

# Uncertainty Quantification and Quasi-Monte Carlo

Wintersemester 2022/23

---

Vesa Kaarnioja  
vesa.kaarnioja@fu-berlin.de

FU Berlin, FB Mathematik und Informatik

Twelfth lecture, January 16, 2023  
Thirteenth lecture, January 23, 2023

We continue studying 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  $\mathbf{y} \in U$ ,  $u(\cdot, \mathbf{y}) \in H_0^1(D)$  such that

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

where the diffusion coefficient has the parameterization

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

where we assume

- (A1)  $a_0 \in L^\infty(D)$  and  $\psi_j \in L^\infty(D)$  for all  $j \in \mathbb{N}$ ,
  - (A2) there exist  $a_{\min}, a_{\max} > 0$  s.t.  $0 < a_{\min} \leq a(\mathbf{x}, \mathbf{y}) \leq a_{\max} < \infty$  for all  $\mathbf{x} \in D$  and  $\mathbf{y} \in U$ ,
  - (A3)  $\sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty(D)}^p < \infty$  for some  $p \in (0, 1)$ .
- (Note that (A3) implies that  $\sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty(D)} < \infty$ .)

Let  $u_s(\cdot, \mathbf{y}) := u_s(\cdot, (y_1, \dots, y_s, 0, 0, \dots))$  denote the dimensionally-truncated PDE solution for  $\mathbf{y} \in U$  (we sometimes also write  $u_s(\cdot, \mathbf{y})$  for  $\mathbf{y} \in [-1/2, 1/2]^s$ ), and let  $u_{s,h}(\cdot, \mathbf{y}) \in V_h$  denote the dimensionally-truncated FE solution in the FE space spanned by piecewise linear FE basis functions. Let  $G: H_0^1(D) \rightarrow \mathbb{R}$  be a bounded linear functional.

During the last lecture, we split the overall approximation error as

$$\begin{aligned}
 & \left| \int_{[-1/2, 1/2]^N} G(u(\cdot, \mathbf{y})) \, d\mathbf{y} - \frac{1}{n} \sum_{i=1}^n G(u_{s,h}(\cdot, \mathbf{t}_i)) \right| \\
 & \leq \left| \int_{[-1/2, 1/2]^N} (G(u(\cdot, \mathbf{y}) - u_s(\cdot, \mathbf{y}))) \, d\mathbf{y} \right| \quad (\text{dimension-truncation error}) \\
 & + \left| \int_{[-1/2, 1/2]^s} G(u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y})) \, d\mathbf{y} \right| \quad (\text{finite element error}) \\
 & + \left| \int_{[-1/2, 1/2]^s} G(u_{s,h}(\cdot, \mathbf{y})) \, d\mathbf{y} - \frac{1}{n} \sum_{i=1}^n G(u_{s,h}(\cdot, \mathbf{t}_i)) \right|, \quad (\text{cubature error})
 \end{aligned}$$

and found that it is possible to construct a QMC point set  $\mathbf{t}_i := \left\{ \frac{i\mathbf{z}}{n} \right\}$  satisfying the QMC cubature error rate  $\mathcal{O}(\varphi(n)^{\max\{-1/p+1/2, -1+\delta\}})$ , where the implied coefficient is independent of  $s$ ,  $n$ , and  $h$ , and  $\delta \in (0, 1/2)$  is arbitrary. Let us consider the other error contributions next.

Some auxiliary results

## Neumann series: “Sufficiently small perturbations of the identity are still invertible”

We will require the following well-known generalization of the geometric series formula, named after 19<sup>th</sup> century mathematician Carl Neumann.

### Theorem (Neumann series)

Let  $H$  be a Hilbert space and let  $A \in \mathcal{L}(H)$  be a bounded linear functional with operator norm  $\|A\| < 1$ . Then  $I - A$  is invertible in  $\mathcal{L}(H)$  with

$$(I - A)^{-1} = I + A + \cdots + A^n + \cdots = \sum_{k=0}^{\infty} A^k,$$

and this series converges in operator norm.

*Proof.* Let  $B_{m,n} := \sum_{k=m}^n A^k$ ,  $m < n$ . Since  $\|A\| < 1$ , we have

$$\|B_{m,n}\| \leq \sum_{k=m}^n \|A\|^k = \|A\|^m \sum_{k=0}^{n-m} \|A\|^k = \|A\|^m \frac{1 - \|A\|^{n-m+1}}{1 - \|A\|} \xrightarrow{m,n \rightarrow \infty} 0.$$

$\therefore$  The partial sums  $\sum_{k=0}^n A^k$  form a Cauchy sequence in  $\mathcal{L}(H)$ .

Since  $H$  is a Hilbert space,  $\mathcal{L}(H)$  is a Banach space and the limit

$$B := \lim_{n \rightarrow \infty} \sum_{k=0}^n A^k \in \mathcal{L}(H)$$

exists. We need to prove that  $(I - A)B = I = B(I - A)$ . Let

$$B_n := I + A + \cdots + A^n.$$

Then

$$\begin{aligned}(I - A)B_n &= I - A^{n+1}, \\ B_n(I - A) &= I - A^{n+1},\end{aligned}$$

and since  $\|A\| < 1$ ,  $\|A^{n+1}\| \leq \|A\|^{n+1} \xrightarrow{n \rightarrow \infty} 0$ , we thus obtain

$$I - A^{n+1} \xrightarrow{n \rightarrow \infty} I \quad \text{in } \mathcal{L}(H)$$

and

$$(I - A)B = \lim_{n \rightarrow \infty} (I - A)B_n = I = \lim_{n \rightarrow \infty} B_n(I - A) = B(I - A). \quad \square$$

## Multinomial theorem

The multinomial theorem is a generalization of Newton's binomial formula. Using multi-index notation, it can be expressed as

$$(x_1 + \cdots + x_s)^k = \sum_{\substack{|\nu|=k \\ \nu \in \mathbb{N}_0^s}} \frac{k!}{\nu!} \mathbf{x}^\nu.$$

In fact, if  $\mathbf{x} := (x_j)_{j=1}^\infty \in \ell^1$ , then we have

$$\left( \sum_{j=1}^{\infty} x_j \right)^k = \sum_{\substack{|\nu|=k \\ \nu \in \mathcal{F}}} \frac{k!}{\nu!} \mathbf{x}^\nu$$

and we will later require the following special case:

$$\left( \sum_{j=s+1}^{\infty} x_j \right)^k = \sum_{\substack{|\nu|=k \\ \nu \in \mathcal{F} \\ \nu_j=0 \ \forall j \leq s}} \frac{k!}{\nu!} \mathbf{x}^\nu. \quad (1)$$

The following lemma frequently appears in the context of best  $N$ -term approximation.

### Lemma (Stechkin's lemma)

Let  $\Lambda$  be a countable index set, let  $0 < p \leq q < \infty$ , and let  $(a_\nu)_{\nu \in \Lambda}$  be a sequence. Let  $\emptyset \neq \Lambda_N \subset \Lambda$  be a set of indices containing the  $N$  largest terms of the sequence  $(a_\nu)_{\nu \in \Lambda}$ . Then

$$\left( \sum_{\nu \in \Lambda \setminus \Lambda_N} |a_\nu|^q \right)^{1/q} \leq N^{-r} \left( \sum_{\nu \in \Lambda} |a_\nu|^p \right)^{1/p}, \quad r = \frac{1}{p} - \frac{1}{q}.$$

*Proof.* WLOG, we can relabel the  $a$ -sequence so that  $(a_j)_{j \geq 1}$  is non-increasing, i.e.,  $a_{j+1} \leq a_j$  for all  $j \geq 1$ . We obtain

$$\begin{aligned} \left( \sum_{j=N+1}^{\infty} |a_j|^q \right)^{1/q} &= \left( \sum_{j=N+1}^{\infty} |a_j|^{q-p} |a_j|^p \right)^{1/q} \leq |a_N|^{1-p/q} \left( \sum_{j=N+1}^{\infty} |a_j|^p \right)^{1/q} \\ &\leq |a_N|^{1-p/q} \left( \sum_{j \geq 1}^{\infty} |a_j|^p \right)^{1/q}. \end{aligned}$$

The key is to bound  $|a_N|^{1-p/q}$  in terms of  $N$ .



Standard technique: the monotonicity of the  $a$ -sequence implies that

$$\begin{aligned} N|a_N|^p &= |a_N|^p + \cdots + |a_N|^p \leq |a_1|^p + \cdots + |a_N|^p \leq \sum_{j \geq 1} |a_j|^p \\ \Rightarrow |a_N|^p &\leq N^{-1} \sum_{j \geq 1} |a_j|^p. \end{aligned}$$

Hence

$$|a_N|^{1-p/q} = |a_N|^{pr} \leq N^{-r} \left( \sum_{j \geq 1} |a_j|^p \right)^r.$$

Plugging this into the inequality on the previous page yields

$$\begin{aligned} \left( \sum_{j=N+1}^{\infty} |a_j|^q \right)^{1/q} &\leq |a_N|^{1-p/q} \left( \sum_{j \geq 1} |a_j|^p \right)^{1/q} \leq N^{-r} \left( \sum_{j \geq 1} |a_j|^p \right)^{r+1/q} \\ &= N^{-r} \left( \sum_{j \geq 1} |a_j|^p \right)^{1/p}, \end{aligned}$$

where the final equality follows from the definition  $r = 1/p - 1/q$ . □

Dimension truncation error

## Remark about infinite-dimensional integrals

Recall that  $U := [-1/2, 1/2]^{\mathbb{N}}$ . We will be discussing infinite-dimensional Lebesgue integrals of the form

$$\int_U f(\mathbf{y}) \, d\mathbf{y},$$

where we have the infinite tensor product measure

$$d\mathbf{y} := \bigotimes_{j=1}^{\infty} dy_j.$$

The  $\sigma$ -algebra  $\mathcal{F}$  for  $d\mathbf{y}$  is generated by finite rectangles  $\prod_{j=1}^{\infty} S_j$ , where only a finite number of  $S_j$  are different from  $[-1/2, 1/2]$  and those that are different are contained in  $[-1/2, 1/2]$ . The resulting triplet  $(U, \mathcal{F}, d\mathbf{y})$  is a probability space.

For in-depth measure-theoretic considerations cf., e.g., “Measure Theory” by Halmos.

For the purposes of this course, we can regard infinite-dimensional integrals as limits of finite-dimensional integrals in the following sense:

$$\int_U f(\mathbf{y}) \, d\mathbf{y} = \lim_{s \rightarrow \infty} \int_{[-1/2, 1/2]^s} f(y_1, \dots, y_s, 0, 0, \dots) \, dy_1 \cdots dy_s. \quad (2)$$

The justification for this can be found, e.g., in “Infinite-dimensional integration and the multivariate decomposition method” by Kuo, Nuyens, Plaskota, Sloan, and Wasilkowski (J. Comput. Appl. Math., 2017). The result is stated below without proof. (Homework: verify that the conditions of the following theorem are valid for our PDE model problem.)

### Theorem (Kuo *et al.* 2017)

Let  $f: U \rightarrow \mathbb{R}$  be integrable w.r.t. the measure  $d\mathbf{y} := \bigotimes_{j=1}^{\infty} dy_j$  which satisfies

$$\lim_{s \rightarrow \infty} f(y_1, \dots, y_s, 0, 0, \dots) = f(\mathbf{y}) \quad \text{for a.e. } \mathbf{y} \in U,$$

$$|f(y_1, \dots, y_s, 0, 0, \dots)| \leq |g(\mathbf{y})| \quad \text{for a.e. } \mathbf{y} \in U$$

for some integrable function  $g: U \rightarrow \mathbb{R}$  w.r.t. the measure  $d\mathbf{y}$ . Then the characterization (2) holds.

The following rate was proved in “Dimension truncation in QMC for affine-parametric operator equations” by Gantner (MCQMC 2016).

### Theorem (Dimension truncation error)

*Suppose that the assumptions (A1)–(A3) hold and  $\|\psi_1\|_{L^\infty(D)} \geq \|\psi_2\|_{L^\infty(D)} \geq \|\psi_3\|_{L^\infty(D)} \geq \dots$ . Then for every  $f \in L^2(D)$  and every bounded linear functional  $G: H_0^1(D) \rightarrow \mathbb{R}$ , there holds*

$$\left| \int_U G(u(\cdot, \mathbf{y}) - u_s(\cdot, \mathbf{y})) \, d\mathbf{y} \right| \leq C \frac{\|f\|_{L^2(D)} \|G\|_{H_0^1(D) \rightarrow \mathbb{R}}}{a_{\min}} s^{-\frac{2}{p}+1},$$

*where the constant  $C > 0$  is independent of  $s$ ,  $f$ , and  $G$ .*

## Intermezzo

The dimension truncation proof is based on recasting the variational formulation as an affine-parametric operator equation. Specifically, if  $u(\cdot, \mathbf{y})$  denotes the parametric PDE solution and  $f$  the source term, we require for the analysis the (linear) *forward operator*

$$A(\mathbf{y}): u(\cdot, \mathbf{y}) \mapsto f$$

and the *solution operator*

$$A(\mathbf{y})^{-1}: f \mapsto u(\cdot, \mathbf{y}).$$

To this end, we need to be careful with the function space setting (the domains and codomains of  $A(\mathbf{y})$  and  $A(\mathbf{y})^{-1}$ ).

First of all, let us denote the dual space of  $H_0^1(D)$  as

$$H^{-1}(D) := (H_0^1(D))' := \{F: H_0^1(D) \rightarrow \mathbb{R} \mid F \text{ is linear and bounded}\}.$$

(This is a Hilbert space as a consequence of Riesz representation theorem.)

Let  $F \in H^{-1}(D)$  and  $v \in H_0^1(D)$ . Then the *duality pairing* of  $F$  and  $v$  is defined as

$$\langle F, v \rangle_{H^{-1}(D), H_0^1(D)} := F(v).$$

In a certain sense, the element  $F \in H^{-1}(D)$  is defined by its *action* on the elements of  $H_0^1(D)$ . For example, fix some  $f \in L^2(D)$ . Then (weighted) integration over (parts of) the domain  $D$ , e.g.,

$$\langle F, v \rangle_{H^{-1}(D), H_0^1(D)} := \int_D f(\mathbf{x})v(\mathbf{x}) \, d\mathbf{x},$$

would be an example of an element of  $H^{-1}(D)$ .

Let  $\mathbf{y} \in U$  and consider the bilinear form

$$B_{\mathbf{y}}(v, w) = \int_D a(\mathbf{x}, \mathbf{y}) \nabla v(\mathbf{x}) \cdot \nabla w(\mathbf{x}) \, d\mathbf{x}, \quad v, w \in H_0^1(D).$$

Now

$$B_{\mathbf{y}}(v, w) \leq a_{\max} \|v\|_{H_0^1(D)} \|w\|_{H_0^1(D)}, \quad v, w \in H_0^1(D), \quad (\text{boundedness})$$

$$|B_{\mathbf{y}}(v, v)| \geq a_{\min} \|v\|_{H_0^1(D)}^2, \quad v \in H_0^1(D). \quad (\text{coercivity})$$

Then the Lax–Milgram lemma implies that for any  $F \in H^{-1}(D)$ , there exists a unique element  $u(\cdot, \mathbf{y}) \in H_0^1(D)$  such that

$$B_{\mathbf{y}}(u(\cdot, \mathbf{y}), v) = F(v) \quad \text{for all } v \in H_0^1(D)$$

and

$$\|u(\cdot, \mathbf{y})\|_{H_0^1(D)} \leq \frac{\|F\|_{H^{-1}(D)}}{a_{\min}}.$$

Especially, the linear map  $A(\mathbf{y}): H_0^1(D) \rightarrow H^{-1}(D)$ ,  $u(\mathbf{y}) \mapsto F$ , is boundedly invertible<sup>†</sup> with

$$\|A(\mathbf{y})\|_{H_0^1(D) \rightarrow H^{-1}(D)} \leq a_{\max} \quad \text{and} \quad \|A(\mathbf{y})^{-1}\|_{H^{-1}(D) \rightarrow H_0^1(D)} \leq \frac{1}{a_{\min}}.$$

---

<sup>†</sup>Not trivial! See, e.g., Remark 2.7 in “Theory and Practice of Finite Elements” by Ern and Guermond.



*Proof (dimension truncation).* Let us introduce the operators  $A(\mathbf{y}), A^s(\mathbf{y}): H_0^1(D) \rightarrow H^{-1}(D)$ ,

$$A(\mathbf{y}) := B_0 + \sum_{j=1}^{\infty} y_j B_j \quad \text{and} \quad A^s(\mathbf{y}) := B_0 + \sum_{j=1}^s y_j B_j,$$

where  $B_j: H_0^1(D) \rightarrow H^{-1}(D)$  are defined by setting

$$\begin{aligned} \langle B_0 v, w \rangle_{H^{-1}(D), H_0^1(D)} &:= \langle a_0 \nabla v, \nabla w \rangle_{L^2(D)}, \\ \langle B_j v, w \rangle_{H^{-1}(D), H_0^1(D)} &:= \langle \psi_j \nabla v, \nabla w \rangle_{L^2(D)} \quad \text{for } j \geq 1. \end{aligned}$$

The variational problem

$$\int_D a(\mathbf{x}, \mathbf{y}) \nabla u(\mathbf{x}, \mathbf{y}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x} = \langle F, v \rangle_{H^{-1}(D), H_0^1(D)} \quad \text{for all } v \in H_0^1(D),$$

$$a(\mathbf{x}, \mathbf{y}) = a_0(\mathbf{x}) + \sum_{j=1}^{\infty} y_j \psi_j(\mathbf{x}),$$

where  $F \in H^{-1}(D)$ , can be expressed as an affine-parametric parametric operator equation

$$A(\mathbf{y})u(\cdot, \mathbf{y}) = F.$$

Our assumptions (A1)–(A3) ensure that both  $A(\mathbf{y})$  and  $A^s(\mathbf{y})$  are

Suppose that  $1 \leq s < s'$ . As a consequence of the *a priori* bound for the PDE, we have

$$\begin{aligned} \int_D G(u(\mathbf{y}) - u_s(\mathbf{y})) \, d\mathbf{y} &\leq \frac{2\|G\|_{H^{-1}(D)}\|F\|_{H^{-1}(D)}}{a_{\min}} \\ &= \frac{2\|G\|_{H^{-1}(D)}\|F\|_{H^{-1}(D)}}{a_{\min}} \frac{s^{-2/p+1}}{s^{-2/p+1}} \leq \frac{2\|G\|_{H^{-1}(D)}\|F\|_{H^{-1}(D)}}{a_{\min}} \frac{s^{-2/p+1}}{(s')^{-2/p+1}}. \end{aligned}$$

Thus it is sufficient to prove the claim for  $s \geq s'$  with  $s'$  large enough. To this end, we assume that  $s \geq s'$  where  $s'$  is chosen to be large enough such that

$$\sum_{j=s+1}^{\infty} b_j \leq \frac{1}{2} \quad \text{for all } s \geq s'. \quad (3)$$

For future reference, note that (3) also implies for all  $s \geq s'$  that

$$b_j \leq \frac{1}{2} \quad \text{for all } j \geq s+1 \quad \text{and} \quad \sum_{j=s+1}^{\infty} b_j^2 \leq \sum_{j=s+1}^{\infty} b_j \leq \frac{1}{2}. \quad (4)$$

We also have for all  $\mathbf{y} \in U$  that

$$\|A(\mathbf{y})\|_{H_0^1(D) \rightarrow H^{-1}(D)} \leq a_{\max}, \quad \|A^s(\mathbf{y})\|_{H_0^1(D) \rightarrow H^{-1}(D)} \leq a_{\max}$$
$$\|A(\mathbf{y})^{-1}\|_{H^{-1}(D) \rightarrow H_0^1(D)} \leq \frac{1}{a_{\min}}, \quad \|A^s(\mathbf{y})^{-1}\|_{H^{-1}(D) \rightarrow H_0^1(D)} \leq \frac{1}{a_{\min}}.$$

For brevity, let us denote

$$u(\mathbf{y}) := u(\cdot, \mathbf{y}), \quad \mathbf{y} \in U,$$
$$u_s(\mathbf{y}) := u_s(\cdot, \mathbf{y}), \quad \mathbf{y} \in U.$$

Now  $u(\mathbf{y}) = A(\mathbf{y})^{-1}F$ ,  $u_s(\mathbf{y}) = A^s(\mathbf{y})^{-1}F$ , and we can write

$$A(\mathbf{y}) - A^s(\mathbf{y}) = \sum_{j=s+1}^{\infty} y_j B_j, \quad \mathbf{y} \in U, \quad s \in \mathbb{N}.$$

Let  $w \in H_0^1(D)$ . Then

$$\begin{aligned} \|A^s(\mathbf{y})^{-1}B_j w\|_{H_0^1(D)} &\leq \frac{\|B_j w\|_{H^{-1}(D)}}{a_{\min}} \\ &= \frac{1}{a_{\min}} \sup_{v \in H_0^1(D) \setminus \{0\}} \frac{\langle \psi_j \nabla w, \nabla v \rangle_{L^2(D)}}{\|v\|_{H_0^1(D)}} \leq b_j \|w\|_{H_0^1(D)}, \end{aligned}$$

where the sequence  $\mathbf{b} = (b_j)_{j \geq 1}$  is defined as  $b_j := \frac{\|\psi_j\|_{L^\infty(D)}}{a_{\min}}$ . In consequence,

$$\sup_{\mathbf{y} \in U} \|A^s(\mathbf{y})^{-1}B_j\|_{\mathcal{L}(H_0^1(D))} \leq b_j,$$

$$\sup_{\mathbf{y} \in U} \|A^s(\mathbf{y})^{-1}(A(\mathbf{y}) - A^s(\mathbf{y}))\|_{\mathcal{L}(H_0^1(D))} \leq \sum_{j=s+1}^{\infty} b_j \stackrel{(3)}{\leq} \frac{1}{2} < 1.$$

It follows from the previous discussion and the assumption  $s \geq s'$  that the Neumann series

$$\begin{aligned} A(\mathbf{y})^{-1} &= (I + A^s(\mathbf{y})^{-1}(A(\mathbf{y}) - A^s(\mathbf{y})))^{-1} A^s(\mathbf{y})^{-1} \\ &= \sum_{k=0}^{\infty} (-A^s(\mathbf{y})^{-1}(A(\mathbf{y}) - A^s(\mathbf{y})))^k A^s(\mathbf{y})^{-1} \end{aligned}$$

is well-defined. Moreover, we have the representation

$$\begin{aligned} \int_U G(u(\mathbf{y}) - u_s(\mathbf{y})) \, d\mathbf{y} &= \int_U G((A(\mathbf{y})^{-1} - A^s(\mathbf{y})^{-1})f) \, d\mathbf{y} \\ &= \sum_{k=1}^{\infty} \int_U G((-A^s(\mathbf{y})^{-1}(A(\mathbf{y}) - A^s(\mathbf{y})))^k u_s(\mathbf{y})) \, d\mathbf{y} \\ &= \sum_{k=1}^{\infty} (-1)^k \int_U G\left(\left(\sum_{j=s+1}^{\infty} y_j A^s(\mathbf{y})^{-1} B_j\right)^k u_s(\mathbf{y})\right) \, d\mathbf{y}. \end{aligned}$$

The integrand can be expanded as

$$\left( \sum_{j=s+1}^{\infty} y_j A^s(\mathbf{y})^{-1} B_j \right)^k = \sum_{\eta_1, \dots, \eta_k = s+1}^{\infty} \left( \prod_{i=1}^k y_{\eta_i} \right) \left( \prod_{i=1}^k A^s(\mathbf{y})^{-1} B_{\eta_i} \right),$$

where the second product symbol is assumed to respect the order of the noncommutative operators. By Fubini's theorem, we obtain

$$\begin{aligned} & \int_U G \left( \left( \sum_{j=s+1}^{\infty} y_j A^s(\mathbf{y})^{-1} B_j \right)^k u_s(\mathbf{y}) \right) d\mathbf{y} \\ &= \sum_{\eta_1, \dots, \eta_k = s+1}^{\infty} \underbrace{\left( \int_U \prod_{i=1}^k y_{\eta_i} d\mathbf{y} \right)}_{=: l_1} \underbrace{\left( \int_{U_s} G \left( \left( \prod_{i=1}^k A^s(\mathbf{y})^{-1} B_{\eta_i} \right) u_s(\mathbf{y}) \right) d\mathbf{y}_{\{1:s\}} \right)}_{=: l_2}. \end{aligned}$$

- $l_1 \geq 0$  can be written as a product of univariate integrals of the form  $0 \leq \int_{-1/2}^{1/2} y_j^m dy_j \leq 1$ ,  $m \in \mathbb{N}$ . Note that this vanishes when  $m = 1$ .
- $|l_2| \leq \|G\|_{H^{-1}(D)} \left( \prod_{i=1}^k \sup_{\mathbf{y} \in U} \|A^s(\mathbf{y})^{-1} B_{\eta_i}\| \right) \|u_s(\mathbf{y})\|_{H_0^1(D)} \leq \frac{\|G\|_{H^{-1}(D)} \|F\|_{H^{-1}(D)}}{a_{\min}} \left( \prod_{i=1}^k b_{\eta_i} \right).$

Earlier we arrived at

$$\int_U G(u(\mathbf{y}) - u_s(\mathbf{y})) d\mathbf{y} = \sum_{k=1}^{\infty} (-1)^k \int_U G\left(\left(\sum_{j=s+1}^{\infty} y_j A^s(\mathbf{y})^{-1} B_j\right)^k u_s(\mathbf{y})\right) d\mathbf{y}$$

We can estimate the summands as

$$\begin{aligned} & \left| (-1)^k \int_U G\left(\left(\sum_{j=s+1}^{\infty} y_j A^s(\mathbf{y})^{-1} B_j\right)^k u_s(\mathbf{y})\right) d\mathbf{y} \right| \\ & \leq \frac{\|G\|_{H^{-1}(D)} \|F\|_{H^{-1}(D)}}{a_{\min}} \sum_{\eta_1, \dots, \eta_k = s+1}^{\infty} \left( \int_U \prod_{k=1}^k y_{\eta_i} d\mathbf{y} \right) \left( \prod_{i=1}^k b_{\eta_i} \right) \\ & = \frac{\|G\|_{H^{-1}(D)} \|F\|_{H^{-1}(D)}}{a_{\min}} \int_U \sum_{\eta_1, \dots, \eta_k = s+1}^{\infty} \left( \prod_{k=1}^k y_{\eta_i} \right) \left( \prod_{i=1}^k b_{\eta_i} \right) d\mathbf{y} \\ & = \frac{\|G\|_{H^{-1}(D)} \|F\|_{H^{-1}(D)}}{a_{\min}} \int_U \left( \sum_{j=s+1}^{\infty} y_j b_j \right)^k d\mathbf{y} \\ & \stackrel{(1)}{=} \frac{\|G\|_{H^{-1}(D)} \|F\|_{H^{-1}(D)}}{a_{\min}} \int_U \sum_{\substack{|\nu|=k \\ \nu_j=0 \forall j \leq s}} \frac{k!}{\nu!} \left( \prod_{j=s+1}^{\infty} y_j^{\nu_j} \right) \left( \prod_{j=s+1}^{\infty} b_j^{\nu_j} \right) d\mathbf{y}. \end{aligned}$$

The integrals vanish whenever  $\nu$  contains an element equal to 1, hence

$$\begin{aligned} & \left| (-1)^k \int_U G \left( \left( \sum_{j=s+1}^{\infty} y_j A^s(\mathbf{y})^{-1} B_j \right)^k u_s(\mathbf{y}) \right) d\mathbf{y} \right| \\ & \leq \frac{\|G\|_{H^{-1}(D)} \|F\|_{H^{-1}(D)}}{a_{\min}} \sum_{\substack{|\nu|=k \\ \nu_j=0 \ \forall j \leq s \\ \nu_j \neq 1 \ \forall j > s}} \frac{k!}{\nu!} \mathbf{b}^\nu. \end{aligned}$$

We arrive at (note that the summand corresponding to  $k = 1$  vanishes!)

$$\begin{aligned} & \left| \int_U G(u(\mathbf{y}) - u_s(\mathbf{y})) d\mathbf{y} \right| \leq \frac{\|G\|_{H^{-1}(D)} \|F\|_{H^{-1}(D)}}{a_{\min}} \sum_{k=1}^{\infty} \sum_{\substack{|\nu|=k \\ \nu_j=0 \ \forall j \leq s \\ \nu_j \neq 1 \ \forall j > s}} \frac{k!}{\nu!} \mathbf{b}^\nu \\ & = \frac{\|G\|_{H^{-1}(D)} \|F\|_{H^{-1}(D)}}{a_{\min}} \left[ \sum_{k=k'}^{\infty} \sum_{\substack{|\nu|=k \\ \nu_j=0 \ \forall j \leq s \\ \nu_j \neq 1 \ \forall j > s}} \frac{k!}{\nu!} \mathbf{b}^\nu + \sum_{k=2}^{k'-1} \sum_{\substack{|\nu|=k \\ \nu_j=0 \ \forall j \leq s \\ \nu_j \neq 1 \ \forall j > s}} \frac{k!}{\nu!} \mathbf{b}^\nu \right], \end{aligned}$$

where we split the sum into two w.r.t.  $k' \geq 3$  to be specified later.



The sum over  $k \geq k'$  can be bounded using the geometric series as

$$\begin{aligned} \sum_{k=k'}^{\infty} \sum_{\substack{|\nu|=k \\ \nu_j=0 \quad \forall j \leq s \\ \nu_j \neq 1 \quad \forall j > s}} \frac{k!}{\nu!} \mathbf{b}^\nu &\leq \sum_{k=k'}^{\infty} \left( \sum_{j=s+1}^{\infty} b_j \right)^k \\ &\leq \left( \sum_{j=s+1}^{\infty} b_j \right)^{k'} \frac{1}{1 - \sum_{j=s+1}^{\infty} b_j} \leq C_S^{k'(-1/p+1)}, \end{aligned}$$

where Stechkin's lemma yields  $\sum_{j=s+1}^{\infty} b_j \leq \left( \sum_{j=1}^{\infty} b_j^p \right)^{1/p} s^{-1/p+1}$  and the resulting constant  $C_1 := 2 \left( \sum_{j=1}^{\infty} b_j^p \right)^{k'/p}$  is independent of  $s$ ,  $f$ , and  $G$ .

On the other hand, for the sum over  $2 \leq k < k'$ , we estimate

$$\sum_{k=2}^{k'-1} \sum_{\substack{|\nu|=k \\ \nu_j=0 \forall j \leq s \\ \nu_j \neq 1 \forall j > s}} \frac{k!}{\nu!} \mathbf{b}^\nu \leq (k'-1)! \sum_{k=2}^{k'-1} \sum_{\substack{|\nu|=k \\ \nu_j=0 \forall j \leq s \\ \nu_j \neq 1 \forall j > s}} \mathbf{b}^\nu.$$

For each  $2 \leq k < k'$ , we obtain

$$\begin{aligned} \sum_{\substack{|\nu|=k \\ \nu_j=0 \forall j \leq s \\ \nu_j \neq 1 \forall j > s}} \mathbf{b}^\nu &\leq \sum_{\substack{0 \neq |\nu|_\infty \leq k \\ \nu_j=0 \forall j \leq s \\ \nu_j \neq 1 \forall j > s}} \mathbf{b}^\nu = \prod_{j=s+1}^{\infty} \left( 1 + \sum_{\ell=2}^k b_j^\ell \right) - 1 \\ &= \prod_{j=s+1}^{\infty} \left( 1 + b_j^2 \frac{1 - b_j^{j-1}}{1 - b_j} \right) - 1 \leq \prod_{j=s+1}^{\infty} (1 + 2b_j^2) - 1 \\ &\leq \exp \left( 2 \sum_{j=s+1}^{\infty} b_j^2 \right) - 1 \leq C_2 s^{-2/p+1}, \quad C_2 := 2(e-1) \left( \sum_{j=1}^{\infty} b_j^p \right)^{1/p}, \end{aligned}$$

where we used  $e^x \leq 1 + (e-1)x$  for  $x \in [0, 1]$  and Stechkin's lemma  $\sum_{j=s+1}^{\infty} b_j^2 \leq \left( \sum_{j=1}^{\infty} b_j^p \right)^{1/p} s^{-2/p+1}$ .  $C_2$  is independent of  $s$ ,  $f$ , and  $G$ .

Putting everything together, we conclude that

$$\left| \int_U G(u(\mathbf{y}) - u_s(\mathbf{y})) \, d\mathbf{y} \right| \leq \frac{\|G\|_{H^{-1}(D)} \|F\|_{H^{-1}(D)}}{a_{\min}} (C_1 s^{k'(-1/p+1)} + k'!(k' - 2) C_2 s^{-2/p+1}).$$

The two terms can be balanced by choosing  $k' := \lceil (2 - p)/(1 - p) \rceil$ , where  $\lceil x \rceil := \min\{k \in \mathbb{Z} \mid k \geq x\}$  is the ceiling function. (Note that  $k' \geq 3$  for all  $p \in (0, 1)$ .)

Since we already know that the result holds for all  $s \leq s'$ , the assertion for all  $s \geq 1$  follows by a trivial adjustment of the constant factors.

Finally, if the source term  $f \in L^2(D)$ , we can associate it with an element  $F \in H^{-1}(D)$  defined by

$$\langle F, v \rangle_{H^{-1}(D), H_0^1(D)} := \int_D f(\mathbf{x}) v(\mathbf{x}) \, d\mathbf{x}, \quad v \in H_0^1(D).$$

Especially,  $\|F\|_{H^{-1}(D)} \leq C_P \|f\|_{L^2(D)}$ , where  $C_P > 0$  is the Poincaré constant. □

Finite element error

Suppose that  $D \subset \mathbb{R}^d$ ,  $d \in \{2, 3\}$ , is a bounded, convex polyhedral domain.

Let  $\{V_h\}_h$  be a family of finite element subspaces of  $H_0^1(D)$ , indexed by the mesh size  $h > 0$  and spanned by continuous, piecewise linear finite element basis functions over a sequence of regular, simplicial meshes in  $D$  obtained from an initial, regular triangulation of  $D$  by recursive, uniform bisection of simplices.

In this setup, it is known (cf., e.g., Gilbarg and Trudinger) that for functions  $v \in H_0^1(D) \cap H^2(D)$ , there exists a constant  $C_1 > 0$  such that

$$\inf_{v_h \in V_h} \|v - v_h\|_{H_0^1(D)} \leq C_1 h \|v\|_{H_0^1(D) \cap H^2(D)} \quad \text{as } h \rightarrow 0, \quad (5)$$

where  $\|v\|_{H_0^1(D) \cap H^2(D)} := (\|v\|_{L^2(D)}^2 + \|\Delta v\|_{L^2(D)}^2)^{1/2}$ .

Note that we need higher  $H^2(D)$  regularity of the PDE solution in order to derive the asymptotic convergence rate as  $h \rightarrow \infty$ . This can be ensured, e.g., when the diffusion coefficient is Lipschitz,  $f \in L^2(D)$ , and the domain  $D$  is a bounded, convex polyhedron.

## Proposition (Elliptic regularity)

Suppose that  $a_0 \in W^{1,\infty}(D)$  and  $\psi_j \in W^{1,\infty}(D)$  for all  $j \geq 1$  such that  $C_\psi := \sum_{j \geq 1} \|\psi_j\|_{W^{1,\infty}(D)} < \infty$ , where

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

Then there exists a constant  $C_2 > 0$  independent of  $\mathbf{y}$  and  $f$  such that the solution  $u(\cdot, \mathbf{y}) \in H_0^1(D)$  of the parametric PDE problem satisfies

$$\|u(\cdot, \mathbf{y})\|_{H_0^1(D) \cap H^2(D)} \leq C_2 \|f\|_{L^2(D)} \quad \text{for all } \mathbf{y} \in U. \quad (6)$$

*Proof (sketch).* Standard ellipticity theory implies that  $u(\cdot, \mathbf{y}) \in H_0^1(D)$  is such that  $\exists \Delta u(\cdot, \mathbf{y}) \in L^2(D)$  for all  $\mathbf{y} \in U$ . Since now  $\|a(\cdot, \mathbf{y})\|_{W^{1,\infty}(D)} < \infty$  for all  $\mathbf{y} \in U$ , we obtain

$$-\nabla \cdot (a(\mathbf{x}, \mathbf{y}) \nabla u(\mathbf{x}, \mathbf{y})) = f(\mathbf{x}) \quad (\nabla \cdot (\psi \nabla \varphi) = \nabla \psi \cdot \nabla \varphi + \psi \Delta \varphi)$$

$$\Leftrightarrow -a(\mathbf{x}, \mathbf{y}) \Delta u(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \nabla a(\mathbf{x}, \mathbf{y}) \cdot \nabla u(\mathbf{x}, \mathbf{y})$$

$$\Rightarrow \|\Delta u(\cdot, \mathbf{y})\|_{L^2(D)} \leq \frac{\|f\|_{L^2(D)}}{a_{\min}} + \frac{\|\nabla a(\cdot, \mathbf{y})\|_{L^\infty(D)}}{a_{\min}} \|u(\cdot, \mathbf{y})\|_{H_0^1(D)}$$

$$\leq \frac{\|f\|_{L^2(D)}}{a_{\min}} + \frac{\|a_0\|_{W^{1,\infty}(D)} + C_\psi}{a_{\min}} \frac{C_P \|f\|_{L^2(D)}}{a_{\min}} =: C_2 \|f\|_{L^2(D)}. \quad \square$$

## Dimensionally-truncated finite element solution

Let  $a_s(\mathbf{x}, \mathbf{y}) := a(\mathbf{x}, (y_1, \dots, y_s, 0, 0, \dots))$  for  $\mathbf{y} \in U$ . For  $\mathbf{y} \in U$ ,  $u_{s,h}(\cdot, \mathbf{y}) \in V_h$  is the dimensionally-truncated finite element solution if

$$\int_D a_s(\mathbf{x}, \mathbf{y}) \nabla u_{s,h}(\mathbf{x}, \mathbf{y}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x} = \int_D f(\mathbf{x}) v(\mathbf{x}) \, d\mathbf{x} \quad \text{for all } v \in V_h.$$

## Finite element error in $H_0^1(D)$

Recall that by Céa's lemma, the finite element solution is a *quasi-optimal* approximation in the following sense:

$$\|u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y})\|_{H_0^1(D)} \leq C(\mathbf{y}) \inf_{v_h \in V_h} \|u_s(\cdot, \mathbf{y}) - v_h\|_{H_0^1(D)},$$

where the constant  $C(\mathbf{y}) := \frac{\sup_{\mathbf{x} \in D} a(\mathbf{x}, \mathbf{y})}{\inf_{\mathbf{x} \in D} a(\mathbf{x}, \mathbf{y})} \leq \frac{a_{\max}}{a_{\min}} =: C_3 < \infty$  can be bounded independently of  $\mathbf{y} \in U$  due to our uniform ellipticity assumption. Combining this with the approximation property (5) and the elliptic regularity shift (6) yields

$$\begin{aligned} \|u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y})\|_{H_0^1(D)} &\leq C_3 \inf_{v_h \in V_h} \|u_s(\cdot, \mathbf{y}) - v_h\|_{H_0^1(D)} \\ &\stackrel{(5)}{\leq} C_3 C_1 h \|u_s(\cdot, \mathbf{y})\|_{H^2(D) \cap H_0^1(D)} \\ &\stackrel{(6)}{\leq} C_3 C_1 C_2 h \|f\|_{L^2(D)} \quad \text{as } h \rightarrow 0. \end{aligned} \quad (7)$$

However, if we measure the error in the  $L^2(D)$  norm, the finite element convergence rate can be improved by an order of magnitude.



## Finite element error in $L^2(D)$

### Proposition

*Under the same assumptions as the previous proposition, there exists a constant  $C > 0$  independent of  $s$ ,  $h$ ,  $f$ , and  $\mathbf{y}$  such that*

$$\|u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y})\|_{L^2(D)} \leq Ch^2 \|f\|_{L^2(D)} \quad \text{as } h \rightarrow 0.$$

*Proof.* Let  $g \in L^2(D)$ . For  $\mathbf{y} \in U$ , let  $u_{g,s}(\cdot, \mathbf{y}) \in H_0^1(D)$  denote the solution to

$$\int_D a_s(\mathbf{x}, \mathbf{y}) \nabla u_{g,s}(\cdot, \mathbf{y}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x} = \int_D g(\mathbf{x}) v(\mathbf{x}) \, d\mathbf{x} \quad \text{for all } v \in H_0^1(D),$$

where  $a_s(\cdot, \mathbf{y}) := a(\cdot, (y_1, \dots, y_s, 0, 0, \dots))$ . We test this against  $v = u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y})$  and let  $v_h \in V_h$  be arbitrary.

It follows from Galerkin orthogonality of the finite element solution that

$$\begin{aligned}
 & \langle g, u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y}) \rangle_{L^2(D)} \\
 &= \int_D a_s(\mathbf{x}, \mathbf{y}) \nabla u_{g,s}(\mathbf{x}, \mathbf{y}) \cdot \nabla (u_s(\mathbf{x}, \mathbf{y}) - u_{s,h}(\mathbf{x}, \mathbf{y})) \, d\mathbf{x} \\
 &= \int_D a_s(\mathbf{x}, \mathbf{y}) \nabla (u_{g,s}(\mathbf{x}, \mathbf{y}) - v_h(\mathbf{x})) \cdot \nabla (u_s(\mathbf{x}, \mathbf{y}) - u_{s,h}(\mathbf{x}, \mathbf{y})) \, d\mathbf{x} \\
 &\leq a_{\max} \|u_{g,s}(\cdot, \mathbf{y}) - v_h\|_{H_0^1(D)} \|u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y})\|_{H_0^1(D)}.
 \end{aligned}$$

In consequence,

$$\begin{aligned}
 & \langle g, u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y}) \rangle_{L^2(D)} \\
 & \leq a_{\max} \|u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y})\|_{H_0^1(D)} \inf_{v_h \in V_h} \|u_{g,s}(\cdot, \mathbf{y}) - v_h\|_{H_0^1(D)}, \quad (8)
 \end{aligned}$$

where  $g \in L^2(D)$  is arbitrary. We now use the *Aubin–Nitsche trick*: recall from the exercises of week 2(!) that the following identity holds

$$\|F\|_{L^2(D)} = \sup_{\substack{g \in L^2(D) \\ \|g\|_{L^2(D)} \leq 1}} \langle g, F \rangle_{L^2(D)} \quad \text{for all } F \in L^2(D).$$

We take the supremum over  $\{g \in L^2(D) : \|g\|_{L^2(D)} \leq 1\}$  in (8) to obtain...

$$\begin{aligned}
& \|u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y})\|_{L^2(D)} \\
&= \sup_{\substack{g \in L^2(D) \\ \|g\|_{L^2(D)} \leq 1}} \langle g, u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y}) \rangle_{L^2(D)} \\
&\leq a_{\max} \underbrace{\|u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y})\|_{H_0^1(D)}}_{\stackrel{(7)}{\leq} C_3 C_1 C_2 h \|f\|_{L^2(D)}} \sup_{\substack{g \in L^2(D) \\ \|g\|_{L^2(D)} \leq 1}} \left( \underbrace{\inf_{v_h \in V_h} \|u_{g,s}(\cdot, \mathbf{y}) - v_h\|_{H_0^1(D)}}_{\stackrel{(5)}{\leq} C_1 h \|u_{g,s}(\cdot, \mathbf{y})\|_{H_0^1(D) \cap H^2(D)}} \right) \\
&\hspace{15em} \stackrel{(6)}{\leq} C_1 C_2 h \|g\|_{L^2(D)} \\
&\leq Ch^2 \|f\|_{L^2(D)},
\end{aligned}$$

where the constant  $C := a_{\max}(C_1 C_2)^2 C_3$  is independent of  $s$ ,  $h$ ,  $f$ , and  $\mathbf{y}$ . □

Note especially that if  $G: L^2(D) \rightarrow \mathbb{R}$  is a bounded linear operator, then

$$\int_U |G(u_s(\cdot, \mathbf{y}) - u_{s,h}(\cdot, \mathbf{y}))| d\mathbf{y} \leq C \|G\|_{L^2(D) \rightarrow \mathbb{R}} \|f\|_{L^2(D)} h^2,$$

where  $C > 0$  is independent of  $s$ ,  $h$ , and  $f$ .

## Overall error

Let  $I(F) := \int_U F(\mathbf{y}) \, d\mathbf{y}$ .

### Theorem

Let  $D \subset \mathbb{R}^d$ ,  $d \in \{2, 3\}$ , be a bounded polyhedron, assume (A1)–(A3),  $\|\psi_1\|_{L^\infty(D)} \geq \|\psi_2\|_{L^\infty(D)} \geq \|\psi_3\|_{L^\infty(D)} \geq \dots$ , and suppose that  $a_0 \in W^{1,\infty}(D)$  and  $\psi_j \in W^{1,\infty}(D)$  with  $\sum_{j=1}^{\infty} \|\psi_j\|_{W^{1,\infty}(D)} < \infty$ . Let  $G: L^2(D) \rightarrow \mathbb{R}$  be a bounded linear functional and define  $b_j := \frac{\|\psi_j\|_{L^\infty(D)}}{a_{\min}}$ . Then using the CBC algorithm with the POD weights

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

as inputs to construct a randomly shifted rank-1 lattice rule

$Q_{n,s}^\Delta(F) := \sum_{k=0}^{n-1} F(\{\frac{kz}{n} + \Delta\} - \frac{1}{2})$ ,  $\Delta \in [0, 1]^s$ , we have the overall error

$$\sqrt{\mathbb{E}_\Delta |I(G(u)) - Q_{n,s}^\Delta(G(u_{s,h}))|^2} \leq C(\varphi(n)^{\max\{-1/p+1/2, -1+\delta\}} + s^{-2/p+1} + h^2),$$

where the constant  $C > 0$  is independent of  $s$ ,  $n$ , and  $h$ .

*Proof.* We have the total error decomposition<sup>†</sup>

$$\begin{aligned}\mathbb{E}_{\Delta}[|I(G(u)) - Q_{n,s}^{\Delta}(u_{s,h})|^2] &\leq 9|(I - I_s)(G(u))|^2 \\ &\quad + 9|I_s(G(u_s - u_{s,h}))|^2 \\ &\quad + 9\mathbb{E}_{\Delta}[|I_s(G(u_{s,h})) - Q_{n,s}^{\Delta}(G(u_{s,h}))|^2].\end{aligned}$$

We have already proved, under the stated assumptions, that there hold

$$\begin{aligned}|(I - I_s)(G(u))| &= \mathcal{O}(s^{-2/p+1}), \\ |I_s(G(u_s - u_{s,h}))| &= \mathcal{O}(h^2), \\ \mathbb{E}_{\Delta}[|I_s(G(u_{s,h})) - Q_{n,s}^{\Delta}(G(u_{s,h}))|^2] &= \mathcal{O}(n^{\max\{-1/p+1/2, -1+\delta\}}),\end{aligned}$$

from which the claim immediately follows. □

---

<sup>†</sup>Let  $a, b, c \geq 0$ . Then  
 $a + b + c \leq 3 \max\{a, b, c\} = 3\sqrt{\max\{a, b, c\}^2} \leq 3\sqrt{a^2 + b^2 + c^2}$ .

## Extension of QMC theory to the full PDE solution *without* a bounded linear quantity of interest $G$

Earlier, we discussed the QMC approximation for integrals of the form

$$\mathbb{E}[G(u_s)] = \int_{U_s} G(u_s(\cdot, \mathbf{y})) \, d\mathbf{y},$$

where  $G: H_0^1(D) \rightarrow \mathbb{R}$  (or  $G: L^2(D) \rightarrow \mathbb{R}$ ) is a bounded linear functional (quantity of interest).

But what if we wanted to approximate

$$\mathbb{E}[u_s(\mathbf{x}, \cdot)] = \int_{U_s} u_s(\mathbf{x}, \mathbf{y}) \, d\mathbf{y}$$

without a linear quantity of interest instead?

Idea: recall the variational characterization

$$\|f\|_{L^2(D)} = \sup_{\substack{G \in L^2(D) \\ \|G\|_{L^2(D)} \leq 1}} \langle G, f \rangle_{L^2(D)}$$

of the  $L^2$  norm from earlier.

By Fubini's theorem, we have that

$$\begin{aligned}
 \|I_s(u_s) - Q_{n,s}^\Delta(u_s)\|_{L^2(D)} &= \sup_{\substack{G \in L^2(D) \\ \|G\|_{L^2(D)} \leq 1}} |\langle G, I_s(u_s) - Q_{n,s}^\Delta(u_s) \rangle_{L^2(D)}| \\
 &= \sup_{\substack{G \in L^2(D) \\ \|G\|_{L^2(D)} \leq 1}} |I_s(\langle G, u_s \rangle_{L^2(D)}) - Q_{n,s}^\Delta(\langle G, u_s \rangle_{L^2(D)})| \\
 &\leq e_{n,s}(\mathbf{z}; \Delta) \sup_{\substack{G \in L^2(D) \\ \|G\|_{L^2(D)} \leq 1}} \|\langle G, u_s \rangle_{L^2(D)}\|_{s,\gamma},
 \end{aligned}$$

where  $e_{n,s}(\mathbf{z}; \Delta)$  denotes the worst-case error of the shifted lattice  $\{\mathbf{t}_i + \Delta : i \in \{1, \dots, n\}\}$ . Especially:

$$\sqrt{\mathbb{E}_\Delta \|I_s(u_s) - Q_{n,s}^\Delta(u_s)\|_{L^2(D)}^2} \leq e_{n,s}^{\text{sh}}(\mathbf{z}) \sup_{\substack{G \in L^2(D) \\ \|G\|_{L^2(D)} \leq 1}} \|\langle G, u_s \rangle_{L^2(D)}\|_{s,\gamma}.$$

The *shift-averaged worst-case error*  $e_{n,s}^{\text{sh}}(\mathbf{z})$  is precisely the same object that we have considered in the past, i.e.,

$$[e_{n,s}^{\text{sh}}(\mathbf{z})]^2 = \frac{1}{n} \sum_{\emptyset \neq u \subseteq \{1:s\}} \gamma_u \sum_{k=0}^{n-1} \prod_{j \in u} B_2\left(\left\{\frac{kz_j}{n}\right\}\right).$$

In summary, even in this setting, we have the CBC search criterion

$$[e_{n,s}^{\text{sh}}(\mathbf{z})]^2 = \frac{1}{n} \sum_{\emptyset \neq u \subseteq \{1:s\}} \gamma_u \sum_{k=0}^{n-1} \prod_{j \in u} B_2 \left( \left\{ \frac{kz_j}{n} \right\} \right).$$

The generating vector obtained using the CBC algorithm satisfies the estimate

$$\begin{aligned} \sqrt{\mathbb{E}_{\Delta} \|I_s(u_s) - Q_{n,s}^{\Delta}(u_s)\|_{L^2(D)}^2} &\leq \left( \frac{1}{\varphi(n)} \sum_{\emptyset \neq u \subseteq \{1:s\}} \gamma_u^{\lambda} \left( \frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}} \right)^{|u|} \right)^{1/\lambda} \\ &\quad \times \sup_{\substack{G \in L^2(D) \\ \|G\|_{L^2(D)} \leq 1}} \|\langle G, u_s \rangle_{L^2(D)}\|_{s,\gamma} \end{aligned}$$

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



Precisely the same analysis that we carried out before shows that choosing the weights

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

with arbitrary  $\delta > 0$ , yields the QMC convergence rate

$$\sqrt{\mathbb{E}_{\mathbf{\Delta}} \|I_s(u_s) - Q_{n,s}^{\mathbf{\Delta}}(u_s)\|_{L^2(D)}^2} = \mathcal{O}(\varphi(n)^{\max\{-1/p+1/2, -1+\delta\}}),$$

where the implied coefficient is independent of the dimension  $s$ .

Naturally, the dimensionally-truncated PDE solution in the above formula can be replaced by the dimensionally-truncated FE solution  $u_{s,h}$  (provided that we use a conforming FE method, i.e., the domain  $D$  is a polygon and we use, e.g., piecewise linear finite element basis functions to span the finite element space  $V_h$ ).