Uncertainty Quantification and Quasi-Monte Carlo

Sommersemester 2025

Vesa Kaarnioja
vesa.kaarnioja@fu-berlin.de

FU Berlin, FB Mathematik und Informatik

Eleventh lecture, July 7, 2025

Summary

Elliptic PDE

Many physical phenomena can be modeled using elliptic partial differential equations of the form

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{x}) \nabla u(\boldsymbol{x})) = f(\boldsymbol{x}), \quad \boldsymbol{x} \in D, \\ +\text{boundary conditions} \end{cases}$$

Uncertainties can appear in the material parameter a, source term f, boundary conditions, or the domain D.

- For the purposes of analysis, we consider the weak formulation of the PDE. Under certain conditions, the solution to the weak formulation can be shown to exist and be uniquely defined.
- When we solve the PDE numerically using the finite element method, we are actually approximating the solution to the *the weak formulation* of the PDE problem.
- Under suitably strong regularity assumptions (D convex Lipschitz domain, $f \in L^2(D)$, and a Lipschitz), the weak solution satisfies $-\nabla \cdot (a(\boldsymbol{x}) \nabla u(\boldsymbol{x})) = f(\boldsymbol{x})$ for a.e. $\boldsymbol{x} \in D$ with $u|_{\partial D} = 0$.

Let $D \subset \mathbb{R}^d$ be a nonempty open set.

$$L^2(D) := \{v \colon D \to \mathbb{R} \mid v \text{ is measurable, } \|v\|_{L^2(D)} := \left(\int_D |v(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} \right)^{1/2} < \infty\},$$

$$H^1(D) := \{v \in L^2(D) \mid \partial_j v \in L^2(D) \text{ for all } j \in \{1, \dots, d\}\},$$

with $\|v\|_{H^1(D)} := (\|v\|^2_{L^2(D)} + \|\nabla v\|^2_{L^2(D)})^{1/2}$,

$$C_0^\infty(D) := \{v \in C^\infty(D) \mid \operatorname{supp}(v) \subset D \text{ is a compact set}\},$$

where $\operatorname{supp}(v) := \overline{\{\boldsymbol{x} \in D \mid v(\boldsymbol{x}) \neq 0\}}$,

$$H_0^1(D) := \operatorname{cl}_{H^1(D)}(C_0^\infty(D)).$$

The spaces $L^2(D)$, $H^1(D)$, and $H_0^1(D)$ are Hilbert spaces.

Poincaré's inequality: if $D \subset \mathbb{R}^d$ is a bounded domain, then there exists a constant $C_P > 0$ (depending on the domain D) such that

$$\|v\|_{L^2(D)} \leq C_P \|\nabla v\|_{L^2(D)} \quad \text{for all } v \in H_0^1(D).$$

Therefore, we can define an equivalent norm in $H_0^1(D)$ by setting

$$\|v\|_{H_0^1(D)} := \|\nabla v\|_{L^2(D)}.$$

This induces exactly the same topology in $H_0^1(D)$ as the usual Sobolev norm $\|\cdot\|_{H^1(D)}$.

Trace theorem and boundary values

Trace theorem: Let D be a bounded Lipschitz domain. Then the trace operator

$$\gamma\colon C^\infty(\overline{D}) \to C^\infty(\partial D),\ \gamma u = u|_{\partial D},$$

has a unique extension to a bounded linear operator $\gamma\colon H^1(D) \to L^2(\partial D)$.

This means that even though $u \in H^1(D)$ is not well-defined over a set of measure zero, we can interpret its restriction to the boundary of the domain D as the trace $\gamma u \in L^2(\partial D)$.

Especially, Sobolev functions $u \in H^1(D)$ with zero trace are precisely the elements of $H^1_0(D)$:

$$u \in H^1_0(D) \quad \Leftrightarrow \quad \gamma u = 0\colon \partial D \to \mathbb{R}.$$

Q: *How to solve such PDE problems in practice?*

A: We consider the *weak formulation* of the PDE problem: given $f \in L^2(D)$, find $u \in H_0^1(D)$ such that

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

where $F \colon H_0^1(D) \to \mathbb{R}$ is a bounded linear functional. If there exist $a_{\min}, a_{\max} > 0$ s.t. $0 < a_{\min} \leq a(\boldsymbol{x}) \leq a_{\max} < \infty$ for all $\boldsymbol{x} \in D$, then the bilinear form $B \colon H_0^1(D) \times H_0^1(D) \to \mathbb{R}$ is bounded, i.e.,

$$|B(u,v)| = \left| \int_D a(\boldsymbol{x}) \nabla u(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \right| \leq a_{\max} \|u\|_{H_0^1(D)} \|v\|_{H_0^1(D)}$$

for all $u, v \in H_0^1(D)$, and coercive, i.e.,

$$B(u,u) = \left| \int_D a(\boldsymbol{x}) \nabla u(\boldsymbol{x}) \cdot \nabla u(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \right| \geq a_{\min} \|u\|_{H_0^1(D)} \quad \text{for all } u \in H_0^1(D),$$

the *Lax–Milgram lemma* ensures that there exists a unique solution $u \in H_0^1(D)$ to (**??**).

Galerkin method

To solve the system approximately, let $V_m \subset H_0^1(D)$ be a finite-dimensional subspace of the solution space $H_0^1(D)$.

The *Galerkin solution* $u_m \in V_m$ of the system (**??**) is the unique solution such that

$$\int_D a(\boldsymbol{x}) \nabla u_m(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_D f(\boldsymbol{x}) v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \quad \text{for all } v \in V_m.$$

Let V_m be spanned by $\psi_1, \ldots, \psi_m$. We can write the solution as $u_m = \sum_{i=1}^m c_i \psi_i$. The above system reduces to the linear system of equations

$$\begin{bmatrix} \int_D \nabla\psi_1(\boldsymbol{x}) \cdot \nabla\psi_1(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} & \cdots & \int_D \nabla\psi_1(\boldsymbol{x}) \cdot \nabla\psi_m(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \\ \vdots & \ddots & \vdots \\ \int_D \nabla\psi_m(\boldsymbol{x}) \cdot \nabla\psi_1(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} & \cdots & \int_D \nabla\psi_m(\boldsymbol{x}) \cdot \nabla\psi_m(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} \int_D f(\boldsymbol{x}) \psi_1(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \\ \vdots \\ \int_D f(\boldsymbol{x}) \psi_m(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \end{bmatrix}.$$

Solving this system and plugging the expansion coefficients back into the expression for u_m yields the Galerkin solution.

Céa's lemma

The solution to the Galerkin system is quasi-optimal in the following sense:

$$\|u - u_m\|_{H_0^1(D)} \leq \frac{a_{\max}}{a_{\min}} \inf_{v_m \in V_m} \|u - v_m\|_{H_0^1(D)}.$$

That is, the $H_0^1(D)$ error between the true PDE solution u and the Galerkin approximation u_m differs from the *optimal approximation* in V_m up to a constant factor.

Finite element method

The finite element method is a particular method of constructing the finite-dimensional subspaces V_m of the solution space $H_0^1(D)$.

- Construct a triangulation for the computational domain D.
- The space V_m is spanned by piecewise linear functions $\psi_1, \ldots, \psi_m$ which are constructed to satisfy

$$\psi_i(\boldsymbol{n}_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases}$$

  where $\boldsymbol{n}_1, \ldots, \boldsymbol{n}_m$ are the *interior* nodes of the triangulation.
- The finite element solution can be written as $u_h(\boldsymbol{x}) = \sum_{i=1}^m c_i \psi_i(\boldsymbol{x}) \in V_h$, where the expansion coefficients are solved from the Galerkin system. Note that $u_h(\boldsymbol{n}_j) = c_j$.
- If $v_h(\boldsymbol{x}) = \sum_{i=1}^m c_i \psi_i(\boldsymbol{x}) \in V_h$, then, e.g., $\|v_h\|_{L^2(D)} = \sqrt{\boldsymbol{c}^{\mathrm{T}} M \boldsymbol{c}}$, where $\boldsymbol{c} := [c_1, \ldots, c_m]^{\mathrm{T}}$ and $M = [M_{i,j}]_{i,j=1}^m$ is the mass matrix defined elementwise by $M_{i,j} := \int_D \psi_i(\boldsymbol{x}) \psi_j(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$, $i, j \in \{1, \ldots, m\}$.

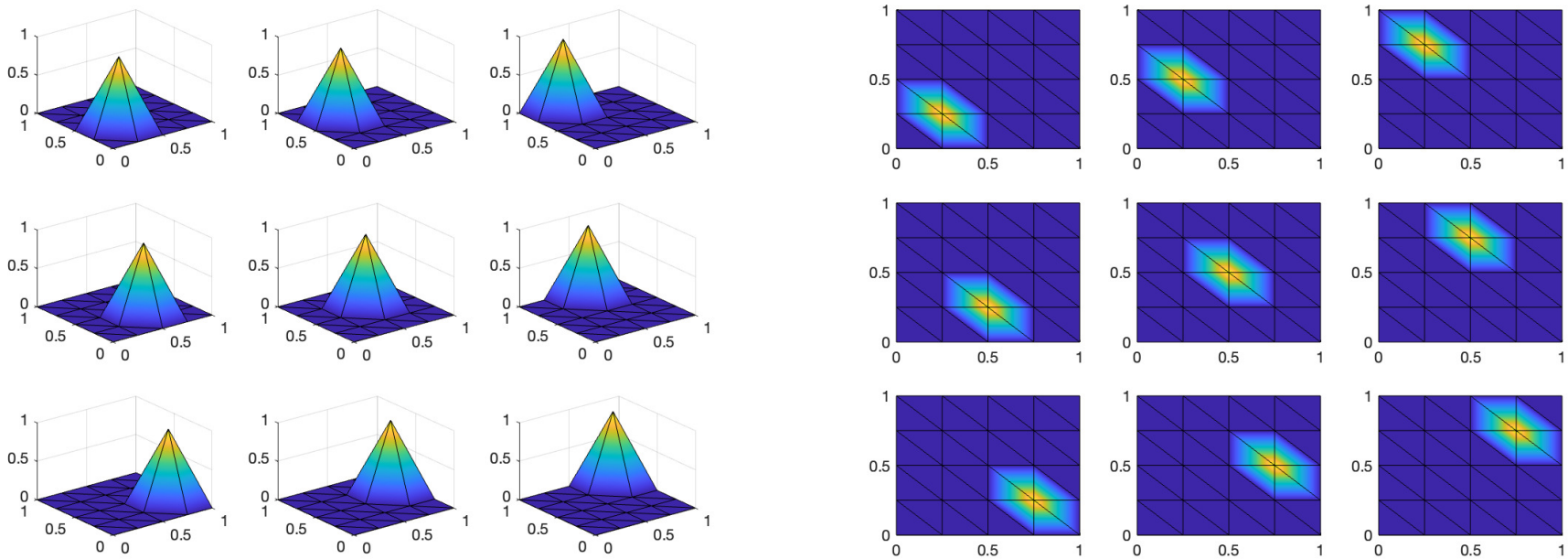


Figure: Left: An illustration of global, piecewise linear FE basis functions spanning V_h over a regular, uniform triangulation of $(0,1)^2$. Right: Bird's-eye view of the same global FE basis functions.

Random field

Definition

Let $D \subset \mathbb{R}^d$ and let $(\Omega, \mathcal{F}, \mu)$ be a probability space. A function $A\colon D \times \Omega \to X$ is called a *random field* if $A(\boldsymbol{x}, \cdot)$ is an X-valued random variable for all $\boldsymbol{x} \in D$.

Definition

We call a random field $A\colon D \times \Omega \to X$ square-integrable if

$$\int_\Omega |A(\boldsymbol{x}, \omega)|^2 \, \mu(\mathrm{d}\omega) < \infty \quad \text{for all } \boldsymbol{x} \in D.$$

Our goal will be to model (infinite-dimensional) input random fields using finite-dimensional expansions with s variables.

Comment on notation: In what follows, s will always refer to the "stochastic dimension" (dimension of the stochastic/parametric space) while d will refer to the "spatial dimension" (dimension of the spatial Lipschitz domain $D \subset \mathbb{R}^d$, $d \in \{2, 3\}$).

Mercer's theorem

Let $a(\boldsymbol{x}, \omega)$ be a square-integrable random field with mean

$$\bar{a}(\boldsymbol{x}) = \int_\Omega a(\boldsymbol{x}, \omega)\, \mu(\mathrm{d}\omega), \quad \boldsymbol{x} \in D,$$

and a continuous, symmetric, positive definite[†] covariance

$$K(\boldsymbol{x}, \boldsymbol{x}') = \int_\Omega (a(\boldsymbol{x}, \omega) - \bar{a}(\boldsymbol{x}))(a(\boldsymbol{x}', \omega) - \bar{a}(\boldsymbol{x}'))\, \mu(\mathrm{d}\omega).$$

Mercer's theorem: the covariance operator $\mathcal{C}\colon L^2(D) \to L^2(D)$,

$$(\mathcal{C}u)(\boldsymbol{x}) = \int_D K(\boldsymbol{x}, \boldsymbol{x}')u(\boldsymbol{x}')\, \mathrm{d}\boldsymbol{x}', \quad \boldsymbol{x} \in D,$$

has a countable sequence of eigenvalues $\{\lambda_k\}_{k\geq 1}$ and corresponding eigenfunctions $\{\psi_k\}_{k\geq 1}$ satisfying $\mathcal{C}\psi_k = \lambda_k\psi_k$ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ and $\lambda_k \to 0$ and the eigenfunctions form an orthonormal basis for $L^2(D)$.
Note that this means:

$$\int_D K(\boldsymbol{x}, \boldsymbol{x}')\psi_k(\boldsymbol{x}')\, \mathrm{d}\boldsymbol{x}' = \lambda_k\psi_k(\boldsymbol{x}), \quad \int_D \psi_k(\boldsymbol{x})\psi_\ell(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \delta_{k,\ell}.$$

[†]In this context, positive definite means: for all choices of finitely many points $x_1, \ldots, x_k \in D$, $k \in \mathbb{N}$, the *Gram matrix* $G := [K(x_i, x_j)]_{i,j=1}^k$ is positive semidefinite.

The Karhunen–Loève (KL) expansion of a random field

Theorem

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space, let $D \subset \mathbb{R}^d$ be closed and bounded, and let $a\colon D \times \Omega \to \mathbb{R}$ be a square-integrable random field with continuous, symmetric, positive definite covariance $K(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(a(\boldsymbol{x}, \cdot) - \bar{a}(\boldsymbol{x}))(a(\boldsymbol{x}', \cdot) - \bar{a}(\boldsymbol{x}'))]$. Then the eigensystem $(\lambda_k, \psi_k) \in \mathbb{R}_+ \times L^2(D)$ of the covariance operator $\mathcal{C}\colon L^2(D) \to L^2(D)$, as described on the previous slide, can be used to write

$$a(\boldsymbol{x}, \omega) = \bar{a}(\boldsymbol{x}) + \sum_{k=1}^{\infty} \sqrt{\lambda_k} \xi_k(\omega) \psi_k(\boldsymbol{x}),$$

$$\text{where } \xi_k(\omega) = \frac{1}{\sqrt{\lambda_k}} \int_D (a(\boldsymbol{x}, \omega) - \bar{a}(\boldsymbol{x})) \psi_k(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x},$$

where the convergence is in L^2 w.r.t. the stochastic parameter and uniform in $\boldsymbol{x}$. Furthermore, the random variables ξ_k are zero-mean uncorrelated random variables with unit variance, i.e.,

$$\mathbb{E}[\xi_k] = 0 \quad \text{and} \quad \mathbb{E}[\xi_k \xi_\ell] = \delta_{k,\ell}.$$

The Karhunen–Loève (KL) expansion of random field $a(\boldsymbol{x}, \omega)$ can be written as

$$a(\boldsymbol{x}, \omega) = \bar{a}(\boldsymbol{x}) + \sum_{k=1}^{\infty} \sqrt{\lambda_k} \xi_k(\omega) \psi_k(\boldsymbol{x}).$$

- The KL expansion minimizes the mean-square truncation error:

  $$\left\| a(\boldsymbol{x}, \omega) - \bar{a}(\boldsymbol{x}) - \sum_{k=1}^{s} \sqrt{\lambda_k} \xi_k(\omega) \psi_k(\boldsymbol{x}) \right\|_{L^2(\Omega, \mu; L^2(D))} = \left(\sum_{k=s+1}^{\infty} \lambda_k \right)^{1/2}.$$

- The random variables ξ_k are centered and uncorrelated, but not necessarily independent.
- If the random field $a(\boldsymbol{x}, \omega)$ is Gaussian – by definition, this means that $(a(\boldsymbol{x}_1, \omega), \ldots, a(\boldsymbol{x}_k, \omega))$ is a multivariate Gaussian random variable for all $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in D$, $k \in \mathbb{N}$ – then the random variables ξ_k are independent.
- The KL expansion is an effective method of representing *input* random fields when their covariance structure is known. If the (infinite-dimensional) input random field has a known covariance (which satisfies the conditions of Mercer's theorem), then we can use the KL expansion to find a finite-dimensional approximation, optimal in the mean-square error sense.

Modeling assumptions

In engineering and practical applications, the idea is that we have some *a priori* knowledge/belief that the unknown input random field is distributed according to some known probability distribution with a known covariance.

- If the input random field is Gaussian with a known, nice covariance function[†], then we use the KL expansion to find a reasonable finite-dimensional approximation of true input. Since the KL expansion decorrelates the stochastic variables, and uncorrelated jointly Gaussian random variables are independent, we can exploit the independence of the stochastic variables to parameterize the model problem.
- If the input random field is *not Gaussian*, then the stochastic variables in the KL expansion are uncorrelated *but not necessarily independent*. For the purposes of mathematical analysis, we typically assume that the random variables in the input random field are independent so that we can parameterize the model problem. (Transforming dependent random variables into independent random variables can be done using, e.g., the Rosenblatt transformation, but this is computationally expensive.)

Note especially that in the Gaussian setting we do not need to make any "extra" effort to ensure the independence of the stochastic variables in the KL expansion.

[†] Matérn covariance is an especially popular choice.

Example (Lognormal input random field)

Let $D \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a Lipschitz domain and consider the PDE problem

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{x}, \omega) \nabla u(\boldsymbol{x}, \omega)) = f(\boldsymbol{x}) & \text{for } \boldsymbol{x} \in D, \\ u(\cdot, \omega)|_{\partial D} = 0, \end{cases}$$

where $f : D \to \mathbb{R}$ is a fixed (deterministic) source term. We can model a lognormally distributed random diffusion coefficient $a \colon D \times \Omega \to \mathbb{R}$ using the KL expansion, e.g., as

$$a(\boldsymbol{x}, \omega) = a_0(\boldsymbol{x}) \exp\left(\sum_{k=1}^{\infty} y_k(\omega) \psi_k(\boldsymbol{x}) \right), \quad y_k \sim \mathcal{N}(0, 1),$$

where $a_0 \in L^\infty(D)$ is such that $a_0(\boldsymbol{x}) > 0$ and the random variables y_k are uncorrelated (and thus independent in the Gaussian case).

Due to the independence, we can consider the above as a *parametric PDE* with $a(\boldsymbol{x}, \boldsymbol{y}) \equiv a(\boldsymbol{x}, \boldsymbol{y}(\omega))$ and $u(\boldsymbol{x}, \boldsymbol{y}) \equiv u(\boldsymbol{x}, \boldsymbol{y}(\omega))$, where (formally) $\boldsymbol{y} \in \mathbb{R}^{\mathbb{N}}$ is a *parametric vector* endowed with a product Gaussian measure.

Example (Uniform and affine input random field)

Let $D \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a Lipschitz domain, $f : D \to \mathbb{R}$ is a fixed (deterministic) source term, and consider the PDE problem

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{x}, \omega) \nabla u(\boldsymbol{x}, \omega)) = f(\boldsymbol{x}) & \text{for } \boldsymbol{x} \in D, \\ u(\cdot, \omega)|_{\partial D} = 0. \end{cases}$$

We can model a uniformly distributed random diffusion coefficient $a : D \times \Omega \to \mathbb{R}$ using the KL expansion, e.g., as

$$a(\boldsymbol{x}, \omega) = a_0(\boldsymbol{x}) + \sum_{k=1}^{\infty} y_k(\omega) \psi_k(\boldsymbol{x}), \quad y_k \sim \mathcal{U}(-\tfrac{1}{2}, \tfrac{1}{2}),$$

where the random variables y_k are uncorrelated. *Unlike the Gaussian setting, the random variables y_k are generally not independent!*

In numerical analysis, the random variables y_k are often **assumed** to be independent – this allows us to consider the above as a parametric PDE with $a(\boldsymbol{x}, \boldsymbol{y}) \equiv a(\boldsymbol{x}, \boldsymbol{y}(\omega))$ and $u(\boldsymbol{x}, \boldsymbol{y}) \equiv u(\boldsymbol{x}, \boldsymbol{y}(\omega))$, where $\boldsymbol{y} \in [-\frac{1}{2}, \frac{1}{2}]^{\mathbb{N}}$ is a *parametric vector* endowed with a uniform probability measure.

To estimate the statistical response, note that in the *lognormal model* the expected value of the PDE solution is given by

$$\mathbb{E}[u(\boldsymbol{x},\cdot)] = \lim_{s\to\infty} \int_{\mathbb{R}^s} u(\boldsymbol{x},\boldsymbol{y}) \prod_{j=1}^{s} \frac{\mathrm{e}^{-\frac{1}{2}y_j^2}}{\sqrt{2\pi}}\,\mathrm{d}\boldsymbol{y}$$

while in the *affine and uniform model* the expected value of the PDE solution is given by

$$\mathbb{E}[u(\boldsymbol{x},\cdot)] = \lim_{s\to\infty} \int_{[-1/2,1/2]^s} u(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.$$

- In practice, we need to truncate these infinite-dimensional integrals into finite-dimensional ones, incurring the so-called *dimension truncation error*. Since the PDE is solved numerically using the finite element method, this also incurs a *finite element discretization error*.
- To compute the resulting high-dimensional integrals for the dimensionally-truncated, finite element discretized PDE solution we use a *quasi-Monte Carlo (QMC) method*.

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 chosen deterministically.

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

Lattice rules

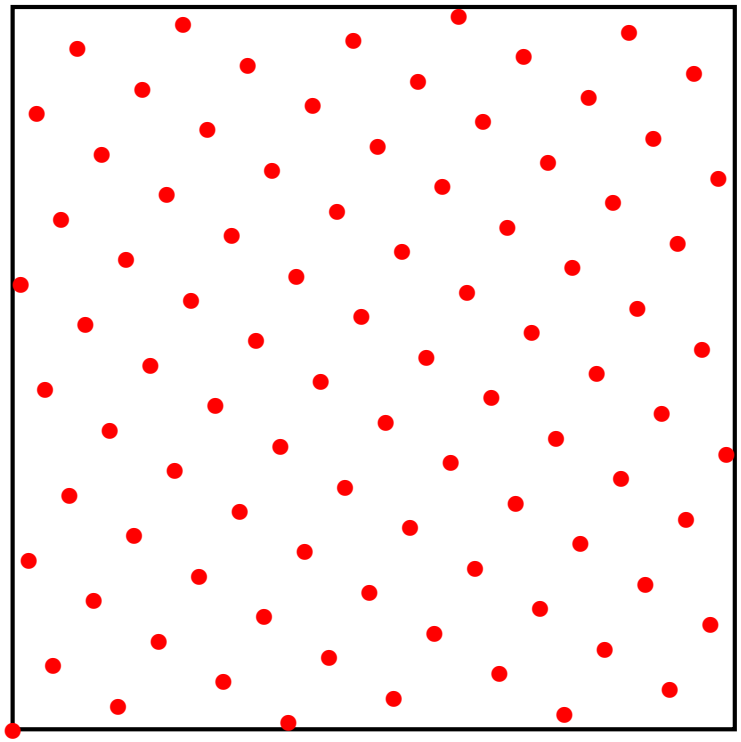
Rank-1 lattice rules

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

have the points

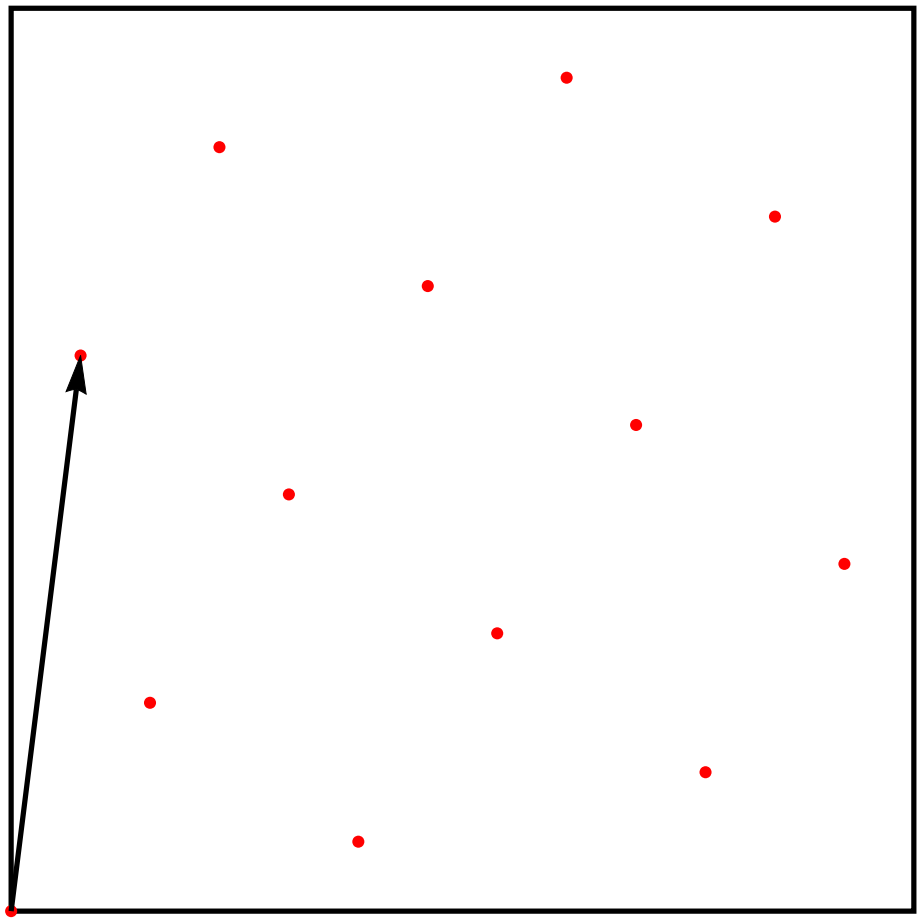
$$\boldsymbol{t}_i = \mathrm{mod}\left(\frac{i\boldsymbol{z}}{n}, 1\right), \quad i \in \{1, \ldots, n\},$$

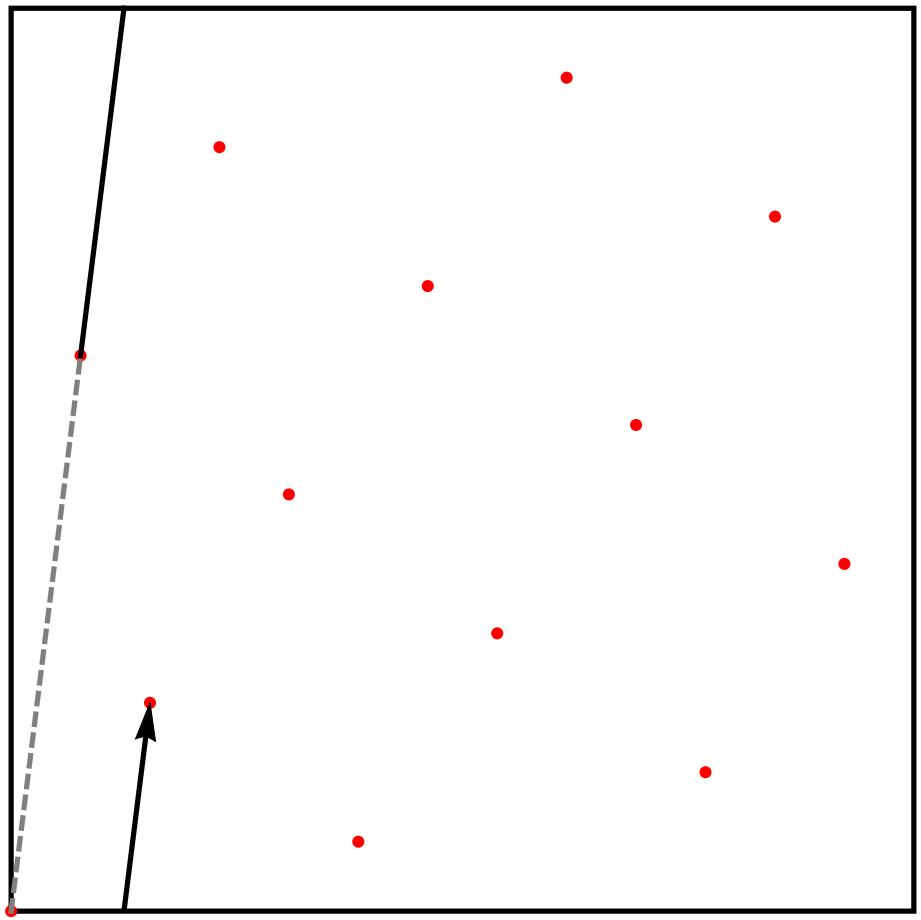
where the entire point set is determined by the *generating vector* $\boldsymbol{z} \in \mathbb{N}^s$, with all components *coprime* to n.

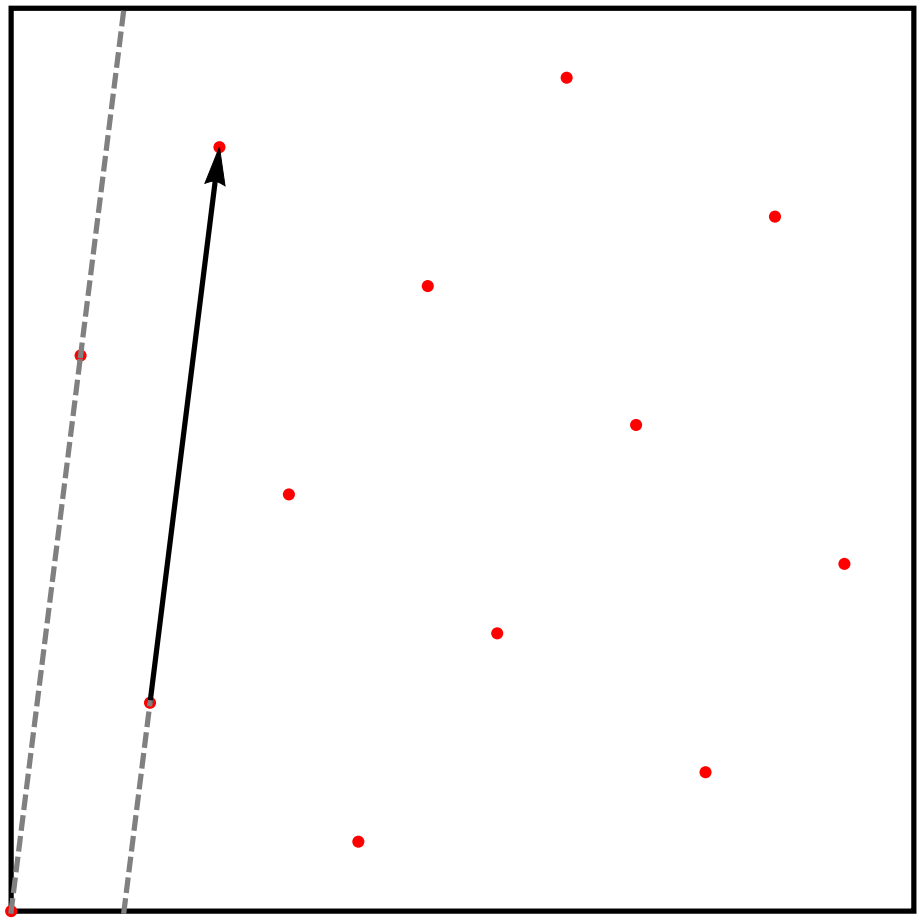


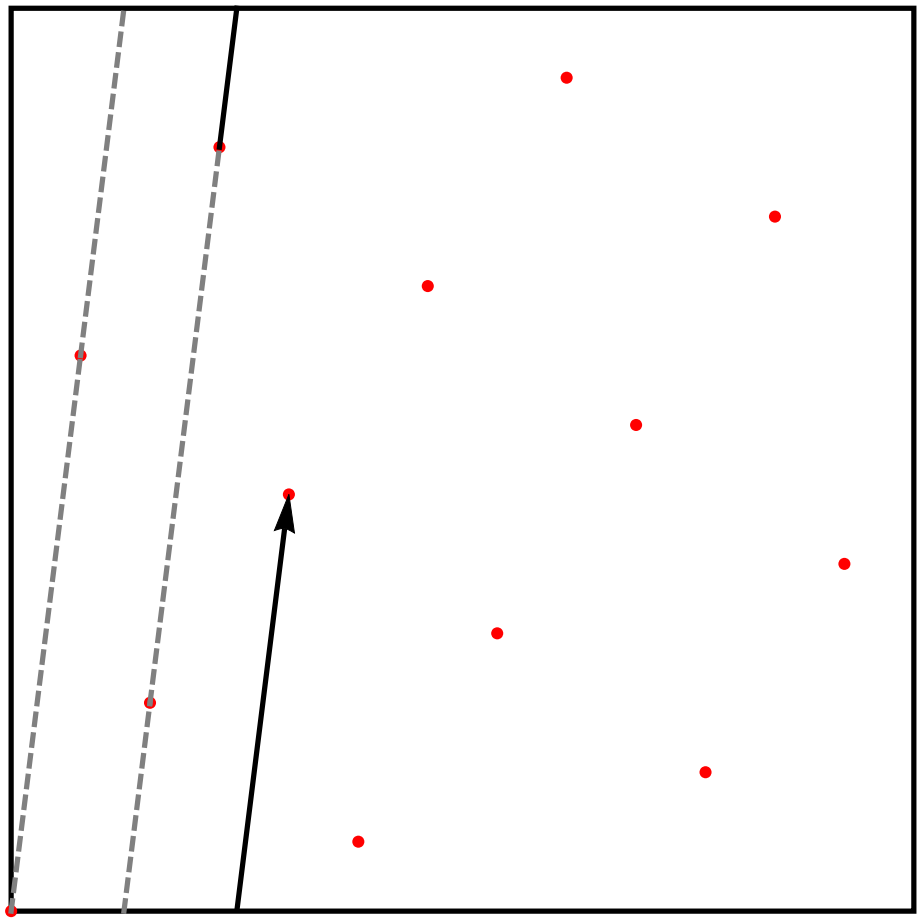
Lattice rule with $\boldsymbol{z} = (1, 55)$ and $n = 89$ nodes in $[0, 1]^2$

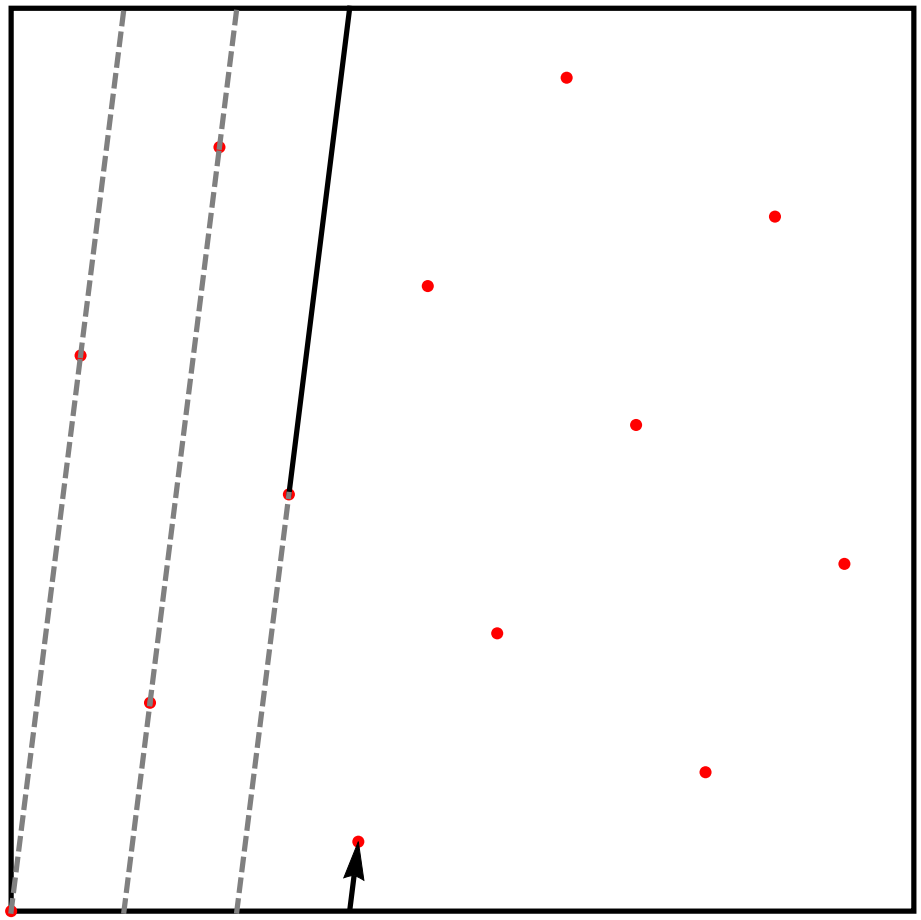
The quality of the lattice rule is determined by the choice of $\boldsymbol{z}$.

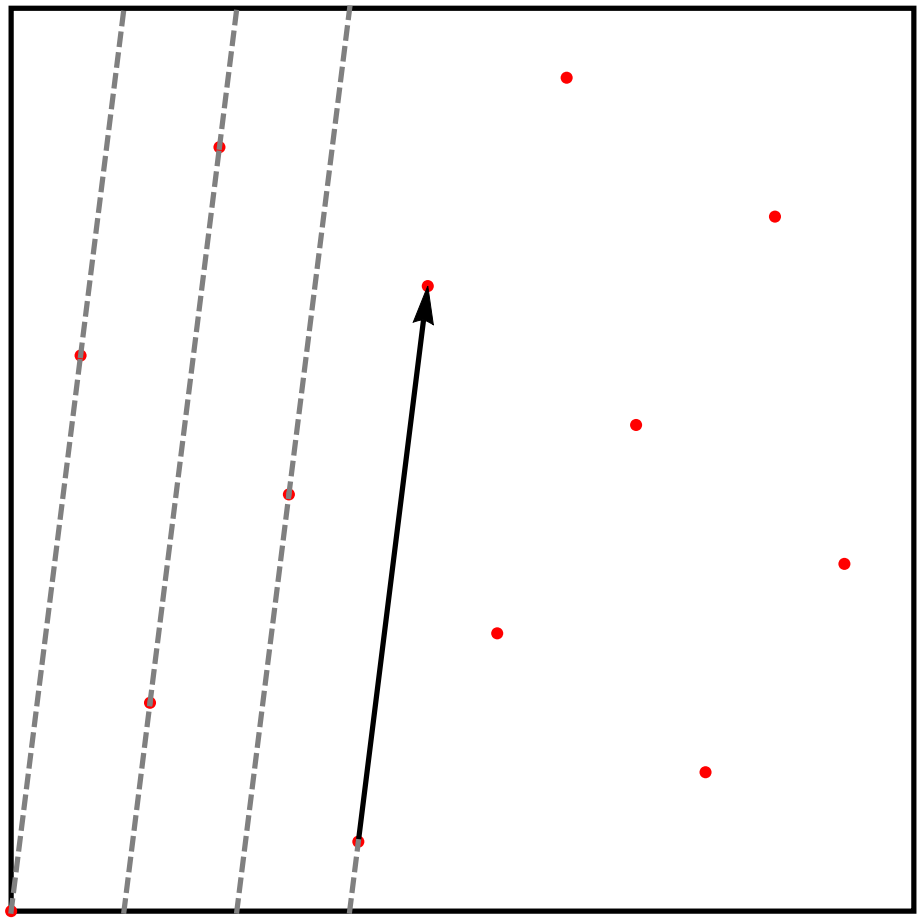


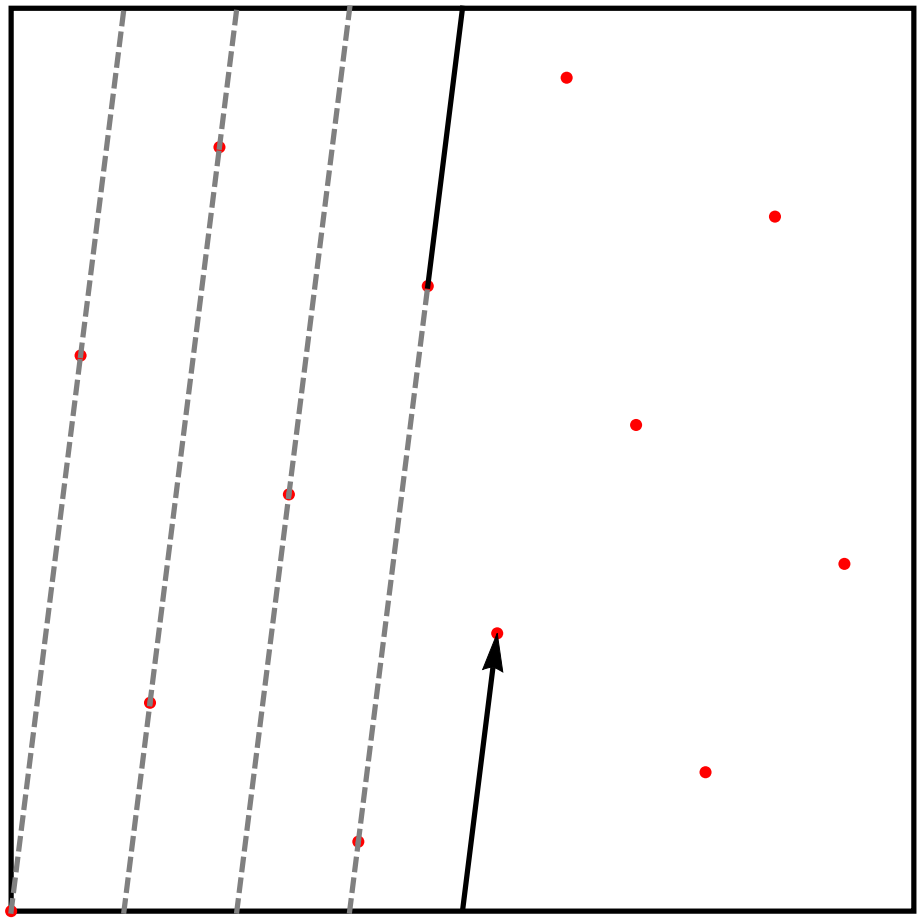


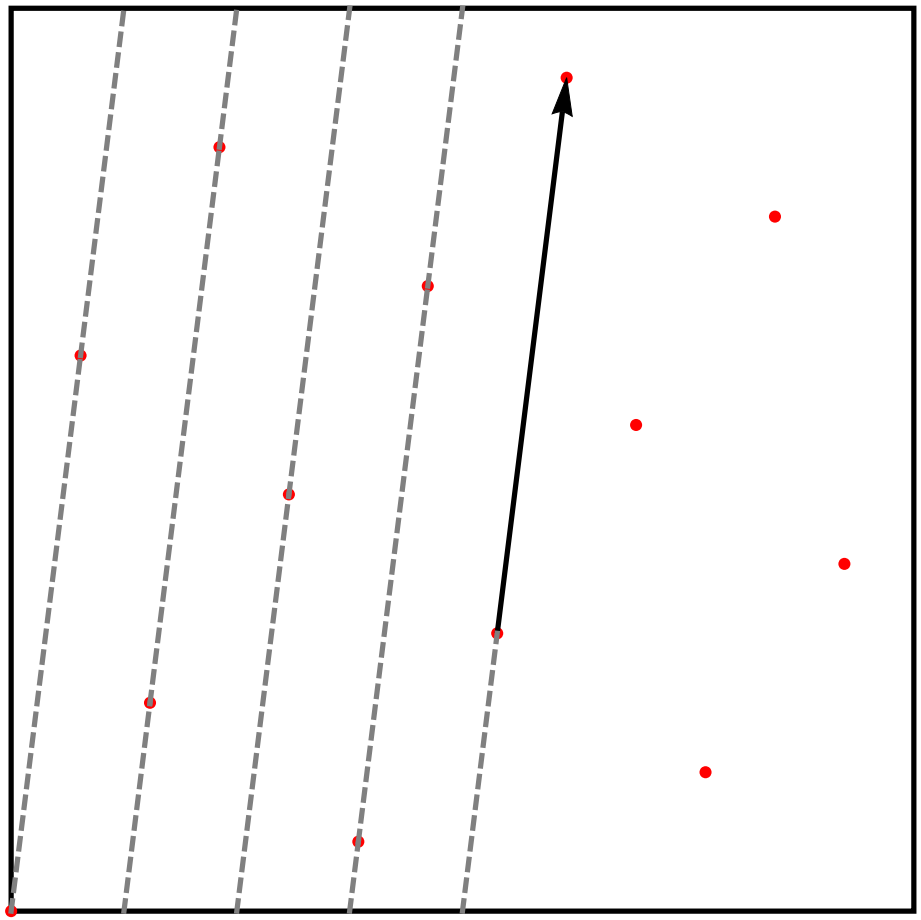


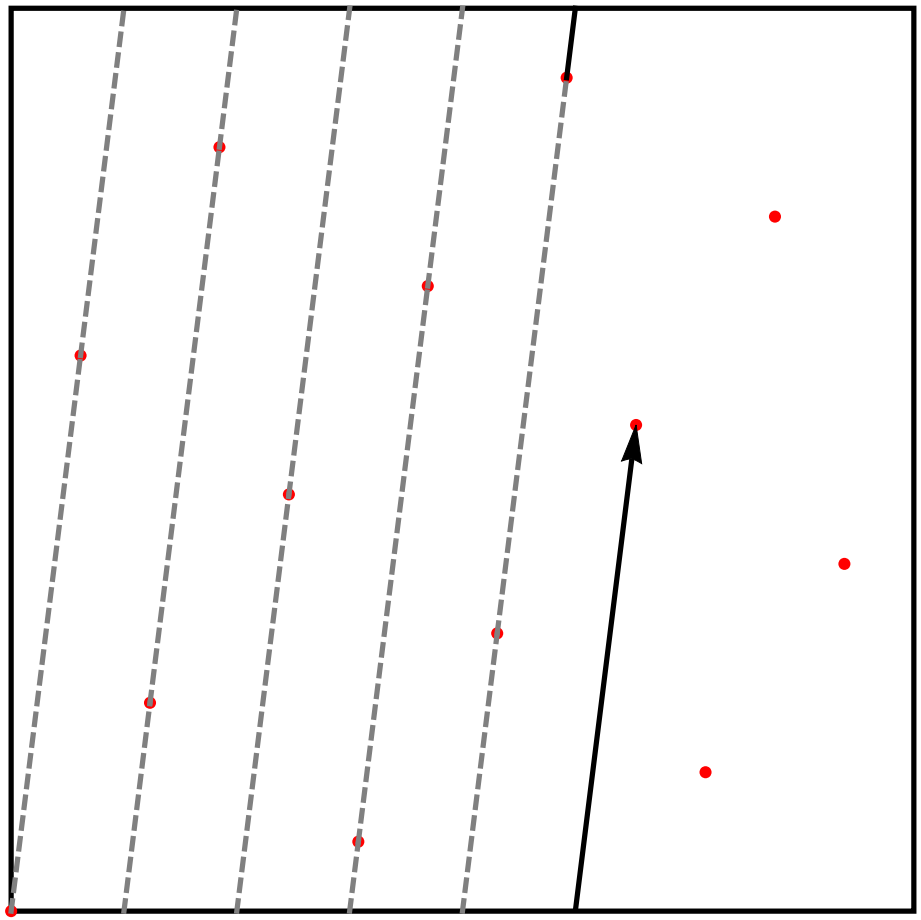


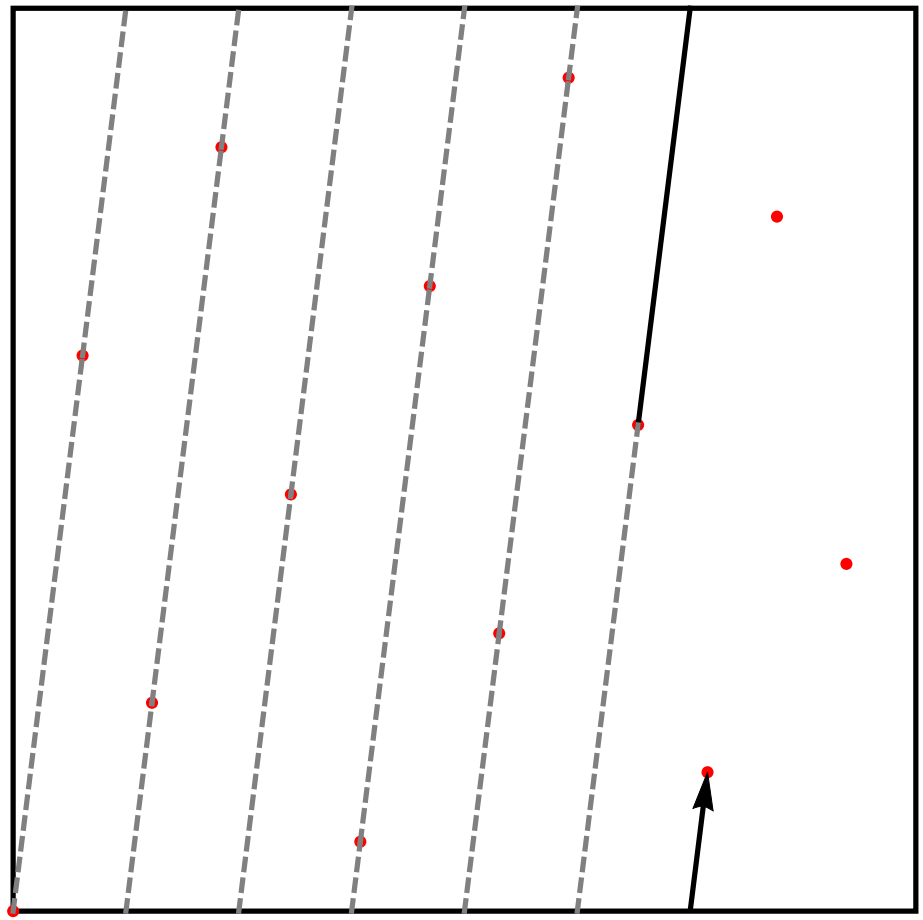


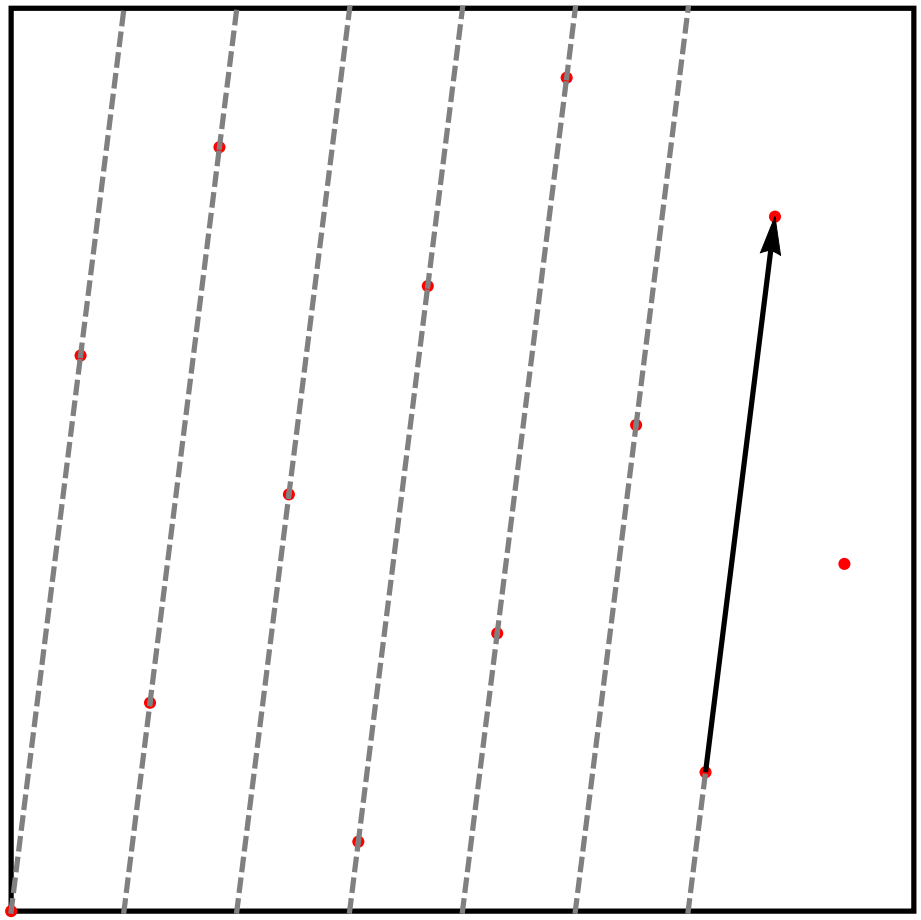


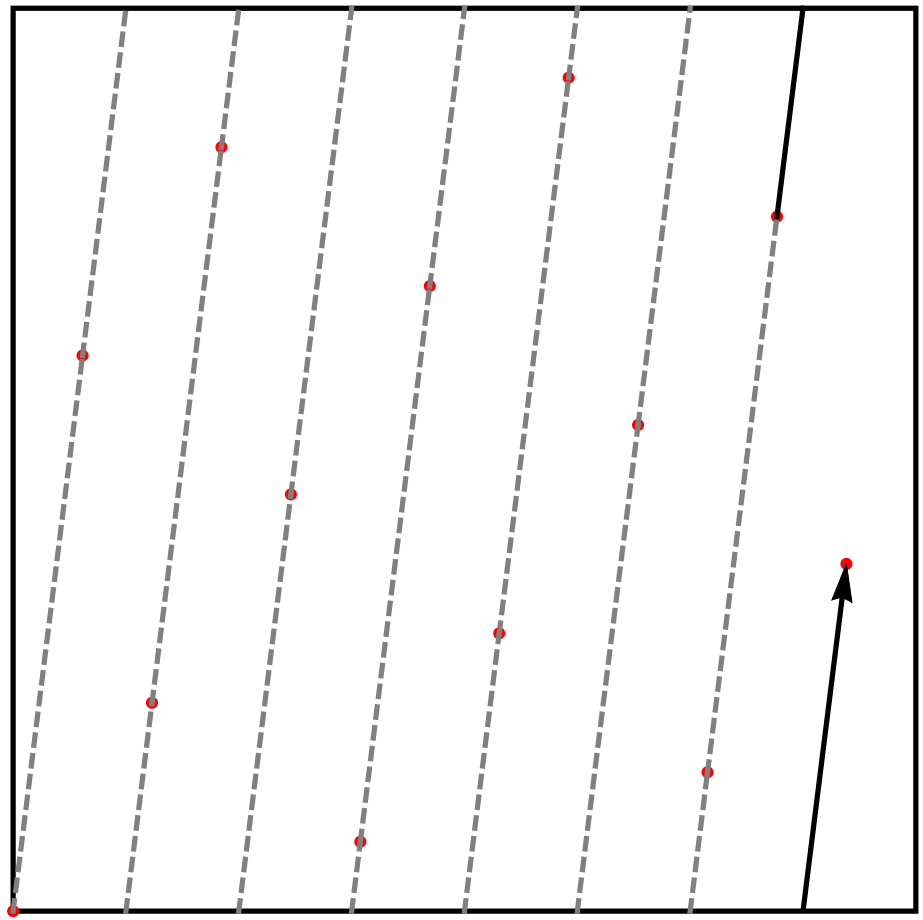












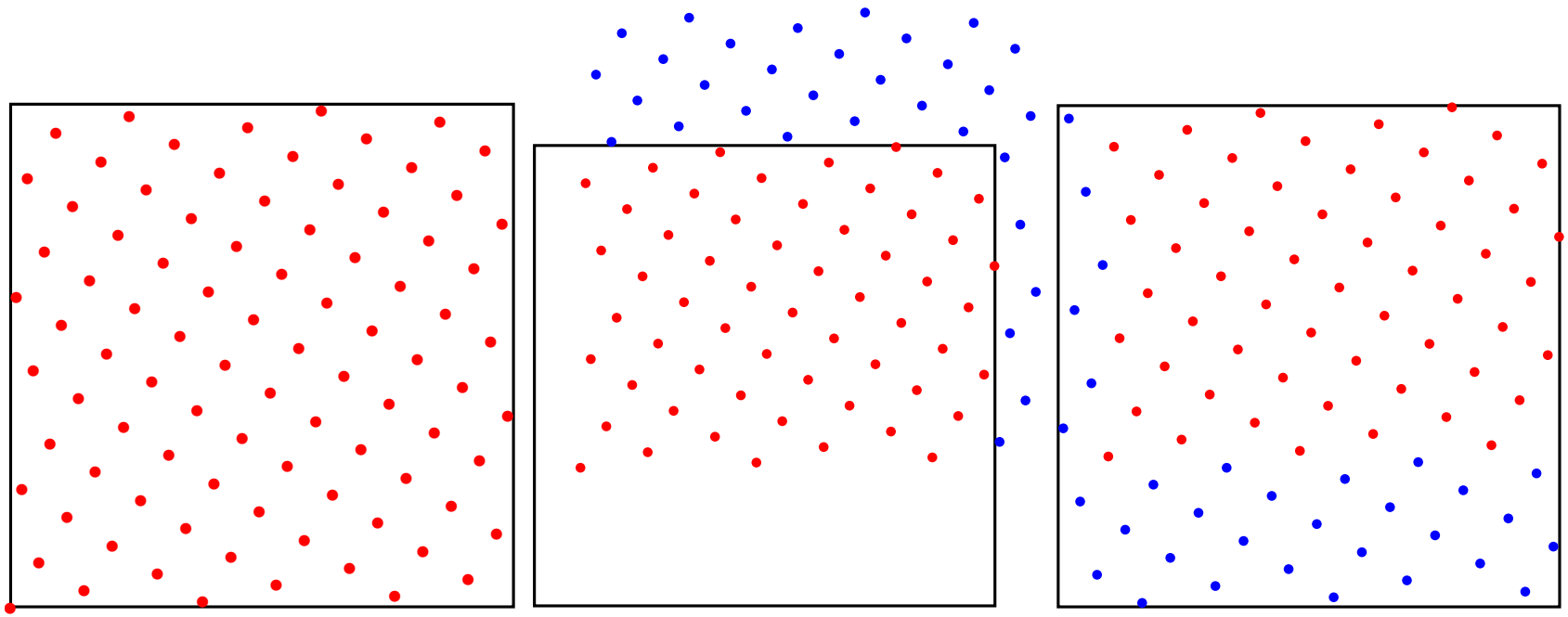
Randomly shifted lattice rules

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

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

Then

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

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

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

Error bound (one random shift):

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

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

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

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

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

and (squared) worst case error

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

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 $z_1 = 1$ (this is a freebie since $P(1) = P(z)$ for all $z \in \mathbb{N}$);
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 $P(z_1, z_2, z_3)$

   $\vdots$

- The CBC algorithm is a *greedy algorithm*: in general, it will not find the generating vector $\boldsymbol{z}$ that minimizes $P(\boldsymbol{z})$. However, it can be shown that the generating vector obtained by the CBC algorithm satisfies an error bound (see next slide).
- For generic $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u} \subseteq \{1:s\}}$, evaluating $P(\boldsymbol{z}) = P(\boldsymbol{\gamma}, \boldsymbol{z})$ takes $\mathcal{O}(2^s)$ operations. For an efficient implementation, it is desirable that the weights $\boldsymbol{\gamma}$ can be characterized by an expression that does not contain too many degrees of freedom.
- Efficient implementation using FFT! (QMC4PDE, QMCPy, etc.)

Theorem (CBC error bound)

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

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

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.
- Fast CBC works for any (composite) number $n \ge 2$, but the implementation is more involved when n is not prime.

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

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

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 (**??**).
- Using the optimized weights $\boldsymbol{\gamma}$ as input, use the CBC algorithm to find a generating vector which *satisfies* the error bound (**??**).

Remarks:

- If n is prime, then $\frac{1}{\varphi(n)} = \frac{1}{n-1}$. If $n = 2^k$, then $\frac{1}{\varphi(n)} = \frac{2}{n}$. For general (composite) $n \ge 3$, $\frac{1}{\varphi(n)} \le \frac{\mathrm{e}^{\gamma} \log\log n + \frac{3}{\log\log n}}{n}$, where $\gamma = 0.57721566\ldots$ (Euler–Mascheroni constant).
- The optimal convergence rate close to $\mathcal{O}(n^{-1})$ is obtained with $\lambda \to 1/2$, but $\lambda = 1/2$ is not permitted since $\zeta(2\lambda) \overset{\lambda \to 1/2+}{\to} \infty$.

Example: applying QMC theory for a simplified parametric PDE

Let $D \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a convex, bounded Lipschitz domain and consider the following (simplified!) elliptic PDE

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{y}) \nabla u(\boldsymbol{x}, \boldsymbol{y})) = f(\boldsymbol{x}), & \boldsymbol{x} \in D,\ \boldsymbol{y} \in [-1/2, 1/2]^s, \\ u(\boldsymbol{x}, \boldsymbol{y}) = 0, & \boldsymbol{x} \in \partial D,\ \boldsymbol{y} \in [-1/2, 1/2]^s, \end{cases}$$

where the source term $f \in L^2(D)$ is fixed and

$$a(\boldsymbol{y}) := 1 + \sum_{j=1}^{s} \beta_j y_j, \quad y_j \in [-1/2, 1/2],$$

where $\beta_j \geq 0$ are assumed to be *constants* for all $j \geq 1$ (i.e., independent of $\boldsymbol{x}$) s.t. $a(\boldsymbol{y}) \geq a_{\min} > 0$ for all $\boldsymbol{y} \in [-1/2, 1/2]^s$ and $\sum_{j=1}^{\infty} \beta_j^p < \infty$ for some $p \in (0, 1)$. Due to the linearity of the PDE problem, we can write

$$u(\boldsymbol{x}, \boldsymbol{y}) = \frac{g(\boldsymbol{x})}{1 + \sum_{j=1}^{s} \beta_j y_j}, \quad \text{where } \begin{cases} -\Delta g(\boldsymbol{x}) = f(\boldsymbol{x}),\ \boldsymbol{x} \in D, \\ g|_{\partial D} = 0. \end{cases}$$

Note that the Poisson problem has a continuous solution $g \in C(D)$.

Clearly,

$$\mathbb{E}[u(\boldsymbol{x},\cdot)] = g(\boldsymbol{x}) \int_{[-1/2,1/2]^s} \underbrace{\frac{1}{1+\sum_{j=1}^{s}\beta_j y_j}}_{=:F(\boldsymbol{y})} \,\mathrm{d}\boldsymbol{y}.$$

(Note the similarity to exercise 2 of week 8!)

Steps of QMC analysis:

- Estimate the (parametric) derivatives $\partial^{\boldsymbol{\nu}} F(\boldsymbol{y})$.
- Using the above, estimate $\|F(\cdot - \frac{\mathbf{1}}{2})\|_{s,\gamma}$.
- Plug the weighted Sobolev norm into QMC error bound and choose the weights $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u}\subseteq\{1:s\}}$ to minimize the resulting error bound.
  - The resulting weights are used as inputs to the CBC algorithm. The generating vector (and the resulting randomly shifted QMC point set) are guaranteed to satisfy the rigorous CBC error bound.
- Analysis: is the coefficient of the CBC error bound independent of the dimension s with the chosen weights?

Step 1: *Parametric regularity.* It is not difficult to see that

$$\frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} F(\boldsymbol{y}) = |\mathfrak{u}|! F(\boldsymbol{y})^{|\mathfrak{u}|+1} \prod_{j \in \mathfrak{u}} (-\beta_j) \quad \text{for all } \mathfrak{u} \subseteq \{1:s\}.$$

Exploiting the fact that we assumed before that $1 + \sum_{j=1}^{s} \beta_j y_j \geq a_{\min} > 0$ for all $\boldsymbol{y} \in [-1/2, 1/2]^s$, we can define

$$b_j := \frac{\beta_j}{a_{\min}} \quad \text{for all } j \geq 1,$$

and estimate the parametric regularity of the first order mixed partial derivatives as

$$\left| \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} F(\boldsymbol{y}) \right| \leq \frac{1}{a_{\min}} |\mathfrak{u}|! \prod_{j \in \mathfrak{u}} b_j \quad \text{for all } \mathfrak{u} \subseteq \{1:s\}.$$

Step 2: *Estimate the weighted Sobolev norm.* It is easy to see that

$$\|F(\cdot - \tfrac{\mathbf{1}}{\mathbf{2}})\|_{s,\gamma}^2 \lesssim \sum_{\mathfrak{u}\subseteq\{1:s\}} \frac{(|\mathfrak{u}|!)^2}{\gamma_{\mathfrak{u}}} \prod_{j\in\mathfrak{u}} b_j^2.$$

Step 3: Plugging this into the CBC error bound

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

yields

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

(We have separated the dependence on the number of QMC nodes n since this is unaffected by the choice of weights. The weights only affect the constant in the error bound, which we try to minimize next.)

Step 4: *Choosing the weights*. Note that the square of the objective functional has the form

$$g(\boldsymbol{\gamma}) := \left(\sum_{\mathfrak{u}} \alpha_{\mathfrak{u}} \gamma_{\mathfrak{u}}^{\lambda} \right)^{1/\lambda} \left(\sum_{\mathfrak{u}} \beta_{\mathfrak{u}} \gamma_{\mathfrak{u}}^{-1} \right),$$

which is minimized by

$$\gamma_{\mathfrak{u}} := c \left(\frac{\beta_{\mathfrak{u}}}{\alpha_{\mathfrak{u}}} \right)^{1/(1+\lambda)} \quad \text{for arbitrary } c > 0.$$

(In fact, with $c = 1$, the minimizer is equivalent to setting the summands equal: $\alpha_{\mathfrak{u}} \gamma_{\mathfrak{u}}^{\lambda} = \beta_{\mathfrak{u}} \gamma_{\mathfrak{u}}^{-1}$.)
Thus the minimizing weights for our problem are the *product-and-order (POD) dependent* weights:

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

(The POD form is important since it doesn't contain "too many degrees of freedom": the cost of fast CBC used to find the generating vector satisfying the CBC error bound is $\mathcal{O}(s\, n \log n + s^2 n)$ with these weights.)

Step 5: Plugging the optimized POD weights into the QMC error bound results in

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

where

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

In complete analogy to the 11^{th} lecture, we have the following:

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. $C(s,\boldsymbol{\gamma},\lambda) \le C(\boldsymbol{\gamma},\lambda) < \infty$.

Using randomly shifted rank-1 lattice rules to estimate the integral

$$\int_{[-1/2,1/2]} F(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \quad F(\boldsymbol{y}) := \frac{1}{1+\sum_{j=1}^{s}\beta_j y_j},$$

we can conclude the following:

For arbitrary $\delta \in (0,1/2)$, we can choose the POD weights

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

as inputs to the CBC algorithm to obtain a generating vector. If the number of QMC nodes n is prime or a prime power, then the resulting randomly shifted rank-1 lattice rule satisfies the root-mean-square error bound

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}|I_s F - Q^{\boldsymbol{\Delta}}_{n,s}F|^2} \lesssim n^{\max\{-1/p+1/2,-1+\delta\}}, \tag{3}$$

where the implied coefficient is independent of the dimension s.
Note that this rate is always better than Monte Carlo, but cannot exceed linear convergence $\mathcal{O}(n^{-1})$ (i.e., double the Monte Carlo rate).

Uniform and affine model

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* $\psi_j : D \to \mathbb{R}$ 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 $p \in (0, 1)$.

Proposition (Parametric regularity for the uniform and affine model)

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* $v \in H_0^1(D)$.

This parametric regularity bound is valid also for the dimensionally-truncated finite element solution $u_{s,h}$. If $G : H_0^1(D) \to \mathbb{R}$ is a bounded linear functional and we define $F(\boldsymbol{y}) := G(u_{s,h}(\cdot, \boldsymbol{y} - \frac{1}{2}))$ for $\boldsymbol{y} \in [0,1]^s$, then

$$\|F\|_{s,\boldsymbol{\gamma}}^2 \lesssim \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} (|\mathfrak{u}|!)^2 \prod_{j \in \mathfrak{u}} b_j^2,$$

and using the POD weights (**??**) as inputs to the CBC algorithm yields a randomly shifted rank-1 lattice rule satisfying the R.M.S. error

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}} |I_s F - Q_{n,s}^{\boldsymbol{\Delta}} F|^2} \lesssim n^{\max\{-1/p+1/2, -1+\delta\}},$$

where the constant is independent of the dimension.

Of course, the truncation of the input random series and the finite element discretization incur additional errors.

- If $\|\psi_1\|_{L^\infty(D)} \geq \|\psi_2\|_{L^\infty(D)} \cdots$, then the error resulting from the dimension truncation has order $\mathcal{O}(s^{-2/p+1})$, where the constant is independent of s.
- If $D \subset \mathbb{R}^d$ is a bounded, convex polyhedron, a_0 and ψ_j are Lipschitz for all $j \geq 1$ with $\sum_{j=1}^{\infty} \|\psi_j\|_{W^{1,\infty}(D)} < \infty$, and $G : L^2(D) \to \mathbb{R}$ is a bounded linear functional, then—if the FE mesh has been obtained from an initial, regular triangulation of D by recursive, uniform bisection of simplices—the L^2 finite element error has order $\mathcal{O}(h^2)$, where $h > 0$ is the mesh size and the implied constant is independent of $\boldsymbol{y}$, s, and h.

Lognormal model

Lognormal model: let $D \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded Lipschitz domain, and let $f \in H^{-1}(D)$. Let $\psi_j \in L^\infty(D)$ and $b_j := \|\psi_j\|_{L^\infty}$ for $j \in \mathbb{N}$ such that $\sum_{j=1}^\infty b_j < \infty$, and set

$$U_{\boldsymbol{b}} := \left\{ \boldsymbol{y} \in \mathbb{R}^{\mathbb{N}} : \sum_{j=1}^{\infty} b_j |y_j| < \infty \right\}.$$

Consider the problem of finding, for all $\boldsymbol{y} \in U$, $u(\cdot, \boldsymbol{y}) \in H_0^1(D)$ such that

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

where the diffusion coefficient is assumed to have the parameterization

$$a(\boldsymbol{x}, \boldsymbol{y}) := a_0(\boldsymbol{x}) \exp\left(\sum_{j=1}^{\infty} y_j \psi_j(\boldsymbol{x}) \right), \quad \boldsymbol{x} \in D, \ \boldsymbol{y} \in U_{\boldsymbol{b}},$$

where $a_0 \in L^\infty(D)$ is such that $a_0(\boldsymbol{x}) > 0$, $\boldsymbol{x} \in D$.

Standing assumptions for the lognormal model

(B1) We have $a_0 \in L^\infty(D)$ and $\sum_{j=1}^\infty b_j < \infty$.

(B2) For every $\boldsymbol{y} \in U_{\boldsymbol{b}}$, the expressions $a_{max}(\boldsymbol{y}) := \max_{\boldsymbol{x} \in \overline{D}} a(\boldsymbol{x}, \boldsymbol{y})$ and $a_{\min}(\boldsymbol{y}) := \min_{\boldsymbol{x} \in \overline{D}} a(\boldsymbol{x}, \boldsymbol{y})$ are well-defined and satisfy $0 < a_{\min}(\boldsymbol{y}) \leq a(\boldsymbol{x}, \boldsymbol{y}) \leq a_{\max}(\boldsymbol{y}) < \infty$.

(B3) $\sum_{j=1}^\infty b_j^p < \infty$ for some $p \in (0, 1)$.

Remark: Note that in the lognormal case, $a(\boldsymbol{x}, \boldsymbol{y})$ can take values which are arbitrarily close to 0 or arbitrarily large. Thus, the best we can do is to find $\boldsymbol{y}$-dependent lower and upper bounds $a_{\min}(\boldsymbol{y})$ and $a_{\max}(\boldsymbol{y})$. This will lead to a $\boldsymbol{y}$-dependent *a priori* bound and, consequently, $\boldsymbol{y}$-dependent parametric regularity bounds. This will make the QMC analysis more involved, leading one to consider "special" weighted, unanchored Sobolev spaces.

In this setting, we have

$$I_s(F) := \int_{\mathbb{R}^s} F(\boldsymbol{y}) \prod_{j=1}^{s} \phi(y_j) \, \mathrm{d}\boldsymbol{y} = \int_{(0,1)^s} F(\Phi^{-1}(\boldsymbol{w})) \, \mathrm{d}\boldsymbol{w}.$$

where $\phi(y) := \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}y^2}$ is the probability density function of $\mathcal{N}(0,1)$ and $\Phi^{-1}(\boldsymbol{w}) = [\Phi^{-1}(w_1), \ldots, \Phi^{-1}(w_s)]^{\mathrm{T}}$ denotes the corresponding (componentwise) inverse cumulative distribution function. We use the randomly shifted QMC rules

$$Q_{n,s}^{\boldsymbol{\Delta}_r}(F) = \frac{1}{n} \sum_{k=1}^{n} F(\Phi^{-1}(\{\boldsymbol{t}_k + \boldsymbol{\Delta}_r\})),$$

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

where we have R independent random shifts $\boldsymbol{\Delta}_1, \ldots, \boldsymbol{\Delta}_R$ drawn from $\mathcal{U}([0,1]^s)$, $\boldsymbol{t}_k := \{\frac{k\boldsymbol{z}}{n}\}$, with generating vector $\boldsymbol{z} \in \mathbb{N}^s$.

The appropriate function space for unbounded integrands is a "special" weighted, unanchored Sobolev space equipped with the norm

$$\|F\|_{s,\boldsymbol{\gamma}} = \Bigg[\sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \int_{\mathbb{R}^{|\mathfrak{u}|}} \Bigg(\int_{\mathbb{R}^{s-|\mathfrak{u}|}} \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} F(\boldsymbol{y}) \Bigg(\prod_{j \in \{1:s\} \setminus \mathfrak{u}} \phi(y_j) \Bigg) \mathrm{d}\boldsymbol{y}_{-\mathfrak{u}} \Bigg)^2 \times \Bigg(\prod_{j \in \mathfrak{u}} \varpi_j^2(y_j) \Bigg) \mathrm{d}\boldsymbol{y}_{\mathfrak{u}} \Bigg]^{1/2}$$

where we have the weights

$$\varpi_j^2(y) := \exp(-2\alpha_j |y_j|), \quad \alpha_j > 0.$$

Theorem (Graham, Kuo, Nichols, Scheichl, Schwab, Sloan (2015))

Let F belong to the special weighted space over $\mathbb{R}^s$ with weights $\boldsymbol{\gamma}$, with ϕ being the standard normal density, and the weight functions ϖ_j defined as above. A randomly shifted lattice rule in s dimensions with n being a prime power can be constructed by a CBC algorithm such that

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}|I_s F - Q_{n,s}^{\boldsymbol{\Delta}} F|^2} \le \left(\frac{2}{n} \sum_{\varnothing \ne \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \prod_{j \in \mathfrak{u}} \varrho_j(\lambda)\right)^{1/(2\lambda)} \|F\|_{s,\boldsymbol{\gamma}},$$

where $\lambda \in (1/2, 1]$ and

$$\varrho_j(\lambda) = 2\left(\frac{\sqrt{2\pi}\exp(\alpha_j^2/\eta_*)}{\pi^{2-2\eta_*}(1-\eta_*)\eta_*}\right)^{\lambda} \zeta(\lambda + \tfrac{1}{2}) \quad \text{and} \quad \eta_* = \frac{2\lambda - 1}{4\lambda},$$

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

The steps for QMC analysis are the same as in the uniform case: (1) estimate $\|\cdot\|_{s,\boldsymbol{\gamma}}$ for a given integrand (2) find weights $\boldsymbol{\gamma}$ which minimize the upper bound (3) plug the weights into the new error bound and estimate the constant (which ideally can be bounded independently of s).

Proposition (Parametric regularity bound for the lognormal model)

For all $\boldsymbol{y} \in U_{\boldsymbol{b}}$ and $\boldsymbol{\nu} \in \mathscr{F}$, there holds

$$\|\partial^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_{H_0^1(D)} \leq \frac{C_P \|f\|_{L^2(D)}}{\min_{\boldsymbol{x} \in \overline{D}} a_0(\boldsymbol{x})} \frac{|\boldsymbol{\nu}|!}{(\log 2)^{|\boldsymbol{\nu}|}} \boldsymbol{b}^{\boldsymbol{\nu}} \prod_{j \geq 1} \exp(b_j |y_j|).$$

This parametric regularity bound is valid also for the dimensionally-truncated finite element solution $u_{s,h}$. If $G : H_0^1(D) \to \mathbb{R}$ is a bounded linear functional and $F(\boldsymbol{y}) := G(u_{s,h}(\cdot, \boldsymbol{y}))$ for $\boldsymbol{y} \in \mathbb{R}^s$, then

$$\|F\|_{s,\boldsymbol{\gamma}}^2 \leq \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{(|\mathfrak{u}|!)^2}{\gamma_{\mathfrak{u}}} \left(\prod_{j=1}^{s} 2 \exp(2b_j^2) \Phi(2b_j) \right) \left(\prod_{j \in \mathfrak{u}} \frac{b_j^2}{2(\log 2)^2 \exp(2b_j^2) \Phi(2b_j)(\alpha_j - b_j)} \right).$$

By choosing $\alpha_j = \frac{1}{2}(b_j + \sqrt{b^2 + 1 - \frac{1}{2\lambda}})$ and using the POD weights

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

as inputs to the CBC algorithm yields a randomly shifted rank-1 lattice rule satisfying the R.M.S. error

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}} |I_s F - Q_{n,s}^{\boldsymbol{\Delta}} F|^2} \lesssim n^{\max\{-1/p+1/2, -1+\delta\}},$$

where the constant is independent of the dimension.

Similarly to the uniform and affine setting, the truncation of the input random series and the finite element discretization incur a *dimension truncation error* and a *finite element discretization error*, respectively. However, the analysis is more complicated in the lognormal case and has been omitted.

Computational implementation

Consider the task of approximating $\int_{[0,1]^s} f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$ using a randomly shifted rank-1 lattice rule with R random shifts.

Once a generating vector $\boldsymbol{z} \in \mathbb{N}^s$ has been obtained for a given number n of QMC nodes and dimension s (using, e.g., the CBC algorithm), then:

for $r = 1, \ldots, R$, **do**
  draw $\boldsymbol{\Delta}^{(r)} \sim \mathcal{U}([0,1]^s)$;
  initialize $Q_r = 0$;
    for $i = 1, \ldots, n$, **do**
      set $\boldsymbol{t}_i = \mathrm{mod}\big(\frac{i\boldsymbol{z}}{n} + \boldsymbol{\Delta}^{(r)}, 1\big)$;
      set $Q_r = Q_r + f(\boldsymbol{t}_i)$;
    end for
  set $Q_r = Q_r/n$;
end for
return $\overline{Q} = \frac{Q_1 + \cdots + Q_R}{R}$;
(This is the QMC estimator with R random shifts.)

Remarks:

- If integrating
$$\int_{\mathbb{R}^s} f(\boldsymbol{y}) \prod_{j=1}^s \frac{\mathrm{e}^{-\frac{1}{2}y_j^2}}{\sqrt{2\pi}}\,\mathrm{d}\boldsymbol{y}$$
then use $\boldsymbol{t}_i = \Phi^{-1}\big(\mathrm{mod}\big(\frac{i\boldsymbol{z}}{n} + \boldsymbol{\Delta}^{(r)}, 1\big)\big)$, where Φ^{-1} is the (componentwise) inverse cumulative distribution function of $\mathcal{N}(0,1)$.
- The R.M.S. error can be estimated by
$$\text{R.M.S. error} \approx \sqrt{\frac{1}{R(R-1)} \sum_{r=1}^R (\overline{Q} - Q_r)^2}.$$

The end!