Quasi-Monte Carlo for Bayesian Optimal Experimental Design Problems Governed by PDEs

Vesa Kaarnioja (FU Berlin) and Claudia Schillings (FU Berlin)
SIAM UQ – February 27, 2024

Table of contents

Part I: Quasi-Monte Carlo methods

High-dimensional numerical integration

$$\int_{[0,1]^s} f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} \approx \sum_{i=1}^{n} w_i f(\boldsymbol{t}_i)$$

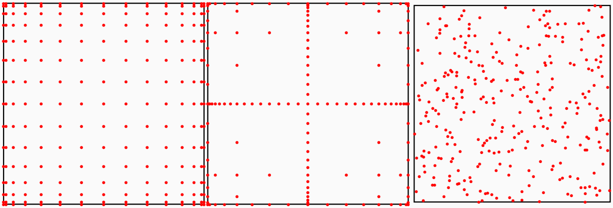


Figure 1: Tensor product grid, sparse grid, Monte Carlo nodes (not QMC rules)

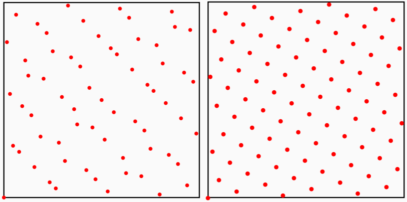


Figure 2: Sobol′ points, lattice rule (examples of QMC rules)

Quasi-Monte Carlo (QMC) methods are a class of *equal weight* cubature rules

$$\int_{[0,1]^s} f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} \approx \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{t}_i),$$

where $(\boldsymbol{t}_i)_{i=1}^n$ is an ensemble of *deterministic* nodes in $[0,1]^s$.

The nodes $(\boldsymbol{t}_i)_{i=1}^n$ are NOT random!! Instead, they are *deterministically chosen*.

QMC methods exploit the smoothness and anisotropy of an integrand in order to achieve better-than-Monte Carlo rates.

Lattice rules

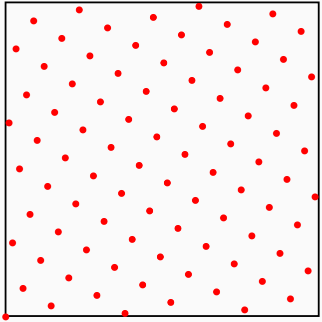
Rank-1 lattice rules

$$Q_{s,n}(f) = \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{t}_i)$$

have the points

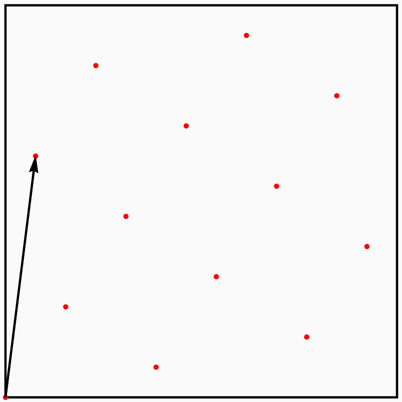
$$\boldsymbol{t}_i = \mathrm{mod}\left(\frac{i\boldsymbol{z}}{n}, 1\right), \quad i \in \{1, \ldots, n\},$$

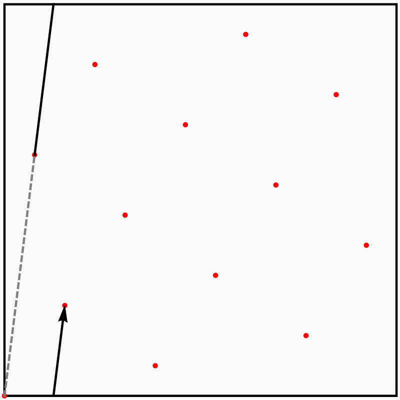
where the entire point set is determined by the *generating vector* $\boldsymbol{z} \in \mathbb{N}^s$, with all components *coprime* to n.

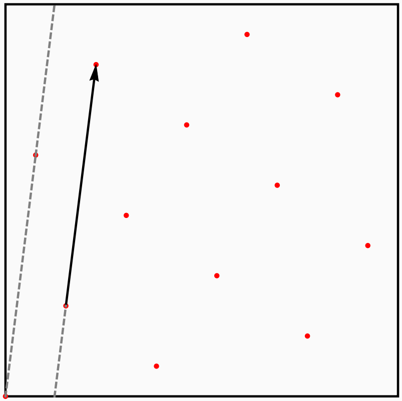


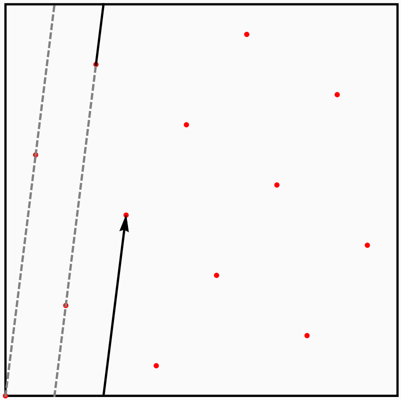
Lattice rule with $\boldsymbol{z} = (1, 55)$ and $n = 89$ nodes in $[0, 1]^2$

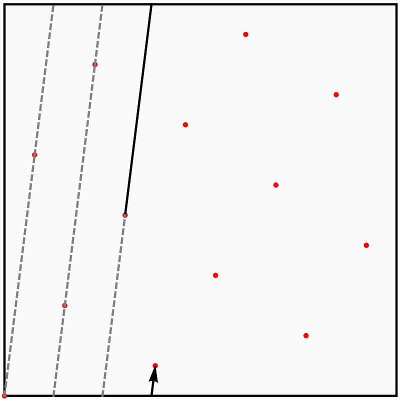
The quality of the lattice rule is determined by the choice of $\boldsymbol{z}$.

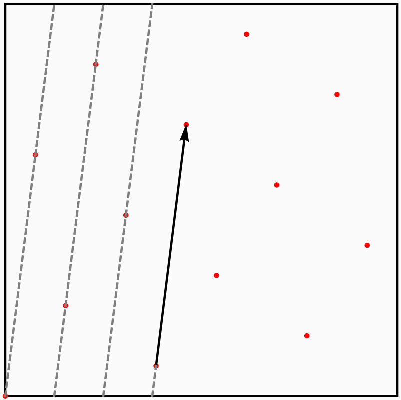


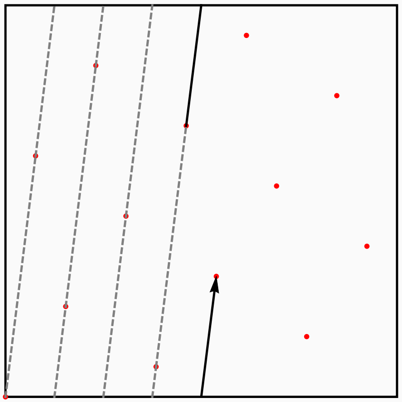


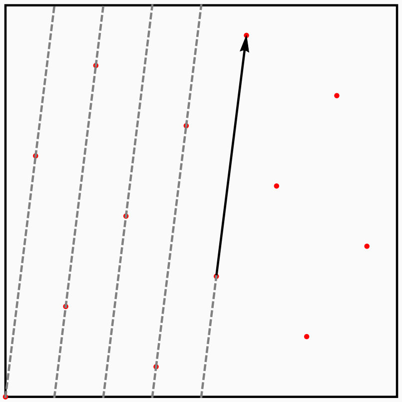


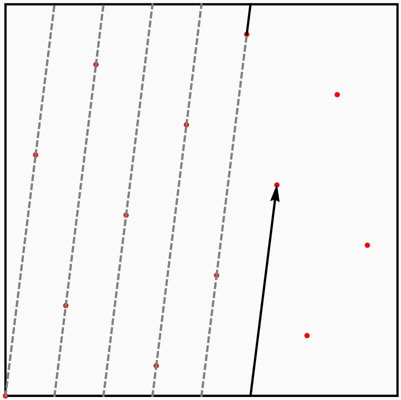


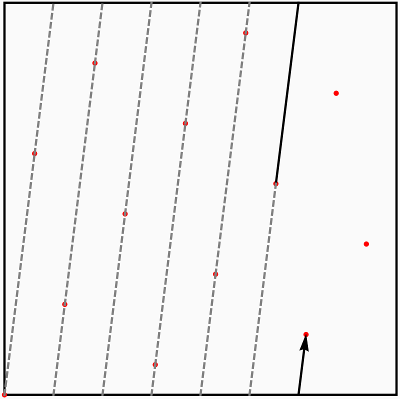


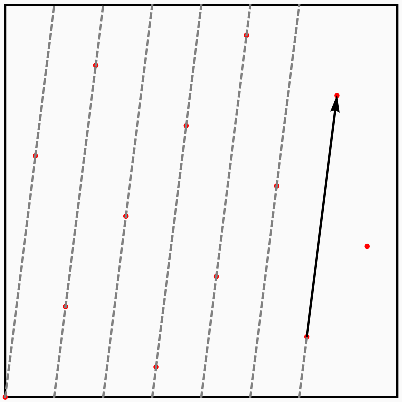


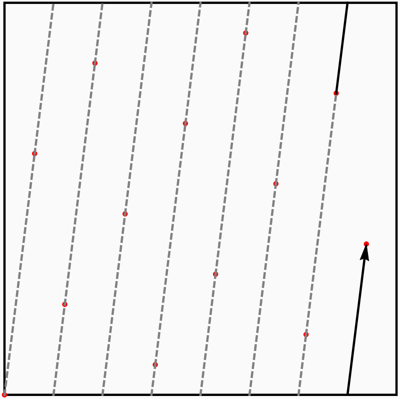












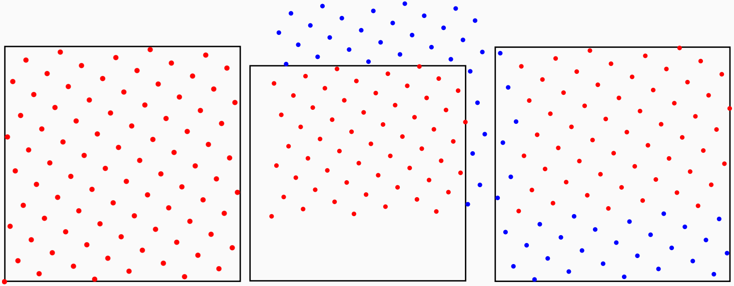
Randomly shifted lattice rules

Shifted rank-1 lattice rules have points

$$\boldsymbol{t}_i = \mathrm{mod}\left(\frac{i\boldsymbol{z}}{n} + \boldsymbol{\Delta}, 1\right), \quad i \in \{1, \ldots, n\}.$$

$\boldsymbol{\Delta} \in [0,1)^s$ is the *shift*

Use a number of random shifts for error estimation.



Lattice rule shifted by $\boldsymbol{\Delta} = (0.1, 0.3)$.

Let $\boldsymbol{\Delta}^{(r)}$, $r = 1, \ldots, R$, be independent random shifts drawn from $U([0,1]^s)$ and define

$$Q_{s,n}^{(r)}(f) := \frac{1}{n} \sum_{i=1}^{n} f(\mathrm{mod}(\boldsymbol{t}_i + \boldsymbol{\Delta}^{(r)}, 1)). \quad \text{(QMC rule with 1 random shift)}$$

Then

$$\overline{Q}_{s,n}(f) = \frac{1}{R} \sum_{r=1}^{R} Q_{s,n}^{(r)} f \qquad \text{(QMC rule with } R \text{ random shifts)}$$

is an unbiased estimator of $I_s(f)$.

Let $f : [0,1]^s \to \mathbb{R}$ be sufficiently smooth.

Error bound (one random shift):

$$|I_s(f) - Q_{s,n}^{\boldsymbol{\Delta}}(f)| \le e_{s,n,\gamma}^{\boldsymbol{\Delta}}(\boldsymbol{z}) \|f\|_{\boldsymbol{\gamma}}.$$

R.M.S. error bound (shift-averaged):

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}[|I_s(f) - \overline{Q}_{s,n}(f)|^2]} \le e_{s,n,\boldsymbol{\gamma}}^{\mathrm{sh}}(\boldsymbol{z}) \|f\|_{\boldsymbol{\gamma}}.$$

We consider weighted Sobolev spaces with dominating mixed smoothness, equipped with norm

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

and (squared) worst case error

$$P(\boldsymbol{z}) := e_{s,n,\boldsymbol{\gamma}}^{\mathrm{sh}}(\boldsymbol{z})^2 = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\varnothing \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}} \prod_{j \in \mathfrak{u}} \omega\left(\left\{ \frac{k z_j}{n} \right\} \right)$$

where $\omega(x) = x^2 - x + \frac{1}{6}$.

CBC algorithm (Sloan, Kuo, Joe 2002)

The idea of the *component-by-component* (CBC) algorithm is to find a good generating vector $\boldsymbol{z} = (z_1, \ldots, z_s)$ by proceeding as follows:

1. Set $z_1 = 1$;
2. With z_1 fixed, choose z_2 to minimize error criterion $P(z_1, z_2)$;
3. With z_1 and z_2 fixed, choose z_3 to minimize error criterion $P(z_1, z_2, z_3)$
   $\vdots$

Efficient implementation using FFT (QMC4PDE, QMCPy, etc.) if weights have certain structure (e.g., POD weights).

Theorem (CBC error bound)
The generating vector $\boldsymbol{z} \in \mathbb{N}^s$ constructed by the CBC algorithm, minimizing the squared shift-averaged worst-case error $[e^{\mathrm{sh}}_{s,n,\boldsymbol{\gamma}}(\boldsymbol{z})]^2$ for the weighted unanchored Sobolev space in each step, satisfies

$$[e^{\mathrm{sh}}_{s,n,\boldsymbol{\gamma}}(\boldsymbol{z})]^2 \le \left(\frac{1}{\varphi(n)} \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}}\right)^{|\mathfrak{u}|}\right)^{1/\lambda} \quad \textit{for all } \lambda \in (1/2, 1],$$

where $\zeta(x) := \sum_{k=1}^{\infty} k^{-x}$ denotes the Riemann zeta function for $x > 1$.

Remarks:

- Optimal rate of convergence $\mathcal{O}(n^{-1+\delta})$ in weighted Sobolev spaces, independently of s under an appropriate condition on the weights.
- Cost of algorithm for POD weights is $\mathcal{O}(s\, n \log n + s^2\, n)$ using FFT.

Significance: Suppose that $f \in H_{s,\boldsymbol{\gamma}}$ for all $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u} \subseteq \{1:s\}}$. Then for any given sequence of weights $\boldsymbol{\gamma}$, we can use the CBC algorithm to obtain a generating vector satisfying the error bound

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

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).

Part II: Bayesian optimal experimental design

Let $G : \Theta \times \Xi \to \mathbb{R}^k$ be a forward mapping depending on a true parameter $\boldsymbol{\theta} \in \Theta$ and a design parameter $\boldsymbol{\xi} \in \Xi$.

Measurement model:

$$\boldsymbol{y} = G(\boldsymbol{\theta}, \boldsymbol{\xi}) + \boldsymbol{\eta},$$

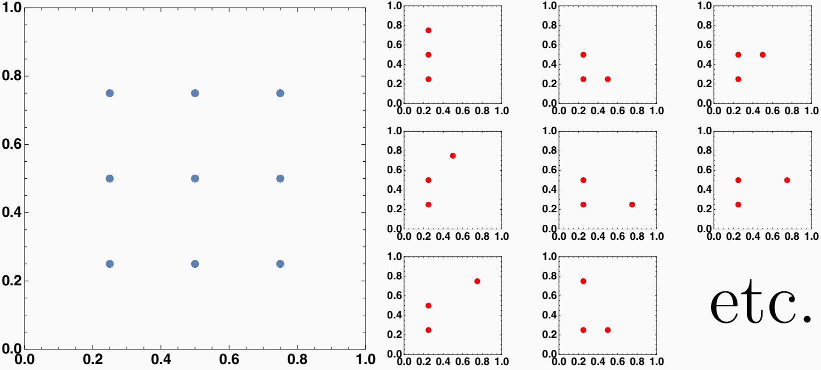
where $\boldsymbol{y} \in \mathbb{R}^k$ is the measurement data and $\boldsymbol{\eta} \in \mathbb{R}^k$ is Gaussian noise such that $\boldsymbol{\eta} \sim \mathcal{N}(0, \Gamma)$ with positive definite covariance matrix $\Gamma \in \mathbb{R}^{k \times k}$.

Goal in Bayesian optimal experimental design: Recover the design parameter $\boldsymbol{\xi}$ for the Bayesian inference of $\boldsymbol{\theta}$, which we model as a random variable endowed with prior distribution $\pi(\boldsymbol{\theta})$.

Example

Suppose we have 9 slots and 3 sensors. Before carrying out the experiment, which 3 slots do we expect to be the most informative for the recovery of the unknown parameter?

$\rightarrow \binom{9}{3} = 84$ possible configurations



Left: 9 slots. Right: We have 84 possible ways to place 3 sensors into 9 slots.

How to rank the 84 different possibilities from most informative to least informative?

A measure of the information gain for a given design $\boldsymbol{\xi}$ and data $\boldsymbol{y}$ is given by the Kullback–Leibler divergence

$$D_{\mathrm{KL}}(\pi(\cdot|\boldsymbol{y},\boldsymbol{\xi})\|\pi(\cdot)) := \int_{\Theta} \log\left(\frac{\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{\xi})}{\pi(\boldsymbol{\theta})}\right)\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\theta}. \tag{2}$$

We wish to maximize the expected utility (2) over the design space Ξ with respect to the data $\boldsymbol{y}$ and model parameters $\boldsymbol{\theta}$:

$$\max_{\boldsymbol{\xi}\in\Xi} \underbrace{\int_{Y}\int_{\Theta} \log\left(\frac{\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{\xi})}{\pi(\boldsymbol{\theta})}\right)\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{\xi})\pi(\boldsymbol{y}|\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{y}}_{=:\mathrm{EIG}},$$

where $\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{\xi})$ corresponds to the posterior distribution of the parameter $\boldsymbol{\theta}$ and $\pi(\boldsymbol{y}|\boldsymbol{\xi}) = \int_{\Theta}\pi(\boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{\xi})\pi(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}$ is the marginal distribution of the data $\boldsymbol{y}$.

The posterior is given by Bayes' theorem

$$\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{\xi}) = \frac{\pi(\boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{\xi})\pi(\boldsymbol{\theta})}{\pi(\boldsymbol{y}|\boldsymbol{\xi})},$$

which means that the expected utility can be written as

$$\begin{aligned}
\mathrm{EIG} &= \int_Y \int_\Theta \log\left(\frac{\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{\xi})}{\pi(\boldsymbol{\theta})}\right)\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\theta}\,\pi(\boldsymbol{y}|\boldsymbol{\xi})\mathrm{d}\boldsymbol{y} \\
&= \int_\Theta \left[\int_Y \log\left(\frac{\pi(\boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{\xi})}{\pi(\boldsymbol{y}|\boldsymbol{\xi})}\right)\pi(\boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{\xi})\mathrm{d}\boldsymbol{y}\right]\pi(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}.
\end{aligned}$$

Approaches taken in the literature:

- Double-loop Monte Carlo (Beck, Mansour, Espath, Long, Tempone)
- MCLA (Beck, Mansour, Espath, Long, Tempone)
- DLMCIS (Beck, Mansour, Espath, Long, Tempone)

Let us assume the following:

A1 The forward model $\boldsymbol{y} = G(\boldsymbol{\theta}, \boldsymbol{\xi}) + \boldsymbol{\eta}$ satisfies

$$\|\partial_{\boldsymbol{\theta}}^{\boldsymbol{\nu}} G(\boldsymbol{\theta}, \boldsymbol{\xi})\| \leq C_0 |\boldsymbol{\nu}|! \boldsymbol{b}^{\boldsymbol{\nu}},$$

where $\boldsymbol{b} := (b_j)_{j \geq 1} \in \ell^p$ are nonnegative real numbers for some $p \in (0,1)$ and $C_0 \geq 1$ is independent of $\boldsymbol{\xi} \in \Xi$.

A2 $\Theta = [-\frac{1}{2}, \frac{1}{2}]^s$ and $\pi(\boldsymbol{\theta}) = 1$ for $\boldsymbol{\theta} \in \Theta$ and 0 otherwise.

A3 The noise covariance is $\Gamma = \sigma^2 I_k$, $0 < \sigma \leq 1$.

Example

Model problem: Let $D \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, be a nonempty, bounded, and convex Lipschitz domain and $z \in L^2(D)$. For each $\boldsymbol{\theta} \in \Theta$, there exists a strong solution $u(\cdot, \boldsymbol{\theta}) \in H^2(D) \cap H_0^1(D)$ to

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{x}, \boldsymbol{\theta}) \nabla u(\boldsymbol{x}, \boldsymbol{\theta})) = z(\boldsymbol{x}), & \boldsymbol{x} \in D, \ \boldsymbol{\theta} \in \Theta, \\ u(\boldsymbol{x}, \boldsymbol{\theta}) = 0, & \boldsymbol{x} \in \partial D, \ \boldsymbol{\theta} \in \Theta, \end{cases}$$

where we assume that $\boldsymbol{\theta} = (\boldsymbol{\theta}_j)_{j \geq 1}$ i.i.d. uniformly distributed in $[-1/2, 1/2]$,

$$a(\boldsymbol{x}, \boldsymbol{\theta}) = a_0(\boldsymbol{x}) + \sum_{j \geq 1} \theta_j \psi_j(\boldsymbol{x}), \quad \boldsymbol{x} \in D, \ \boldsymbol{\theta} \in [-1/2, 1/2]^{\mathbb{N}},$$

with $a_0 \in W^{1,\infty}(D)$ and $\psi_j \in W^{1,\infty}(D)$, $j \geq 1$, such that $\sum_{j \geq 1} \|\psi_j\|_{W^{1,\infty}(D)} < \infty$, $0 < a_{\min} \leq a(\boldsymbol{x}, \boldsymbol{\theta}) \leq a_{\max} < \infty$ for all $\boldsymbol{x} \in D$, $\boldsymbol{\theta} \in [-1/2, 1/2]^{\mathbb{N}}$, and $b_j := \|\psi_j\|_{L^\infty(D)} / a_{\min}$.

Then $G(\boldsymbol{\theta}, \boldsymbol{\xi}) := (u(\boldsymbol{x}, \boldsymbol{\theta}))_{\boldsymbol{x} \in \boldsymbol{\xi}}$, where $\boldsymbol{\xi} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \subset D$, satisfies A1.

If $\boldsymbol{\theta} \perp \boldsymbol{\eta}$, then the likelihood is given by

$$\pi(\boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{\xi}) = C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}, \quad C = \frac{1}{(2\pi\sigma^2)^{k/2}}.$$

Under these conditions, it is easy to see that

$$\mathrm{EIG} = \int_Y \int_\Theta \log\left(\frac{\pi(\boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{\xi})}{\pi(\boldsymbol{y}|\boldsymbol{\xi})}\right)\pi(\boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{\xi})\pi(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{y}$$
$$= \log C - 1 - \int_Y \log\left(\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\right)\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{y}.$$

Observations:

- The **inner integral** can be approximated *independently* of the dimension s using QMC.
- In general, the data dimension k affects the QMC cubature error bound of the **outer integral**.
- How to efficiently approximate the nested integrals?

Consider

$$\int_Y \log\left(\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\right)\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{y}.$$

- Parametric regularity of the integrand in

$$\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}$$

  is well-understood (as long as the parametric regularity of G can be quantified); straightforward modification of Herrmann–Keller–Schwab (2021).
- More generally, in many applications of UQ, we are interested in parametric integrals of quantities of the general form

$$f(G(\boldsymbol{\theta},\cdot)), \tag{3}$$

  where $f:\mathbb{R}^k\to\mathbb{R}$ is a (somewhat) smooth nonlinear quantity of interest†, and

$$\|\partial_{\boldsymbol{\theta}}^{\boldsymbol{\nu}} G(\boldsymbol{\theta},\cdot)\|\leq C_0|\boldsymbol{\nu}|!\boldsymbol{b}^{\boldsymbol{\nu}}.$$

  What can be said about the parametric regularity of (3)?

†E.g., $f(\boldsymbol{x})=\|\boldsymbol{y}-\boldsymbol{x}\|^2$ or $f(\boldsymbol{x})=\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-\boldsymbol{x}\|^2}$.

Faà di Bruno's formula:

$$\partial_{\boldsymbol{\theta}}^{\boldsymbol{\nu}} f(G(\boldsymbol{\theta},\cdot)) = \sum_{\substack{1\le|\boldsymbol{\lambda}|\le|\boldsymbol{\nu}| \\ \boldsymbol{\lambda}\in\mathbb{N}_0^k}} \partial_{\boldsymbol{x}}^{\boldsymbol{\lambda}} f(\boldsymbol{x})\Big|_{\boldsymbol{x}=G(\boldsymbol{\theta},\cdot)} \kappa_{\boldsymbol{\nu},\boldsymbol{\lambda}}(\boldsymbol{\theta}), \quad \boldsymbol{\nu}\neq\mathbf{0},$$

where the sequence $(\kappa_{\boldsymbol{\nu},\boldsymbol{\lambda}})$ depends only on G via

$\kappa_{\boldsymbol{\nu},\mathbf{0}} \equiv \delta_{\boldsymbol{\nu},\mathbf{0}}$,

$\kappa_{\boldsymbol{\nu},\boldsymbol{\lambda}} \equiv 0$ if $|\boldsymbol{\nu}| < |\boldsymbol{\lambda}|$ or $\boldsymbol{\lambda} \not\geq \mathbf{0}$ (i.e., if $\boldsymbol{\lambda}$ contains negative entries),

$$\kappa_{\boldsymbol{\nu}+\boldsymbol{e}_j,\boldsymbol{\lambda}}(\boldsymbol{\theta}) = \sum_{\ell\in\mathrm{supp}(\boldsymbol{\lambda})} \sum_{\mathbf{0}\le\boldsymbol{m}\le\boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial^{\boldsymbol{m}+\boldsymbol{e}_j}[G(\boldsymbol{\theta},\cdot)]_\ell \kappa_{\boldsymbol{\nu}-\boldsymbol{m},\boldsymbol{\lambda}-\boldsymbol{e}_\ell}(\boldsymbol{\theta}) \quad \text{otherwise.}$$

Since $\|\partial_{\boldsymbol{\theta}}^{\boldsymbol{\nu}} G(\boldsymbol{\theta},\cdot)\| \le C_0|\boldsymbol{\nu}|!\boldsymbol{b}^{\boldsymbol{\nu}}$, we obtain the following *uniform* bound.

Lemma
For all $1 \le |\boldsymbol{\lambda}| \le |\boldsymbol{\nu}|$, *there holds*

$$|\kappa_{\boldsymbol{\nu},\boldsymbol{\lambda}}(\boldsymbol{\theta})| \le C_0^{|\boldsymbol{\lambda}|} \frac{|\boldsymbol{\nu}|!(|\boldsymbol{\nu}|-1)!}{\boldsymbol{\lambda}!(|\boldsymbol{\nu}|-|\boldsymbol{\lambda}|)!(|\boldsymbol{\lambda}|-1)!} \boldsymbol{b}^{\boldsymbol{\nu}}.$$

Proof.
By induction w.r.t. the order of the multi-index $\boldsymbol{\nu}$. □

For the present problem, plugging the upper bound for $|\kappa_{\boldsymbol{\nu},\boldsymbol{\lambda}}(\boldsymbol{\theta})|$ into the expression

$$\partial_{\boldsymbol{\theta}}^{\boldsymbol{\nu}} f(G(\boldsymbol{\theta},\cdot)) = \sum_{\substack{1\le|\boldsymbol{\lambda}|\le|\boldsymbol{\nu}| \\ \boldsymbol{\lambda}\in\mathbb{N}_0^k}} \partial_{\boldsymbol{x}}^{\boldsymbol{\lambda}} f(\boldsymbol{x})\Big|_{\boldsymbol{x}=G(\boldsymbol{\theta},\cdot)} \kappa_{\boldsymbol{\nu},\boldsymbol{\lambda}}(\boldsymbol{\theta}), \quad \boldsymbol{\nu}\neq\mathbf{0},$$

and making some simple estimates yields altogether that

$$|\partial_{\boldsymbol{\theta}}^{\boldsymbol{\nu}} \mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\cdot)\|^2}| \le 3.82^k \cdot C_0^{|\boldsymbol{\nu}|} 2^{|\boldsymbol{\nu}|-1}\sigma^{-|\boldsymbol{\nu}|}|\boldsymbol{\nu}|!\boldsymbol{b}^{\boldsymbol{\nu}} \quad \text{for all } \boldsymbol{\nu}\neq\mathbf{0}.$$

Decomposing the high-dimensional integral

Suppose that $Y = \mathbb{R}^k$. Then

$$
\begin{aligned}
&\int_{\mathbb{R}^k} \log\left(\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\right)\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\widetilde{\boldsymbol{\theta}},\boldsymbol{\xi})\|^2}\,\mathrm{d}\widetilde{\boldsymbol{\theta}}\,\mathrm{d}\boldsymbol{y}\\
&= \int_{[-K,K]^k} \log\left(\int_\Theta C_{k,\sigma}\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\right)\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\widetilde{\boldsymbol{\theta}},\boldsymbol{\xi})\|^2}\,\mathrm{d}\widetilde{\boldsymbol{\theta}}\,\mathrm{d}\boldsymbol{y}\\
&\quad + \int_{\mathbb{R}^k\setminus[-K,K]^k} \log\left(\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\right)\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\widetilde{\boldsymbol{\theta}},\boldsymbol{\xi})\|^2}\,\mathrm{d}\widetilde{\boldsymbol{\theta}}\,\mathrm{d}\boldsymbol{y}\\
&=: \mathcal{I}_K + \widetilde{\mathcal{I}}_K.
\end{aligned}
$$

For $K \gg 1$, there holds

$$
\begin{aligned}
|\widetilde{\mathcal{I}}_K| \le\ & C^{1/2}\mathrm{e}^{\frac{\|\overline{G}\|^2}{4\sigma^2}}(4\pi\sigma^2)^{k/2}\\
&- C^{1/2}\mathrm{e}^{\frac{\|\overline{G}\|^2}{4\sigma^2}}(\pi\sigma^2)^{k/2}\prod_{j=1}^{k}\left(\operatorname{erf}\left(\frac{\overline{G}_j+K}{2\sigma}\right)-\operatorname{erf}\left(\frac{\overline{G}_j-K}{2\sigma}\right)\right),
\end{aligned}
$$

where $\overline{G} := (\overline{G}_j)_{j=1}^k := G(\boldsymbol{\theta}^*,\boldsymbol{\xi}^*)$, $(\boldsymbol{\theta}^*,\boldsymbol{\xi}^*) := \arg\max_{(\boldsymbol{\theta},\boldsymbol{\xi})\in(\Theta,\Xi)}\|G(\boldsymbol{\theta},\boldsymbol{\xi})\|$.

$$\int_{[-K,K]^k} \log\left(\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\right)\int_\Theta C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{y}$$

QMC weights for the **inner integral**:

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

with $\beta_j := \frac{2C_0}{\sigma}b_j$, $j\in\{1,\ldots,s\}$, and $\delta>0$ arbitrary, yields the QMC convergence rate

$$\mathcal{O}(\varphi(n)^{\max\{-1/p+1/2,-1+\delta\}})$$

independently of the dimension s with $\varphi(n)$ denoting the Euler totient function.

$$\int_{[-K,K]^k} \log\left(\int_\Theta C e^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\right)\int_\Theta C e^{-\frac{1}{2\sigma^2}\|\boldsymbol{y}-G(\boldsymbol{\theta},\boldsymbol{\xi})\|^2}\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{y}$$

QMC weights for the **outer integral**:

$$\widetilde{\gamma}_{\mathfrak{u}} = \left(|\mathfrak{u}|!\prod_{j\in\mathfrak{u}}\frac{1.1^k k\sigma^{-1}e^{\frac{1}{2\sigma^2}(kK^2+2\sqrt{k}KC+C^2)}}{\log(2)\sqrt{2\zeta(2\widetilde{\lambda})/(2\pi^2)^{\widetilde{\lambda}}}}\right)^{\frac{2}{1+\widetilde{\lambda}}}, \quad \widetilde{\lambda} = \frac{1}{2-2\widetilde{\delta}},$$

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

$$\mathcal{O}(\varphi(n)^{-1+\widetilde{\delta}})$$

with $\varphi(n)$ denoting the Euler totient function. Note that the implied coefficient depends on k.

The nested integral

Goal of computation:

$$\mathcal{I}_K(f) = \int_{Y_K} g\bigg(\int_{\Theta} f(\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\theta} \bigg) \, \mathrm{d}\boldsymbol{y},$$

where $g(x) := x \log x$, $Y_K = [-K, K]^k$, and $f(\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{\xi}) := C\mathrm{e}^{-\frac{1}{2\sigma^2}\|\boldsymbol{y} - G(\boldsymbol{\theta}, \boldsymbol{\xi})\|^2}$.

Define a hierarchy of QMC cubature operators for the **outer integral**, i.e.,

$$I^{(1)}F := \int_{Y_K} F(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \approx 2^{-\ell} \sum_{k=1}^{2^\ell} F(\boldsymbol{y}_k^{(\ell)}) =: Q_\ell^{(1)}F, \; \ell = \ell_0^{(1)}, \ell_0^{(1)}+1, \ell_0^{(1)}+2, \ldots,$$

for a given function $F \in \widetilde{H}_{k,\widetilde{\gamma}}$, and likewise for the **inner integral**

$$I^{(2)}F := \int_{\Theta} F(\boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{\theta} \approx 2^{-\ell} \sum_{k=1}^{2^\ell} F(\boldsymbol{\theta}_k^{(\ell)}) =: Q_\ell^{(2)}F, \; \ell = \ell_0^{(2)}, \ell_0^{(2)}+1, \ell_0^{(2)}+2, \ldots,$$

for a given function $F \in H_{s,\boldsymbol{\gamma}}$.

Why full tensor product cubature is a bad idea

Approximating the integral

$$\mathcal{I}_K(f) = \int_{Y_K} g\bigg(\int_{\Theta} f(\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\theta} \bigg)\, \mathrm{d}\boldsymbol{y},$$

by

$$\mathcal{I}_K(f) \approx Q_\ell^{(1)} g(Q_\ell^{(2)} f) \tag{4}$$

is inefficient. A hand-wavy argument would be as follows:

- Suppose that we have the approximation rates (recall $n = 2^\ell$)

$$|I^{(1)}F - Q_\ell^{(1)}F| \asymp n^{-\alpha} \quad \text{and} \quad |I^{(2)}F - Q_\ell^{(2)}F| \asymp n^{-\alpha}.$$

- Evaluating (4) takes $N = n^2$ function calls, but the cubature accuracy will not be better than $\mathcal{O}(n^{-\alpha}) = \mathcal{O}(N^{-\alpha/2})$
  $\rightarrow$ the convergence rate is effectively halved! ("Curse of dimensionality")

Sparse tensor product cubature in the vein of Gilch, Griebel, Oettershagen (2022)

Define the difference cubature operator corresponding to the **outer integral**

$$\Delta_\ell^{(1)} F := \begin{cases} Q_\ell^{(1)} F - Q_{\ell-1}^{(1)} F & \text{if } \ell \geq 1, \\ Q_0^{(1)} F & \text{if } \ell = 0, \end{cases}$$

as well as the *generalized* difference cubature operators corresponding to the **inner integral**

$$\Delta_\ell^{(2)} F := \begin{cases} g(Q_\ell^{(2)} F) - g(Q_{\ell-1}^{(2)} F) & \text{if } \ell \geq 1, \\ g(Q_0^{(2)} F) & \text{if } \ell = 0. \end{cases}$$

Generalized sparse grid cubature operator:

$$\mathcal{Q}_{L,\varsigma}(f) := \sum_{\varsigma \ell_1 + \frac{\ell_2}{\varsigma} \leq L} \Delta_{\ell_1}^{(1)} \Delta_{\ell_2}^{(2)}(f) = \sum_{\ell_1=0}^{L/\varsigma} \Delta_{\ell_1}^{(1)} g(Q_{\varsigma L - \varsigma^2 \ell_1}^{(2)} f).$$

Sparse tensor product cubature

$$\mathcal{Q}_{L,\varsigma}(f) := \sum_{\varsigma\ell_1+\frac{\ell_2}{\varsigma}\leq L} \Delta_{\ell_1}^{(1)}\Delta_{\ell_2}^{(2)}(f) = \sum_{\ell_1=0}^{L/\varsigma} \Delta_{\ell_1}^{(1)} g(Q_{\varsigma L-\varsigma^2\ell_1}^{(2)} f).$$

Sparse grid error: Our **inner** and **outer** QMC cubatures have essentially linear convergence rates, i.e.,

$$|I^{(1)}f - Q_\ell^{(1)}f| \lesssim 2^{-(1-\delta)\ell} \quad \text{and} \quad |I^{(2)}f - Q_\ell^{(2)}f| \lesssim 2^{-(1-\delta)\ell}.$$

For an *isotropic* ($\varsigma = 1$) sparse tensor product cubature operator, we obtain

$$\|\mathcal{I}_K(f) - \mathcal{Q}_{L,\varsigma}(f)\|_{\mathbf{\Delta}} \lesssim 2^{-(1-\delta)L}(L+1)$$

under some additional technical assumptions.

Numerical experiment

Let $D = (0,1)^2$. We consider the elliptic PDE

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{x},\boldsymbol{\theta})\nabla u(\boldsymbol{x},\boldsymbol{\theta})) = 10x_1, & \boldsymbol{x} \in D,\ \boldsymbol{\theta} \in [-1/2,1/2]^{100}, \\ u(\cdot,\boldsymbol{\theta})|_{\partial D} = 0, & \boldsymbol{\theta} \in [-1/2,1/2]^{100}, \end{cases}$$

equipped with the parametric diffusion coefficient

$$a(\boldsymbol{x},\boldsymbol{\theta}) = 1 + 0.1\sum_{j=1}^{100} j^{-2}\theta_j \sin(\pi j x_1)\sin(\pi j x_2), \boldsymbol{\theta} \in [-1/2,1/2]^{100}.$$

Numerical experiment

The goal is to find a design $\boldsymbol{\xi}^*$ from the set

$$\Xi = \{(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) \in \Upsilon^3 \mid \boldsymbol{x}_i \neq \boldsymbol{x}_j \text{ for } i \neq j\},$$

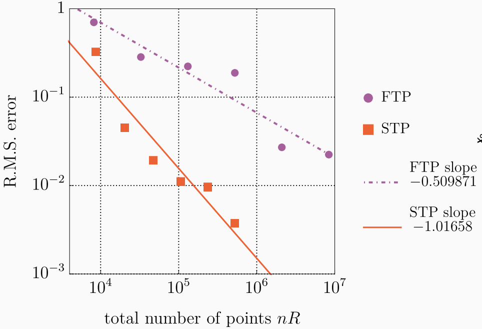
where

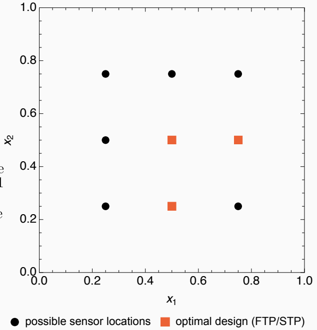
$$\begin{aligned}\Upsilon = & \{(0.25, 0.25), (0.25, 0.50), (0.25, 0.75),\\ & (0.50, 0.25), (0.50, 0.50), (0.50, 0.75),\\ & (0.75, 0.25), (0.75, 0.50), (0.75, 0.75)\},\end{aligned}$$

maximizing the expected information gain subject to the observation operator

$$G(\boldsymbol{\theta}, \boldsymbol{\xi}) = (u(\boldsymbol{x}, \boldsymbol{\theta}))_{\boldsymbol{x} \in \boldsymbol{\xi}}, \quad \boldsymbol{\theta} \in [-1/2, 1/2]^{100},\ \boldsymbol{\xi} \in \Xi.$$

Numerical experiment





Left: R.M.S. errors for the full tensor product (FTP) and sparse tensor product (STP) cubatures of the nested integral subject to affine and uniform parameterization of the input random field with $R = 8$ random shifts.
Right: the optimal design corresponding to the cubature rule with the largest number of points.

Conclusions and outlook

- QMC for DOE: sparse approach can recover almost the optimal rate.
- Related QMC analysis for PDE-constrained optimal control under uncertainty (including also the optimization process):

P. A. Guth, K., F. Y. Kuo, C. Schillings, and I. H. Sloan. A Quasi-Monte Carlo Method for Optimal Control Under Uncertainty. *SIAM/ASA J. Uncertain. Quantif.* **9**(2):354–383, 2021

P. A. Guth, K., F. Y. Kuo, C. Schillings, and I. H. Sloan. *Parabolic PDE-Constrained Optimal Control Under Uncertainty With Entropic Risk Measure Using Quasi-Monte Carlo Integration*. To appear in *Numer. Math.*, 2024

Thank you for your attention!