *minimize* the error bound (1).
- Using the optimized weights $\boldsymbol{\gamma}$ as input, use the CBC algorithm to find a generating vector which *satisfies* the error bound (1).

Application to parametric PDE problems

For the application of QMC methods to parametric PDE problems, we follow the survey papers

- F. Y. Kuo and D. Nuyens. Application of quasi-Monte Carlo methods to elliptic PDEs with random diffusion coefficients - a survey of analysis and implementation. *Found. Comput. Math.* **16**:1631–1696, 2016. arXiv version: https://arxiv.org/abs/1606.06613
- F. Y. Kuo and D. Nuyens. Application of quasi-Monte Carlo methods to PDEs with random coefficients – an overview and tutorial. In: A. Owen and P. Glynn (eds), *Monte Carlo and Quasi-Monte Carlo Methods 2016*, pp. 53–71. arXiv version: https://arxiv.org/abs/1710.10984

Let us first consider applying QMC for the uniform and affine model problem discussed during week 4.

Recall the *uniform and affine model*: let $D \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded Lipschitz domain, let $f \in L^2(D)$, and let $U := [-1/2, 1/2]^{\mathbb{N}} := \{(a_j)_{j\geq 1} : -1/2 \leq a_j \leq 1/2\}$ be a set of parameters. Consider the problem of finding, for all $\boldsymbol{y} \in U$, $u(\cdot, \boldsymbol{y}) \in H_0^1(D)$ such that

$$\int_D a(\boldsymbol{x}, \boldsymbol{y}) \nabla u(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_D f(\boldsymbol{x}) v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \quad \text{for all } v \in H_0^1(D),$$

where the diffusion coefficient has the parameterization

$$a(\boldsymbol{x}, \boldsymbol{y}) := a_0(\boldsymbol{x}) + \sum_{j=1}^{\infty} y_j \psi_j(\boldsymbol{x}), \quad \boldsymbol{x} \in D,\ \boldsymbol{y} \in U,$$

where $a_0 \in L^\infty(D)$, there exist $a_{\min}, a_{\max} > 0$
s.t. $0 < a_{\min} \leq a(\boldsymbol{x}, \boldsymbol{y}) \leq a_{\max} < \infty$ for all $\boldsymbol{x} \in D$ and $\boldsymbol{y} \in U$, and the *stochastic fluctuations* $\psi_j \colon D \to \mathbb{R}$ are functions of the spatial variable such that

- $\psi_j \in L^\infty(D)$ for all $j \in \mathbb{N}$,
- $\sum_{j=1}^\infty \|\psi_j\|_{L^\infty(D)} < \infty$,
- $\sum_{j=1}^\infty \|\psi_j\|_{L^\infty(D)}^p < \infty$ for some $p \in (0,1)$.

Total error decomposition

In practice, we need to truncate the infinite-dimensional parametric vector $\boldsymbol{y} \in [-1/2, 1/2]^{\mathbb{N}}$ to a finite number of terms. Moreover, the PDE needs to be discretized spatially using, e.g., the finite element method.

Let $u_s(\boldsymbol{y}) := u_s(y_1, \ldots, y_s, 0, 0, \ldots)$ denote the dimensionally-truncated PDE solution for $\boldsymbol{y} \in [-1/2, 1/2]^{\mathbb{N}}$ (we often abuse notation and also write $u_s(\boldsymbol{y})$ for $\boldsymbol{y} \in [-1/2, 1/2]^s$), and let $u_{s,h}(\cdot, \boldsymbol{y}) \in V_h$ denote the dimensionally-truncated FE solution in the FE subspace spanned by piecewise linear FE basis functions. Furthermore, let $\{\boldsymbol{t}_i\}_{i=1}^n$ be a QMC point set in $[-1/2, 1/2]^s$.

Total error decomposition

For simplicity, let us consider the problem of computing $\mathbb{E}[G(u)]$, where $u(\cdot,\boldsymbol{y}) \in H_0^1(D)$ is the PDE solution for $\boldsymbol{y} \in U$ and $G\colon H_0^1(D) \to \mathbb{R}$ is a linear functional (quantity of interest). We decompose the total error as

$$\begin{aligned}
&\int_{[-1/2,1/2]^{\mathbb{N}}} G(u(\cdot,\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y} - \frac{1}{n}\sum_{i=1}^{n} G(u_{s,h}(\cdot,\boldsymbol{t}_i))\\
&= \int_{[-1/2,1/2]^{\mathbb{N}}} (G(u(\cdot,\boldsymbol{y}) - u_s(\cdot,\boldsymbol{y})))\,\mathrm{d}\boldsymbol{y}\\
&+ \int_{[-1/2,1/2]^{s}} G(u_s(\cdot,\boldsymbol{y}) - u_{s,h}(\cdot,\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y}\\
&+ \int_{[-1/2,1/2]^{s}} G(u_{s,h}(\cdot,\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y} - \frac{1}{n}\sum_{i=1}^{n} G(u_{s,h}(\cdot,\boldsymbol{t}_i)).
\end{aligned}$$

Using the triangle inequality, we are left with the total error decomposition

$$\bigg|\int_{[-1/2,1/2]^{\mathbb{N}}} G(u(\cdot,\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y} - \frac{1}{n}\sum_{i=1}^{n} G(u_{s,h}(\cdot,\boldsymbol{t}_i))\bigg|$$

$$\leq \bigg|\int_{[-1/2,1/2]^{\mathbb{N}}} (G(u(\cdot,\boldsymbol{y}) - u_s(\cdot,\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y}\bigg| \quad \text{(dimension-truncation error)}$$

$$+ \bigg|\int_{[-1/2,1/2]^{s}} G(u_s(\cdot,\boldsymbol{y}) - u_{s,h}(\cdot,\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y}\bigg| \quad \text{(finite element error)}$$

$$+ \bigg|\int_{[-1/2,1/2]^{s}} G(u_{s,h}(\cdot,\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y} - \frac{1}{n}\sum_{i=1}^{n} G(u_{s,h}(\cdot,\boldsymbol{t}_i))\bigg|. \quad \text{(cubature error)}$$

Let us focus today on the cubature error.

Remarks:

- We'll discuss the other error contributions (dimension truncation and finite element errors) later. Furthermore, we'll see how the analysis differs in the lognormal setting.
- It turns out that if we can control the error for all linear quantities of interest $G\colon H_0^1(D)\to\mathbb{R}$, we can control the error for the full PDE solution with respect to the $\|\cdot\|_{H_0^1(D)}$ norm using a duality argument.

Multi-index notation

We introduce the set of *finitely-supported multi-indices*

$$\mathscr{F} := \{\boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}} : |\mathrm{supp}(\boldsymbol{\nu})| < \infty\},$$

where the *support* of a multi-index $\boldsymbol{\nu}$ is defined as the set

$$\mathrm{supp}(\boldsymbol{\nu}) := \{i \in \mathbb{N} : \nu_i \neq 0\}.$$

As before, the *order* of a multi-index is defined as

$$|\boldsymbol{\nu}| := \sum_{j \geq 1} \nu_j$$

and we use the special multi-index notations

$$\partial^{\boldsymbol{\nu}} := \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} := \prod_{j \in \mathrm{supp}(\boldsymbol{\nu})} \frac{\partial^{\nu_j}}{\partial y_j^{\nu_j}}, \ \boldsymbol{x}^{\boldsymbol{\nu}} := \prod_{j \in \mathrm{supp}(\boldsymbol{\nu})} x_j^{\nu_j}, \ \binom{\boldsymbol{\nu}}{\boldsymbol{m}} := \prod_{j \in \mathrm{supp}(\boldsymbol{\nu})} \binom{\nu_j}{m_j}.$$

Recursive bound

Consider the weak formulation

$$\int_D a(\boldsymbol{x},\boldsymbol{y})\nabla u(\boldsymbol{x},\boldsymbol{y})\cdot\nabla v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_D f(\boldsymbol{x})v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{2}$$

Noting that

$$\partial^{\boldsymbol{\nu}} a(\boldsymbol{x},\boldsymbol{y}) = \begin{cases} a(\boldsymbol{x},\boldsymbol{y}) & \text{if } \boldsymbol{\nu}=\mathbf{0},\\ \psi_j(\boldsymbol{x}) & \text{if } \boldsymbol{\nu}=\boldsymbol{e}_j,\\ 0 & \text{otherwise,}\end{cases}$$

then by differentiating (2) on both sides with $\partial^{\boldsymbol{\nu}}$ and using the Leibniz product rule† yields

$$\begin{aligned}
&\partial^{\boldsymbol{\nu}}\int_D a(\boldsymbol{x},\boldsymbol{y})\nabla u(\boldsymbol{x},\boldsymbol{y})\cdot\nabla v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0\\
\Leftrightarrow\quad & \sum_{\boldsymbol{m}\le\boldsymbol{\nu}}\binom{\boldsymbol{\nu}}{\boldsymbol{m}}\int_D \partial^{\boldsymbol{m}}a(\boldsymbol{x})\nabla\partial^{\boldsymbol{\nu}-\boldsymbol{m}}u(\boldsymbol{x},\boldsymbol{y})\cdot\nabla v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0\\
\Leftrightarrow\quad & \int_D a(\boldsymbol{x},\boldsymbol{y})\nabla\partial^{\boldsymbol{\nu}}u(\boldsymbol{x},\boldsymbol{y})\cdot\nabla v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = -\sum_{j\in\mathrm{supp}(\boldsymbol{\nu})}\nu_j\int_D\psi_j(\boldsymbol{x})\nabla\partial^{\boldsymbol{\nu}-\boldsymbol{e}_j}u(\boldsymbol{x},\boldsymbol{y})\cdot\nabla v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

† $\partial^{\boldsymbol{\nu}}(fg) = \sum_{\boldsymbol{m}\le\boldsymbol{\nu}}\binom{\boldsymbol{\nu}}{\boldsymbol{m}}\partial^{\boldsymbol{m}}f\,\partial^{\boldsymbol{\nu}-\boldsymbol{m}}g$ (exercise)

Testing this against $v = \partial^{\boldsymbol{\nu}} u(\boldsymbol{x}, \boldsymbol{y})$ yields

$$
\begin{aligned}
&a_{\min} \|\partial^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_{H_0^1(D)}^2 \\
&\leq \int_D a(\boldsymbol{x}, \boldsymbol{y}) \|\nabla \partial^{\boldsymbol{\nu}} u(\boldsymbol{x}, \boldsymbol{y})\|^2 \, \mathrm{d}\boldsymbol{x} \\
&\leq \sum_{j \in \operatorname{supp}(\boldsymbol{\nu})} \nu_j \|\psi_j\|_{L^\infty(D)} \|\partial^{\boldsymbol{\nu} - \boldsymbol{e}_j} u(\cdot, \boldsymbol{y})\|_{H_0^1(D)} \|\partial^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_{H_0^1(D)}
\end{aligned}
$$

Thus we obtain the recursive relation

$$
\|\partial^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_{H_0^1(D)} \leq \sum_{j \in \operatorname{supp}(\boldsymbol{\nu})} \nu_j \underbrace{\frac{\|\psi_j\|_{L^\infty(D)}}{a_{\min}}}_{=: b_j} \|\partial^{\boldsymbol{\nu} - \boldsymbol{e}_j} u(\cdot, \boldsymbol{y})\|_{H_0^1(D)}.
$$

For later convenience, we introduce here the sequence $\boldsymbol{b} := (b_j)_{j \geq 1}$ defined by $b_j := \frac{\|\psi_j\|_{L^\infty(D)}}{a_{\min}}$. Recall that by the assumptions we placed on the uniform and affine model, there holds $\boldsymbol{b} \in \ell^p$ for some $p \in (0, 1)$.

Parametric regularity

Proposition

For all $\boldsymbol{y} \in [-1/2, 1/2]^{\mathbb{N}}$ *and* $\boldsymbol{\nu} \in \mathscr{F}$*, there holds*

$$\|\partial^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_{H_0^1(D)} \leq \frac{C_P \|f\|_{L^2(D)}}{a_{\min}} \boldsymbol{b}^{\boldsymbol{\nu}} |\boldsymbol{\nu}|!,$$

where C_P *is the Poincaré constant satisfying* $\|v\|_{L^2(D)} \leq C_P \|v\|_{H_0^1(D)}$ *for all* $v \in H_0^1(D)$*.*

Proof. By induction w.r.t. the order of the multi-index $\boldsymbol{\nu} \in \mathscr{F}$. If $\boldsymbol{\nu} = \mathbf{0}$, then this is the ordinary Lax–Milgram *a priori* bound

$$a_{\min} \underbrace{\int_D |\nabla u(\boldsymbol{x}, \boldsymbol{y})|^2 \, \mathrm{d}\boldsymbol{x}}_{=\|u(\cdot,\boldsymbol{y})\|^2_{H_0^1(D)}} \leq \int_D a(\boldsymbol{x}, \boldsymbol{y}) \nabla u(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla u(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} = \int_D f(\boldsymbol{x}) u(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x}$$

$$\leq \|f\|_{L^2(D)} \|u(\cdot, \boldsymbol{y})\|_{L^2(D)} \leq C_P \|f\|_{L^2(D)} \|u(\cdot, \boldsymbol{y})\|_{H_0^1(D)}$$

whence

$$\|u(\cdot, \boldsymbol{y})\|_{H_0^1(D)} \leq \frac{C_P \|f\|_{L^2(D)}}{a_{\min}}.$$

Next, let $\boldsymbol{\nu} \in \mathscr{F}$ and suppose that the claim has been proved for all multi-indices with order less than $|\boldsymbol{\nu}|$. Then using the recursive relation we derived previously, we obtain

$$\begin{aligned}
\|\partial^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_{H_0^1(D)} &\leq \sum_{j \in \operatorname{supp}(\boldsymbol{\nu})} \nu_j b_j \|\partial^{\boldsymbol{\nu} - \boldsymbol{e}_j} u(\cdot, \boldsymbol{y})\|_{H_0^1(D)} \\
&\leq \frac{C_P \|f\|_{L^2(D)}}{a_{\min}} \sum_{j \in \operatorname{supp}(\boldsymbol{\nu})} \nu_j b_j |\boldsymbol{\nu} - \boldsymbol{e}_j|! \boldsymbol{b}^{\boldsymbol{\nu} - \boldsymbol{e}_j} \\
&= \frac{C_P \|f\|_{L^2(D)}}{a_{\min}} \boldsymbol{b}^{\boldsymbol{\nu}} (|\boldsymbol{\nu}| - 1)! \sum_{j \geq 1} \nu_j \\
&= \frac{C_P \|f\|_{L^2(D)}}{a_{\min}} \boldsymbol{b}^{\boldsymbol{\nu}} |\boldsymbol{\nu}|!,
\end{aligned}$$

as desired. □

Remark. Note that the same regularity bound holds for the dimensionally-truncated FE solution $u_{s,h}$ as long as a (conforming) Galerkin FE discretization has been used to construct the FE approximation. This is due to the fact that the weak formulation of the Galerkin discretization is exactly the same (only the function space differs).

Now that we know the regularity of the PDE problem, we can analyze the QMC cubature error! Let $G\colon H_0^1(D) \to \mathbb{R}$ be a linear and bounded functional, $u_{s,h}$ the dimensionally-truncated FE solution, and define $F(\boldsymbol{y}) := G(u_{s,h}(\cdot, \boldsymbol{y} - \frac{\mathbf{1}}{\mathbf{2}}))$ for $\boldsymbol{y} \in [0,1]^s$. Let $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u} \subseteq \{1:s\}}$ be a sequence of positive weights. Then we know that the generating vector obtained by the CBC algorithm satisfies the error bound

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}|I_s F - Q_{n,s}^{\boldsymbol{\Delta}} F|^2} \leq \bigg(\frac{1}{\varphi(n)} \sum_{\varnothing \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \bigg(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}}\bigg)^{|\mathfrak{u}|}\bigg)^{1/(2\lambda)} \|F\|_{s,\boldsymbol{\gamma}}$$

for all $\lambda \in (1/2, 1]$, where

$$\begin{aligned}\|F\|_{s,\boldsymbol{\gamma}}^2 &= \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \int_{[0,1]^{|\mathfrak{u}|}} \bigg(\int_{[0,1]^{s-|\mathfrak{u}|}} \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{x}_{\mathfrak{u}}} F(\boldsymbol{y}) \,\mathrm{d}\boldsymbol{y}_{-\mathfrak{u}}\bigg)^2 \mathrm{d}\boldsymbol{y}_{\mathfrak{u}} \\ &\leq \bigg(\frac{C_P \|G\|_{H_0^1(D)\to\mathbb{R}} \|f\|_{L^2(D)}}{a_{\min}}\bigg)^2 \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} (|\mathfrak{u}|!)^2 \prod_{j \in \mathfrak{u}} b_j^2.\end{aligned}$$

Plugging this norm bound back into the QMC error bound yields...

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}|I_s F - Q_{n,s}^{\boldsymbol{\Delta}} F|^2} \lesssim \left(\frac{1}{\varphi(n)}\right)^{1/(2\lambda)} \left(\sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}}\right)^{|\mathfrak{u}|}\right)^{1/(2\lambda)}$$
$$\times \left(\sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} (|\mathfrak{u}|!)^2 \prod_{j \in \mathfrak{u}} b_j^2\right)^{1/2}.$$

The upper bound can be *minimized* by choosing the *POD weights*

$$\gamma_{\mathfrak{u}} := \left(|\mathfrak{u}|! \prod_{j \in \mathfrak{u}} \frac{b_j}{\sqrt{\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}}}}\right)^{2/(1+\lambda)},$$

as explained by the following lemma.

Lemma

Let (α_i) *and* (β_i) *be sequences of positive real numbers. The expression*

$$g(\boldsymbol{\gamma}) := \left(\sum_i \alpha_i \gamma_i^{\lambda}\right)^{1/\lambda} \left(\sum_i \beta_i \gamma_i^{-1}\right)$$

is minimized by $\gamma_i = c\left(\frac{\beta_i}{\alpha_i}\right)^{1/(1+\lambda)}$ *for arbitrary* $c > 0$.

Proof. Let us find out when the gradient vanishes:

$$0 = \partial_j g(\gamma) = \frac{1}{\lambda}\left(\sum_i \alpha_i \gamma_i^\lambda\right)^{1/\lambda - 1} \lambda \alpha_j \gamma_j^{\lambda-1}\left(\sum_i \beta_i \gamma_i^{-1}\right) - \beta_j \gamma_j^{-2}\left(\sum_i \alpha_i \gamma_i^\lambda\right)^{1/\lambda}.$$

After some trivial simplifications, we can see that this is equivalent to

$$\gamma_j^{\lambda+1} = \frac{\beta_j}{\alpha_j}\frac{\sum_i \alpha_i \gamma_i^\lambda}{\sum_i \beta_i \gamma_i^{-1}}.$$

Furthermore, this condition is satisfied if

$$\gamma_j = c\left(\frac{\beta_j}{\alpha_j}\right)^{1/(1+\lambda)},$$

where $c > 0$ is arbitrary. □

Note that plugging $\gamma_i = c\left(\frac{\beta_i}{\alpha_i}\right)^{1/(1+\lambda)}$ into $\left(\sum_i \alpha_i \gamma_i^{\lambda}\right)^{1/(2\lambda)}\left(\sum_i \beta_i \gamma_i^{-1}\right)^{1/2}$ yields the expression $\left(\sum_i \alpha_i^{1/(1+\lambda)} \beta_i^{\lambda/(1+\lambda)}\right)^{(1+\lambda)/(2\lambda)}$. Thus, plugging the optimal POD weights into the QMC error bound results in

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}|I_s F - Q_{n,s}^{\boldsymbol{\Delta}} F|^2} \lesssim \left(\frac{1}{\varphi(n)}\right)^{1/(2\lambda)} C(s,\boldsymbol{\gamma},\lambda)^{(1+\lambda)/(2\lambda)},$$

where

$$C(s,\boldsymbol{\gamma},\lambda) := \sum_{\mathfrak{u}\subseteq\{1:s\}} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}}\right)^{|\mathfrak{u}|/(1+\lambda)} (|\mathfrak{u}|!)^{2\lambda/(1+\lambda)} \prod_{j\in\mathfrak{u}} b_j^{2\lambda/(1+\lambda)}.$$

This is the punchline:

Lemma

By choosing

$$\lambda = \begin{cases} \frac{p}{2-p} & \text{when } p \in (2/3,1) \\ \frac{1}{2-2\delta} \text{ for arbitrary } \delta \in (0,1/2) & \text{when } p \in (0,2/3], \end{cases}$$

there exists a constant $C(\boldsymbol{\gamma},\lambda) < \infty$ independently of s s.t. $C(s,\boldsymbol{\gamma},\lambda) \leq C(\boldsymbol{\gamma},\lambda) < \infty$.

Proof. First observe that

$$C(s,\gamma,\lambda) = \sum_{\mathfrak{u}\subseteq\{1:s\}} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^\lambda}\right)^{|\mathfrak{u}|/(1+\lambda)} (|\mathfrak{u}|!)^{2\lambda/(1+\lambda)} \prod_{j\in\mathfrak{u}} b_j^{2\lambda/(1+\lambda)}$$

$$= \sum_{\ell=0}^{s} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^\lambda}\right)^{\ell/(1+\lambda)} (\ell!)^{2\lambda/(1+\lambda)} \sum_{\substack{|\mathfrak{u}|=\ell \\ \mathfrak{u}\subseteq\{1:s\}}} \prod_{j\in\mathfrak{u}} b_j^{2\lambda/(1+\lambda)}$$

$$\leq \sum_{\ell=0}^{\infty} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^\lambda}\right)^{\ell/(1+\lambda)} (\ell!)^{2\lambda/(1+\lambda)-1} \left(\sum_{j\geq 1} b_j^{2\lambda/(1+\lambda)}\right)^{\ell}$$

where we used the inequality $\sum_{|\mathfrak{u}|=\ell,\mathfrak{u}\subseteq\mathbb{Z}_+} \prod_{j\in\mathfrak{u}} c_j \leq \frac{1}{\ell!}\big(\sum_{j\geq 1} c_j\big)^{\ell}$.

Case 1: $p \in (2/3, 1)$**.** We choose $p = \frac{2\lambda}{1+\lambda} \Leftrightarrow \lambda = \frac{p}{2-p} \in (1/2, 1)$, and

$$C(s,\gamma,\lambda) \leq \sum_{\ell=0}^{\infty} \underbrace{\left(\frac{2\zeta(2\lambda)}{(2\pi^2)^\lambda}\right)^{\ell/(1+\lambda)} (\ell!)^{p-1} \left(\sum_{j\geq 1} b_j^{p}\right)^{\ell}}_{=:a_\ell}$$

It is easy to see that $\frac{a_{\ell+1}}{a_\ell} \xrightarrow{\ell\to\infty} 0$. By the ratio test, this upper bound is finite independently of s.

Case 2: $p \in (0, 2/3]$. Let $\delta \in (0, 1/2)$ be arbitrary. We choose $\lambda = \frac{1}{2-2\delta} \in (1/2, 1)$. Now $\frac{2\lambda}{1+\lambda} = \frac{2}{3-2\delta} \in (2/3, 1)$. Especially, $\|\boldsymbol{b}\|_{\ell^{2\lambda/(1+\lambda)}} \leq \|\boldsymbol{b}\|_{\ell^p}$, and we obtain from the estimate on the previous slide that

$$C(s,\gamma,\lambda) \leq \sum_{\ell=0}^{\infty} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}}\right)^{\ell/(1+\lambda)} (\ell!)^{2\lambda/(1+\lambda)-1} \left(\sum_{j\geq 1} b_j^{2\lambda/(1+\lambda)}\right)^{\ell}$$

$$\leq \sum_{\ell=0}^{\infty} \underbrace{\left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}}\right)^{\ell/(1+\lambda)} (\ell!)^{2/(3-2\delta)-1} \left(\sum_{j\geq 1} b_j^{p}\right)^{2\ell/((3-2\delta)p)}}_{=:a_\ell}$$

Again, $\frac{a_{\ell+1}}{a_\ell} \xrightarrow{\ell\to\infty} 0$, so by the ratio test this upper bound is finite independently of s. $\square$

Theorem

Let $\delta \in (0,1/2)$ be arbitrary. By choosing the POD weights

$$\gamma_{\mathfrak{u}} := \left(|\mathfrak{u}|! \prod_{j\in\mathfrak{u}} \frac{b_j}{\sqrt{\frac{2\zeta(2\lambda)}{(2\pi^2)^\lambda}}} \right)^{2/(1+\lambda)}, \quad \lambda := \begin{cases} \frac{p}{2-p} & \text{if } p \in (2/3,1), \\ \frac{1}{2-2\delta} & \text{if } p \in (0,2/3], \end{cases}$$

then the QMC approximation for the expected value of the PDE problem satisfies

$$\text{R.M.S. error} \lesssim \begin{cases} \left(\frac{1}{\varphi(n)}\right)^{1/p-1/2} & \text{if } p \in (2/3,1), \\ \left(\frac{1}{\varphi(n)}\right)^{1-\delta} & \text{if } p \in (0,2/3], \end{cases}$$

where the implied coefficient is independent of the dimension s.

Remark: We have the following dimension-independent convergence rates:

- n is prime $\Rightarrow \frac{1}{\varphi(n)} = \frac{1}{n-1} \Rightarrow$ QMC rate $\mathcal{O}(n^{\max\{-1/p+1/2,-1+\delta\}})$.
- $n = 2^k \Rightarrow \frac{1}{\varphi(n)} = \frac{2}{n} \Rightarrow$ QMC rate $\mathcal{O}(n^{\max\{-1/p+1/2,-1+\delta\}})$.
- For general composite n, the dimension-independent QMC rate is at best essentially linear up to a double logarithmic factor of n.

Remarks on implementation

Let $G\colon H_0^1(D) \to \mathbb{R}$ be a bounded linear functional. Consider the problem of approximating

$$\mathbb{E}[G(u_{s,h})] = \int_{[-1/2,1/2]^s} G(u_{s,h}(\cdot,\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y},$$

where $u_{s,h}$ is the dimensionally-truncated FE approximation to the elliptic PDE with a uniform and affine diffusion coefficient.

Our QMC approximation is guaranteed to satisfy the R.M.S. error bound from the previous slide if we plug the theoretically derived weights as input to the fast CBC algorithm. This produces a generating vector $\boldsymbol{z} \in \mathbb{N}^s$. The generating vector is designed to be used to compute the estimate

$$\overline{Q}_{n,s,R}G(u_{s,h}) := \frac{1}{R}\sum_{r=0}^{R-1} Q_{n,s}^{\boldsymbol{\Delta}_r} G(u_{s,h}),$$

where $Q_{n,s}^{\boldsymbol{\Delta}_r}F := \frac{1}{n}\sum_{i=0}^{n-1} f(\{\boldsymbol{t}_i + \boldsymbol{\Delta}_r\} - \frac{\mathbf{1}}{\mathbf{2}})$, $\boldsymbol{t}_k := \{\frac{k\boldsymbol{z}}{n}\}$, and $\boldsymbol{\Delta}_0,\ldots,\boldsymbol{\Delta}_{R-1}$ are independent random shifts drawn from $\mathcal{U}([0,1]^s)$.

- Typically, the number of random shifts is taken to be rather small, e.g., $8 \leq R \leq 64$.
- A practical estimate for the R.M.S. error is given by the formula

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}|I_s F - Q_{n,s}^{\boldsymbol{\Delta}} F|^2} \approx \sqrt{\frac{1}{R(R-1)} \sum_{r=0}^{R-1} (Q_{n,s}^{\boldsymbol{\Delta}_r} F - \overline{Q}_{n,s,R} F)^2}.$$

- For the computation of the variance, note that

$$\mathrm{Var}[G(u_{s,h})] = \mathbb{E}[G(u_{s,h})^2] - \mathbb{E}[G(u_{s,h})]^2.$$

  We already know how to approximate $\mathbb{E}[G(u_{s,h})]$ using QMC, but the weights need to be updated if we wish to construct a QMC rule with a dimension-independent convergence rate for $\mathbb{E}[G(u_{s,h})^2]$ (exercise).
- If a QMC rule converges independently of s for the approximation of $\mathbb{E}[G(u_{s,h})^2]$, then the same rule will have dimension-independent convergence for $\mathbb{E}[G(u_{s,h})]$ as well.
- If we instead wish to estimate $\mathbb{E}[u_{s,h}(\boldsymbol{x},\cdot)]$ or $\mathrm{Var}[u_{s,h}(\boldsymbol{x},\cdot)]$ (i.e., leave out the *quantity of interest* $G\colon H_0^1(D) \to \mathbb{R}$), the same weights can be used as input to the CBC algorithm (but we still need to prove this).