



Quasi-Monte Carlo for Bayesian inverse problems governed by PDEs

March 4, 2025

Vesa Kaarnioja (FU Berlin)

Thanks to collaborators: Laura Bazahica (LUT University), Ana Djurdjevac (FU Berlin), Max Orteu (FU Berlin), Lassi Roininen (LUT University), Claudia Schillings (FU Berlin), and André-Alexander Zepernick (FU Berlin)

Table of contents

Part I: Quasi-Monte Carlo methods

Part II: Parameterization of input uncertainty

Part III: Application to Bayesian inverse problems

Part I: Quasi-Monte Carlo methods

High-dimensional numerical integration

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

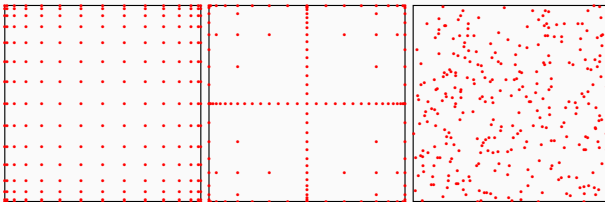


Figure 1: Tensor product grid, sparse grid, Monte Carlo nodes (not QMC rules)

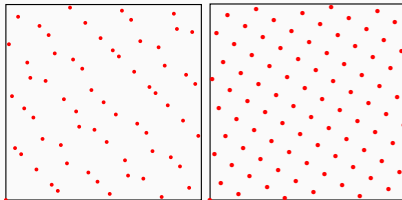


Figure 2: Sobol' points, lattice rule (examples of QMC rules)

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

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

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

The nodes $(\mathbf{t}_i)_{i=1}^n$ 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

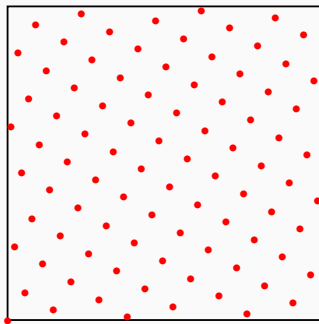
Rank-1 lattice rules

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

have the points

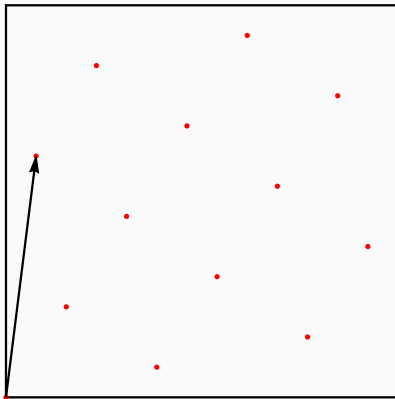
$$\mathbf{t}_i = \text{mod} \left(\frac{i\mathbf{z}}{n}, 1 \right), \quad i \in \{1, \dots, n\},$$

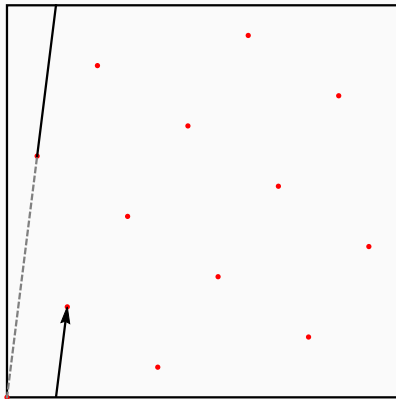
where the entire point set is determined by the *generating vector* $\mathbf{z} \in \mathbb{N}^s$, with all components *coprime* to n .

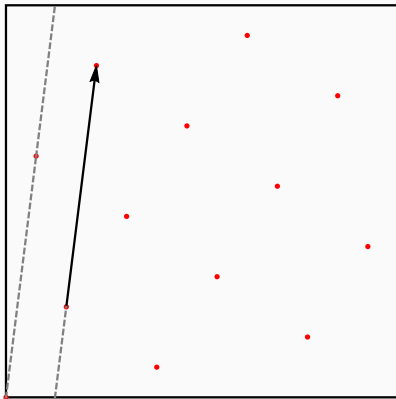


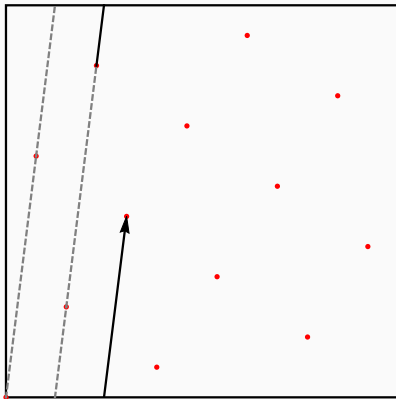
Lattice rule with $\mathbf{z} = (1, 55)$ and $n = 89$
nodes in $[0, 1]^2$

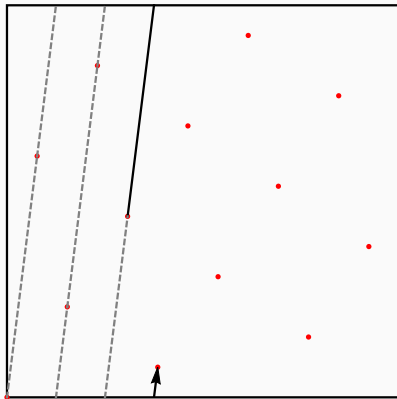
The quality of the lattice rule is determined by the choice of \mathbf{z} .

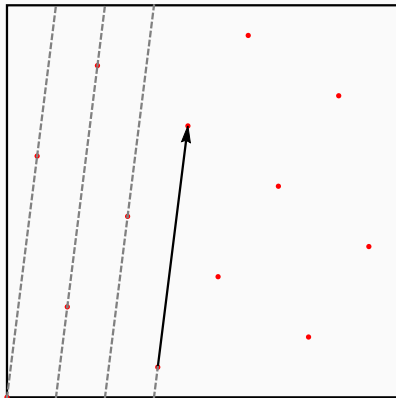


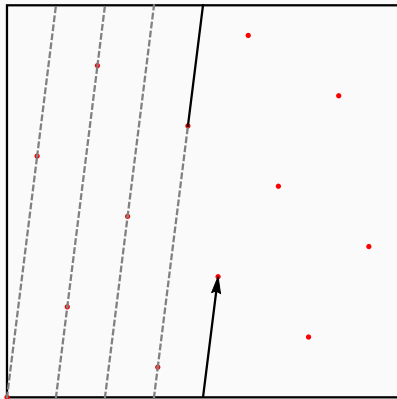


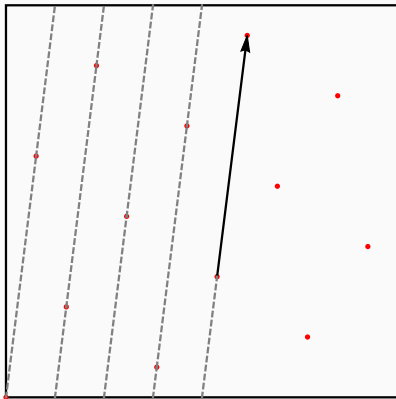


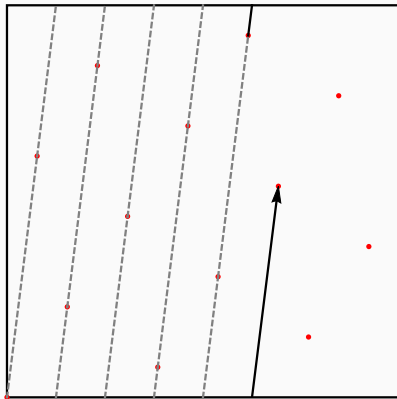


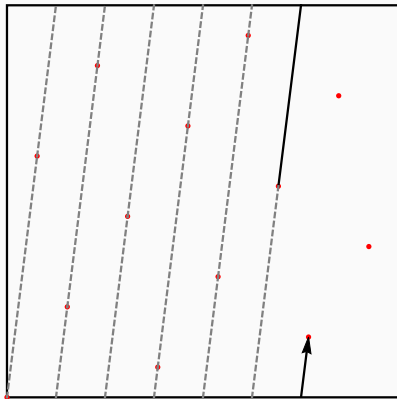


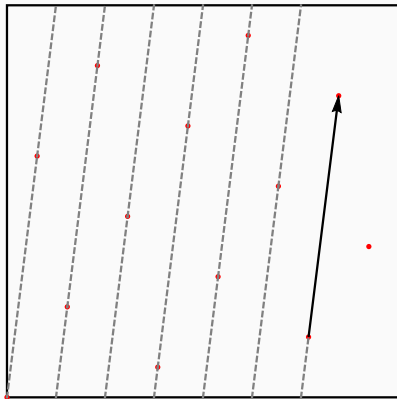


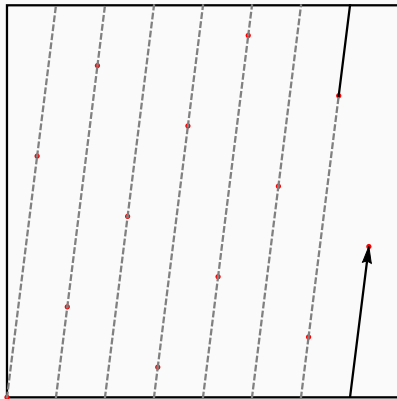












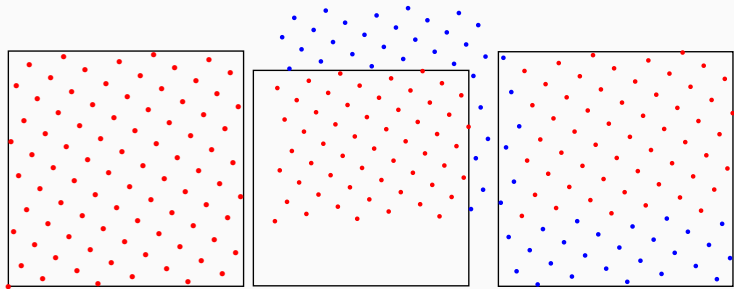
Randomly shifted lattice rules

Shifted rank-1 lattice rules have points

$$\mathbf{t}_i = \text{mod}\left(\frac{i\mathbf{z}}{n} + \mathbf{\Delta}, 1\right), \quad i \in \{1, \dots, n\}.$$

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

Use a number of random shifts for error estimation.



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

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

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

Then

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

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

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

Error bound (one random shift):

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

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

$$\sqrt{\mathbb{E}_{\Delta}[|I_s(f) - \bar{Q}_{s,n}(f)|^2]} \leq e_{s,n,\gamma}^{\text{sh}}(\mathbf{z}) \|f\|_{\gamma}.$$

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

$$\|f\|_{\gamma}^2 = \sum_{u \subseteq \{1, \dots, s\}} \frac{1}{\gamma_u} \int_{[0,1]^{|u|}} \left(\int_{[0,1]^{s-|u|}} \frac{\partial^{|u|} f}{\partial \mathbf{y}_u}(\mathbf{y}) d\mathbf{y}_{-u} \right)^2 d\mathbf{y}_u$$

and (squared) worst case error

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

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 $\mathbf{z} = (z_1, \dots, z_s)$ 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 the weights have certain structure – typically, **product-and-order dependent (POD) weights** are used in practice.

Theorem (CBC error bound)

Let $n = 2^k$ be the number of lattice points. The generating vector $\mathbf{z} \in \mathbb{N}^s$ constructed by the CBC algorithm, minimizing the squared shift-averaged worst-case error $[e_{s,n,\gamma}^{\text{sh}}(\mathbf{z})]^2$ for the weighted unanchored Sobolev space in each step, satisfies

$$[e_{s,n,\gamma}^{\text{sh}}(\mathbf{z})]^2 \leq \left(\frac{2}{n} \sum_{\emptyset \neq \mathbf{u} \subseteq \{1, \dots, s\}} \gamma_{\mathbf{u}}^{\lambda} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}} \right)^{|\mathbf{u}|} \right)^{1/\lambda} \quad \text{for all } \lambda \in (1/2, 1],$$

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

Remarks:

- Optimal rate of convergence $\mathcal{O}(n^{-1+\varepsilon})$ 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 $f \in H_{s,\gamma}$ for all $\gamma = (\gamma_u)_{u \subseteq \{1,\dots,s\}}$. Then for any given sequence of weights γ , we can use the CBC algorithm to obtain a generating vector satisfying the error bound

$$\sqrt{\mathbb{E}_{\Delta} |I_s f - Q_{s,n}^{\Delta} f|^2} \leq \left(\frac{2}{n} \sum_{\emptyset \neq u \subseteq \{1,\dots,s\}} \gamma_u^{\lambda} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}} \right)^{|u|} \right)^{1/(2\lambda)} \|f\|_{s,\gamma} \quad (1)$$

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

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

Part II: Parameterization of input uncertainty

Consider the elliptic PDE problem:

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

In practice, one or several of the material/system parameters may be uncertain or incompletely known and modeled as random fields:

- PDE coefficient a may be uncertain;
- Source term f may be uncertain;
- Boundary conditions may be uncertain;
- The domain D itself may be uncertain.

In **forward uncertainty quantification**, one is interested in assessing how uncertainties in the inputs of a mathematical model affect the output.

In **inverse uncertainty quantification**, one is typically interested in computing the statistics for the posterior distribution of unknown model parameters conditioned on measurements of the system response.

Background

A popular model in the literature: the uniform and affine model.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. For $\mathbf{x} \in D$ and $\omega \in \Omega$,

$$a(\mathbf{x}, \omega) = \bar{a}(\mathbf{x}) + \sum_{j=1}^s Y_j(\omega) \psi_j(\mathbf{x}), \quad Y_j \text{ i.i.d. uniform on } [-\frac{1}{2}, \frac{1}{2}].$$

Computing $\mathbb{E}[u(\mathbf{x}, \cdot)]$ (or some quantity of interest $\mathbb{E}[G(u)]$) using

- Rank-1 lattice cubature rules with random shifts
 \Rightarrow **dimension-independent** cubature error $\mathcal{O}(n^{-1+\varepsilon})$ at best.
(Kuo, Schwab, Sloan 2012)
- Interlaced polynomial lattice rules
 \Rightarrow higher order **dimension-independent** convergence $\mathcal{O}(n^{-1/p})$ for $0 < p < 1$ (p is a summability exponent s.t. $(\|\psi_j\|_{L^\infty})_{j \geq 1} \in \ell^p$).
(Dick, Kuo, Le Gia, Nuyens, Schwab 2014)

Gevrey regular random inputs

Chernov and Lê (2024a and 2024b) observed that the input random field can be much more general while retaining dimension-independent QMC convergence rates for the PDE response. It is enough that the input random field satisfies a certain parametric regularity bound.

Identifying $y_j \equiv Y_j(\omega)$ as parameters, the parametric coefficient $a: D \times [-\frac{1}{2}, \frac{1}{2}]^s \rightarrow \mathbb{R}$ is called **Gevrey regular** if it satisfies

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} a(\cdot, \mathbf{y})\|_{L^\infty(D)} \leq C(|\boldsymbol{\nu}|!)^\beta \mathbf{b}^{\boldsymbol{\nu}} \quad \text{for all } \boldsymbol{\nu} \in \mathbb{N}_0^s, \mathbf{y} \in [-\frac{1}{2}, \frac{1}{2}]^s$$

for some constant $C > 0$ and Gevrey parameter $\beta \geq 1$.

Here, we use multi-index notation

$$\partial_{\mathbf{y}}^{\boldsymbol{\nu}} = \prod_{j=1}^s \frac{\partial^{\nu_j}}{\partial y_j^{\nu_j}}, \quad \mathbf{b}^{\boldsymbol{\nu}} = \prod_{j=1}^s b_j^{\nu_j}, \quad |\boldsymbol{\nu}| = \sum_{j=1}^s \nu_j,$$
$$\boldsymbol{\nu}! = \prod_{j \geq 1} \nu_j, \quad \binom{\boldsymbol{\nu}}{\mathbf{m}} = \prod_{j \geq 1} \binom{\nu_j}{m_j}.$$

Part III: Application to Bayesian inverse problems

We consider a mathematical measurement model

$$\delta = G(\mathbf{y}) + \boldsymbol{\eta},$$

where $\delta \in \mathbb{R}^k$ are the measurements, $\mathbf{y} \in U \subset \mathbb{R}^s$ the unknown, $\boldsymbol{\eta} \sim \mathcal{N}(0, \Gamma)$ is k -dimensional additive Gaussian noise with covariance matrix $\Gamma \in \mathbb{R}^{k \times k}$, $\boldsymbol{\eta}$ is assumed to be independent of \mathbf{y} , and $G: U \rightarrow \mathbb{R}^k$ is the **parameter-to-observation map**.

The likelihood of δ , given \mathbf{y} , is

$$\pi(\delta|\mathbf{y}) \propto \exp\left(-\frac{1}{2}(\delta - G(\mathbf{y}))^T \Gamma^{-1}(\delta - G(\mathbf{y}))\right) = \exp\left(-\frac{1}{2}\|\delta - G(\mathbf{y})\|_{\Gamma^{-1}}^2\right).$$

If \mathbf{y} has the **prior density** $\pi(\mathbf{y})$, then Bayes' formula can be used to obtain the **posterior density** of \mathbf{y} , given δ , as

$$\pi(\mathbf{y}|\delta) = \frac{\pi(\delta|\mathbf{y})\pi(\mathbf{y})}{Z(\delta)}, \quad Z(\delta) = \int_U \pi(\delta|\mathbf{y})\pi(\mathbf{y}) \, d\mathbf{y}.$$

We can obtain information about the unknown parameter \mathbf{y} given some indirect observations δ by computing the statistics of the posterior distribution.

Assumptions about the forward model:

A1 The forward mapping satisfies

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} G(\mathbf{y})\| \leq C_0(|\boldsymbol{\nu}|!)^\beta \mathbf{b}^{\boldsymbol{\nu}},$$

where $\mathbf{b} := (b_j)_{j \geq 1} \in \ell^p$ are nonnegative real numbers for some $p \in (0, 1)$ and $C_0, \beta \geq 1$.

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

A3 There exists $0 < \mu_{\min} \leq 1$ such that $\lambda_{\min}(\Gamma) \geq \mu_{\min}$.

Example

Model problem: Let $D \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, be a nonempty and bounded Lipschitz domain and $z \in H^{-1}(D)$. For each $\mathbf{y} \in U$, there exists a weak solution $u(\cdot, \mathbf{y}) \in H_0^1(D)$ to

$$\begin{cases} -\nabla \cdot (a(\mathbf{x}, \mathbf{y}) \nabla u(\mathbf{x}, \mathbf{y})) = z(\mathbf{x}), & \mathbf{x} \in D, \mathbf{y} \in U, \\ u(\mathbf{x}, \mathbf{y}) = 0, & \mathbf{x} \in \partial D, \mathbf{y} \in U, \end{cases}$$

where we assume that $\mathbf{y} = (y_j)_{j=1}^s$ i.i.d. uniformly distributed in $[-1/2, 1/2]$,

$$a(\mathbf{x}, \mathbf{y}) = a_0(\mathbf{x}) + \sum_{j=1}^s y_j \psi_j(\mathbf{x}), \quad \mathbf{x} \in D, \mathbf{y} \in [-1/2, 1/2]^s,$$

with $a_0 \in L^\infty(D)$ and $\psi_j \in L^\infty(D)$, $j \geq 1$, such that

$0 < a_{\min} \leq a(\mathbf{x}, \mathbf{y}) \leq a_{\max} < \infty$ for all $\mathbf{x} \in D$, $\mathbf{y} \in [-1/2, 1/2]^s$.

Then $G(\mathbf{y}) := \mathcal{O}(u(\cdot, \mathbf{y}))$, where $\mathcal{O}: H_0^1(D) \rightarrow \mathbb{R}^k$ is bounded and linear, satisfies A1 with $b_j := \|\psi_j\|_{L^\infty(D)} / a_{\min}$.

We are interested in conditional mean estimators of the form

$$\hat{\mathbf{F}} = \mathbb{E}_{\pi(\cdot|\delta)}[\mathbf{F}(\mathbf{y})] = \int_U \mathbf{F}(\mathbf{y})\pi(\mathbf{y}|\delta) \mathrm{d}\mathbf{y} = \frac{\int_U \mathbf{F}(\mathbf{y})\pi(\delta|\mathbf{y})\pi(\mathbf{y}) \mathrm{d}\mathbf{y}}{\int_U \pi(\delta|\mathbf{y})\pi(\mathbf{y}) \mathrm{d}\mathbf{y}} =: \frac{Z'}{Z},$$

where $\mathbf{F}: U \rightarrow \mathbb{R}^L$ is some quantity of interest.

For example, if $\mathbf{F}(\mathbf{y}) = \mathbf{y}$, then $\hat{\mathbf{F}}$ is the posterior mean.

- We want to use QMC to approximate the numerator and denominator of this ratio estimator.
- For the design of efficient QMC cubature rules, we need to characterize the parametric regularity of integrands corresponding to both the numerator and denominator, provided that \mathbf{F} is sufficiently smooth (e.g., belonging to a Gevrey class).

Theorem

There holds for all $\mathbf{y} \in U$ and all multi-indices $\boldsymbol{\nu} \neq \mathbf{0}$ that

$$|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} e^{-\frac{1}{2}\|\boldsymbol{\delta} - G(\mathbf{y})\|_{\Gamma}^2}| \leq C_1 C_2^{|\boldsymbol{\nu}|} (|\boldsymbol{\nu}|!)^{\beta} \mathbf{b}^{\boldsymbol{\nu}},$$

where

$$C_1 = \frac{1}{2^{\beta}} \cdot 3.47^k, \quad C_2 = 2^{\beta} \frac{C_0}{\sqrt{\lambda_{\min}}}.$$

Proof (idea): The integrand can be expressed as a composition of functions

$$(f \circ h)(\mathbf{y}),$$

where $f(\mathbf{x}) = e^{-\mathbf{x}^T \mathbf{x} / 2}$ and $h(\mathbf{y}) = \Gamma^{-1/2}(\boldsymbol{\delta} - G(\mathbf{y}))$.

Cramér's inequality: $|\partial_{\mathbf{x}}^{\boldsymbol{\nu}} e^{-\mathbf{x}^T \mathbf{x} / 2}| \leq \sqrt{\boldsymbol{\nu}!}$ for all $\mathbf{x} \in \mathbb{R}^k$ and $\boldsymbol{\nu} \in \mathbb{N}_0^k$.

We already know from the regularity bound of the **forward problem** that

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} h(\mathbf{y})\| \leq \frac{C_0}{\sqrt{\mu_{\min}}} (|\boldsymbol{\nu}|!)^{\beta} \mathbf{b}^{\boldsymbol{\nu}} \text{ for } \boldsymbol{\nu} \neq \mathbf{0}.$$

We use the multivariate chain rule (**Faà di Bruno's formula**) to put everything together.

Faà di Bruno's formula (recursive version; Savits 2006):

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

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

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

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

$$\kappa_{\boldsymbol{\nu} + \mathbf{e}_j, \boldsymbol{\lambda}}(\mathbf{y}) = \sum_{\ell \in \text{supp}(\boldsymbol{\lambda})} \sum_{\mathbf{0} \leq \mathbf{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\mathbf{m}} \partial_{\mathbf{y}}^{\mathbf{m} + \mathbf{e}_j} [h(\mathbf{y})]_{\ell} \kappa_{\boldsymbol{\nu} - \mathbf{m}, \boldsymbol{\lambda} - \mathbf{e}_\ell}(\mathbf{y}) \quad \text{otherwise.}$$

Since $\|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} h(\mathbf{y})\| \leq \frac{C_0}{\sqrt{\mu_{\min}}} (|\boldsymbol{\nu}|!)^{\beta} \mathbf{b}^{\boldsymbol{\nu}}$, we obtain the following *uniform* bound.

Lemma

For all $1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|$, there holds

$$|\kappa_{\boldsymbol{\nu}, \boldsymbol{\lambda}}(\mathbf{y})| \leq \left(\frac{C_0}{\sqrt{\mu_{\min}}} \right)^{|\boldsymbol{\nu}|} \left(\frac{|\boldsymbol{\nu}|! (|\boldsymbol{\nu}| - 1)!}{\boldsymbol{\lambda}! (|\boldsymbol{\nu}| - |\boldsymbol{\lambda}|)! (|\boldsymbol{\lambda}| - 1)!} \right)^{\beta} \mathbf{b}^{\boldsymbol{\nu}}.$$

Proof.

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

□

For the present problem, we obtain

$$\begin{aligned}
& |\partial_{\mathbf{y}}^{\boldsymbol{\nu}} e^{-\frac{1}{2}\|\mathbf{y}-G(\mathbf{y})\|_{\Gamma}^2}| \\
& \leq \left(\frac{C_0}{\sqrt{\mu_{\min}}} \right)^{|\boldsymbol{\nu}|} \\
& \quad \times \sum_{\substack{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}| \\ \boldsymbol{\lambda} \in \mathbb{N}_0^k}} |\partial_{\mathbf{x}}^{\boldsymbol{\lambda}} e^{-\mathbf{x}^T \mathbf{x}/2}| \Big|_{\mathbf{x}=\Gamma^{-1/2}(\mathbf{y}-G(\mathbf{y}))} \left(\frac{|\boldsymbol{\nu}|!(|\boldsymbol{\nu}|-1)!}{\boldsymbol{\lambda}!(|\boldsymbol{\nu}|-|\boldsymbol{\lambda}|)!(|\boldsymbol{\lambda}|-1)!} \right)^{\beta} \mathbf{b}^{\boldsymbol{\nu}} \\
& \leq \left(\frac{C_0}{\sqrt{\mu_{\min}}} \right)^{|\boldsymbol{\nu}|} \mathbf{b}^{\boldsymbol{\nu}} (|\boldsymbol{\nu}|!)^{\beta} ((|\boldsymbol{\nu}|-1)!)^{\beta} \sum_{\substack{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}| \\ \boldsymbol{\lambda} \in \mathbb{N}_0^k}} \left(\frac{\sqrt{\boldsymbol{\lambda}!}}{\boldsymbol{\lambda}!(|\boldsymbol{\nu}|-|\boldsymbol{\lambda}|)!(|\boldsymbol{\lambda}|-1)!} \right)^{\beta}.
\end{aligned}$$

It remains to estimate the multi-index sum, but fortunately this is not too difficult:

$$\begin{aligned}
& \sum_{\substack{1 \leq |\lambda| \leq |\nu| \\ \lambda \in \mathbb{N}_0^k}} \left(\frac{\sqrt{\lambda}!}{\lambda!(|\nu| - |\lambda|)!(|\lambda| - 1)!} \right)^\beta \\
&= \sum_{\ell=1}^{|\nu|} \left(\frac{1}{(|\nu| - \ell)!(\ell - 1)!} \right)^\beta \sum_{\substack{\lambda \in \mathbb{N}_0^k \\ |\lambda| = \ell}} \left(\frac{1}{\sqrt{\lambda}!} \right)^\beta \\
&\leq \left(\sum_{\ell=1}^{|\nu|} \left(\frac{1}{(|\nu| - \ell)!(\ell - 1)!} \right)^\beta \right) \underbrace{\left(\sum_{\lambda=0}^{\infty} \frac{1}{\sqrt{\lambda}!} \right)^k}_{=3.469506} \\
&\leq 3.47^k \left(\sum_{\ell=1}^{|\nu|} \frac{1}{(|\nu| - \ell)!(\ell - 1)!} \right)^\beta = 3.47^k \cdot \frac{2^{\beta|\nu| - \beta}}{((|\nu| - 1)!)^\beta},
\end{aligned}$$

where we made use of $\sum_k a_k \leq (\sum_k a_k^{1/\beta})^\beta$ for $a_k \geq 0$ and $\beta \geq 1$ as well as the summation identity $\sum_{\ell=1}^{\nu} \frac{1}{(\nu - \ell)!(\ell - 1)!} = \frac{2^{\nu-1}}{(\nu-1)!}$.

This yields the desired result. □

Theorem

Suppose that the quantity of interest $\mathbf{F}: U \rightarrow \mathbb{R}^L$ satisfies

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} \mathbf{F}(\mathbf{y})\| \leq C_0(|\boldsymbol{\nu}|!)^{\beta} \mathbf{b}^{\boldsymbol{\nu}}.$$

There holds for all $\mathbf{y} \in U$ and all multi-indices $\boldsymbol{\nu} \neq \mathbf{0}$ that

$$|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} (\mathbf{F}(\mathbf{y}) e^{-\frac{1}{2}\|\boldsymbol{\delta}-G(\mathbf{y})\|_{\Gamma-1}^2})| \leq C_0 C_1 C_2^{|\boldsymbol{\nu}|} ((|\boldsymbol{\nu}| + 1)!)^{\beta} \mathbf{b}^{\boldsymbol{\nu}}.$$

Proof. By the Leibniz product rule:

$$\begin{aligned} |\partial_{\mathbf{y}} (\mathbf{F}(\mathbf{y}) e^{-\frac{1}{2}\|\boldsymbol{\delta}-G(\mathbf{y})\|_{\Gamma-1}^2})| &= \left| \sum_{\mathbf{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\mathbf{m}} \partial_{\mathbf{y}}^{\boldsymbol{\nu}-\mathbf{m}} \mathbf{F}(\mathbf{x}, \mathbf{y}) \partial^{\mathbf{m}} e^{-\frac{1}{2}\|\boldsymbol{\delta}-G(\mathbf{y})\|_{\Gamma-1}^2} \right| \\ &\leq \sum_{\mathbf{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\mathbf{m}} C_0((|\boldsymbol{\nu}| - |\mathbf{m}|)!)^{\beta} \mathbf{b}^{\boldsymbol{\nu}-\mathbf{m}} C_1 C_2^{|\mathbf{m}|} (|\mathbf{m}|!)^{\beta} \mathbf{b}^{\mathbf{m}} \\ &\leq C_0 C_1 C_2^{|\boldsymbol{\nu}|} \mathbf{b}^{\boldsymbol{\nu}} \sum_{\mathbf{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\mathbf{m}} ((|\boldsymbol{\nu}| - |\mathbf{m}|)!)^{\beta} (|\mathbf{m}|!)^{\beta} \\ &\leq \cdots \leq C_0 C_1 C_2^{|\boldsymbol{\nu}|} \mathbf{b}^{\boldsymbol{\nu}} ((|\boldsymbol{\nu}| + 1)!)^{\beta}. \quad \square \end{aligned}$$

We want to apply QMC to approximate both the numerator and denominator of the ratio estimator $\frac{Z'}{Z}$.

The ratio estimator satisfies the following bound:

$$\begin{aligned} \left| \frac{Z'}{Z} - \frac{Z'_n}{Z_n} \right| &= \left| \frac{Z'Z_n - Z'_nZ}{ZZ_n} \right| = \left| \frac{Z'Z_n - Z'Z + Z'Z - Z'_nZ}{ZZ_n} \right| \\ &\leq \frac{|Z'| |Z - Z_n|}{|ZZ_n|} + \frac{|Z' - Z'_n|}{|Z_n|} \\ &\lesssim |Z - Z_n| + |Z' - Z'_n|, \end{aligned}$$

meaning that we can simply use the larger derivative bound to inform our choice of weights:

$$\gamma_u := \left(((|u| + 1)!)^\beta \prod_{j \in u} \frac{C_2 b_j}{\sqrt{2\zeta(2\lambda)/(2\pi^2)^\lambda}} \right)^{\frac{2}{1+\lambda}} \quad \text{for } u \subset \{1, \dots, s\},$$

where we set $\lambda := \frac{p}{2-p}$ if $p \in (\frac{2}{3}, \frac{1}{\beta})$ and $\lambda = \frac{1}{2-2\varepsilon}$ for arbitrary $\varepsilon \in (0, \frac{1}{2})$ if $p \in (0, \min\{\frac{2}{3}, \frac{1}{\beta}\}]$, $p \neq \frac{1}{\beta}$. We obtain QMC convergence rate

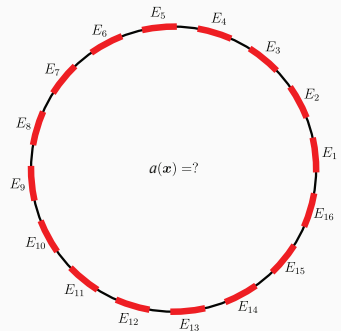
$\sqrt{\mathbb{E}_\Delta \left| \frac{Z'}{Z} - \frac{Z'_{n,\Delta}}{Z_{n,\Delta}} \right|^2} = \mathcal{O}(n^{\max\{-\frac{1}{p} + \frac{1}{2}, -1+\varepsilon\}})$ where the implied coefficient can be shown to be independent of the dimension s .

Application 1: Electrical impedance tomography

Electrical impedance tomography

Use measurements of current and voltage collected at electrodes covering part of the boundary to infer the interior conductivity of an object/body.

$$\begin{cases} \nabla \cdot (a \nabla u) = 0 & \text{in } D, \\ a \frac{\partial u}{\partial \mathbf{n}} = 0 & \text{on } \partial D \setminus \bigcup_{k=1}^L \overline{E_k}, \\ u + z_k a \frac{\partial u}{\partial \mathbf{n}} = \mathcal{U}_k & \text{on } E_k, \quad k \in \{1, \dots, L\}, \\ \int_{E_k} a \frac{\partial u}{\partial \mathbf{n}} \, dS = I_k, & k \in \{1, \dots, L\}, \end{cases}$$



Model specifics

- $D \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$ nonempty physical domain with Lipschitz boundary, $y \in U := [-1/2, 1/2]^s$.
- Assumptions on a :
 1. $a(\cdot, \mathbf{y}) \in L^\infty(D)$ for all $\mathbf{y} \in U$.
 2. There exist constants $C_a, \sigma \geq 1$ and a sequence $\boldsymbol{\rho} = (\rho_j)_{j \geq 1} \in \ell^p(\mathbb{N})$, $\rho \in (0, 1)$, of non-negative real numbers such that

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} a(\cdot, \mathbf{y})\|_{L^\infty(D)} \leq C_a(|\boldsymbol{\nu}|!)^\beta \boldsymbol{\rho}^{\boldsymbol{\nu}} \quad \text{for all } \boldsymbol{\nu} \in \mathbb{N}_0^s \text{ and } \mathbf{y} \in U.$$

3. There exist positive constants a_{\min} and a_{\max} such that

$$0 < a_{\min} \leq a(\mathbf{x}, \mathbf{y}) \leq a_{\max} < \infty \quad \text{for all } \mathbf{x} \in D \text{ and } \mathbf{y} \in U.$$

- $\{E_k\}_{k=1}^L$, $L \geq 2$, open, nonempty, connected subsets of D s.t.
 $\overline{E}_i \cap \overline{E}_j = \emptyset$ for $i \neq j$
- Quotient Hilbert space $\mathcal{H} := (H^1(D) \oplus \mathbb{R}^L)/\mathbb{R}$ with the norm

$$\|(v, V)\|_{\mathcal{H}}^2 := \int_D |\nabla v|^2 d\mathbf{x} + \sum_{k=1}^L \int_{E_k} (v - V_m)^2 dS, \quad (v, V) \in \mathcal{H}.$$

- Contact impedances $\{z_m\}_{m=1}^M$ satisfy $0 < \varsigma_- \leq z_m \leq \varsigma_+ < \infty \forall m$.

Forward problem

For $\mathbf{y} \in U$ find the electromagnetic potential $u(\cdot, \mathbf{y})$ and the potentials on the electrodes $\mathcal{U}(\mathbf{y})$ such that the model equations hold.

Variational formulation: Find $(u(\cdot, \mathbf{y}), \mathcal{U}(\mathbf{y})) \in \mathcal{H}$ such that

$$\int_D a(\cdot, \mathbf{y}) \nabla u(\cdot, \mathbf{y}) \cdot \nabla v \, dx + \sum_{m=1}^L \frac{1}{z_m} \int_{E_m} (u(\cdot, \mathbf{y}) - \mathcal{U}_m(\mathbf{y}))(v - V_m) \, dS = \sum_{m=1}^L I_m V_m$$

for all $(v, V) \in \mathcal{H}$.

Lemma

Let $\nu \in \mathcal{F}$ and $\mathbf{y} \in U$. Then there exists a constant $C = C_{D,E} \geq 1$ s.t.

$$\|(u(\cdot, \mathbf{y}), \mathcal{U}(\mathbf{y}))\|_{\mathcal{H}} \leq \frac{C_{D,E}|I|}{\min\{a_{\min}, \varsigma_+^{-1}\}} \quad \text{for all } \mathbf{y} \in U,$$

and, moreover,

$$\|\partial^\nu(u(\cdot, \mathbf{y}), \mathcal{U}(\mathbf{y}))\|_{\mathcal{H}} \leq C_u(|\nu|!)^\beta \mathbf{b}^\nu,$$

where $C_u = \frac{C_{D,E}|I|}{\min\{a_{\min}, \varsigma_+^{-1}\}}$ and $b_j := \left(1 + \frac{C_a}{\min\{a_{\min}, \varsigma_+^{-1}\}}\right) \rho_j$.

This fits into our theoretical framework!

Inverse problem

Assume there are some voltage measurements taken at the L electrodes placed on the boundary of the computational domain D corresponding to the observation operator $G: U \rightarrow \mathbb{R}^L$

$$G(\mathbf{y}) := [\mathcal{U}_1(\mathbf{y}), \dots, \mathcal{U}_L(\mathbf{y})]^T.$$

Consider the problem of finding an unknown $\mathbf{y} \in U$ using the data $\boldsymbol{\delta} \in \mathbb{R}^L$, where \mathbf{y} and $\boldsymbol{\delta}$ are related by the equation

$$\boldsymbol{\delta} = \mathcal{O}(\mathbf{y}) + \boldsymbol{\eta}.$$

- $\boldsymbol{\eta} \in \mathbb{R}^L$ Gaussian observational noise in $\boldsymbol{\delta}$, let $\mathbf{y}, \boldsymbol{\delta}, \boldsymbol{\eta}$ to be random variables independent of each other

To reconstruct $a(\cdot, \mathbf{y})$ with the unknown parameter $\mathbf{y} \in [-1/2, 1/2]^s$, we consider the estimator

$$\hat{a}(\mathbf{x}) = \frac{\int_{[-1/2, 1/2]^s} a(\mathbf{x}, \mathbf{y}) \pi(\boldsymbol{\delta}|\mathbf{y}) \pi(\mathbf{y}) d\mathbf{y}}{\int_{[-1/2, 1/2]^s} \pi(\boldsymbol{\delta}|\mathbf{y}) \pi(\mathbf{y}) d\mathbf{y}}.$$

Bayes' formula:

$$\pi(\mathbf{y}|\boldsymbol{\delta}) = \frac{\pi(\boldsymbol{\delta}|\mathbf{y})\pi(\mathbf{y})}{Z(\boldsymbol{\delta})}, \quad \pi(\boldsymbol{\delta}|\mathbf{y}) \propto e^{-\frac{1}{2}\|\boldsymbol{\delta}-G(\mathbf{y})\|_{\Gamma}^2},$$

where $\pi(\mathbf{y})$ denotes the prior density of \mathbf{y} , here $\mathbf{y} \sim \mathcal{U}([-1/2, 1/2]^s)$, and $\pi(\boldsymbol{\delta}|\mathbf{y})$ the likelihood of $\boldsymbol{\delta}$ given \mathbf{y} .

Generate synthetic measurement data δ using the parametrization

$$a(\mathbf{x}, \mathbf{y}) = \exp \left(\sum_{j=1}^{20} y_j \psi_j(\mathbf{x}) \right), \quad \mathbf{x} \in D, \mathbf{y} \in [-1/2, 1/2]^{20},$$

with

$$\psi_j(\mathbf{x}) := \frac{5}{(k_j^2 + \ell_j^2)^2} \sin(\tfrac{1}{14}\pi k_j x_1) \sin(\tfrac{1}{14}\pi \ell_j x_2)$$

where the sequence $(k_j, \ell_j)_{j \geq 1}$ is an ordering of the elements of $\mathbb{N} \times \mathbb{N}$.

The measurements were contaminated by 10% relative Gaussian noise.

Reconstruction

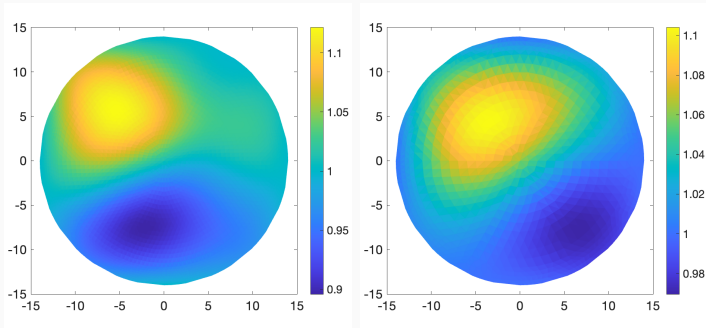
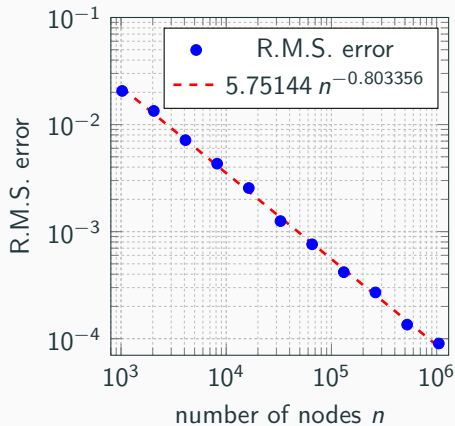


Figure 3: Randomly generated ground truth conductivity field on the left, reconstruction corresponding to noisy voltage measurements on the right. The reconstruction corresponds to $n = 2^{20}$ QMC nodes with a single random shift.

Error plot

R.M.S. errors for the reconstruction corresponding to the reconstruction with $R = 16$ random shifts plotted alongside the corresponding least squares fits.



Application 2: Bayesian shape inversion

Consider the Poisson problem

$$\begin{cases} -\Delta u(\mathbf{x}, \omega) = f(\mathbf{x}) & \text{for } \mathbf{x} \in D(\omega), \\ u(\mathbf{x}, \omega) = 0 & \text{for } \mathbf{x} \in \partial D(\omega), \end{cases}$$

where the bounded domain $D(\omega) \subset \mathbb{R}^d$, $d \in \{2, 3\}$, is assumed to be *uncertain*.

Domain mapping method: Let $D_{\text{ref}} \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a fixed reference domain. Define perturbation field $\mathbf{V}(\omega): \overline{D_{\text{ref}}} \rightarrow \mathbb{R}^d$, which we assume is given explicitly.

Uncertain domains studied by many authors in the literature: Harbrecht, Multerer, Siebenmorgen, Schwab, Gantner, Zech...

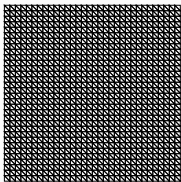


Figure 4: Reference domain

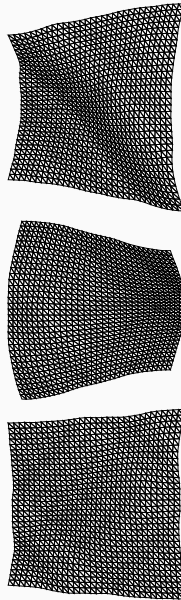


Figure 5: Three realizations of the random domain

Model specifics

- **Fixed reference domain** $D_{\text{ref}} \subset \mathbb{R}^d$, $d \in \{2, 3\}$ is a nonempty physical domain with Lipschitz boundary.
- The family of **admissible domains** $D(\mathbf{y}) := \mathbf{V}(D_{\text{ref}}, \mathbf{y})$ for $\mathbf{y} \in U := [-1/2, 1/2]^s$. **Hold-all domain** $\mathcal{D} = \bigcup_{\mathbf{y} \in U} D(\mathbf{y})$.
- Assumptions on \mathbf{V} :
 1. $\mathbf{V}(\cdot, \mathbf{y}): \overline{D_{\text{ref}}} \rightarrow \overline{D(\mathbf{y})}$ is a \mathcal{C}^2 -diffeomorphism for all $\mathbf{y} \in U$.
 2. There exist constants $C_{\mathbf{V}}, \beta \geq 1$ and a sequence $\mathbf{b} = (b_j)_{j \geq 1} \in \ell^p(\mathbb{N})$, $p \in (0, 1)$, of non-negative real numbers such that

$$\|\partial_{\mathbf{y}}^{\nu} \mathbf{V}(\cdot, \mathbf{y})\|_{W^{1,\infty}(D_{\text{ref}})} \leq C_{\mathbf{V}}(|\nu|!)^{\beta} \mathbf{b}^{\nu} \quad \text{for all } \nu \in \mathbb{N}_0^s \text{ and } \mathbf{y} \in U.$$

3. There exist positive constants $0 < \sigma_{\min} \leq 1 \leq \sigma_{\max}$ such that

$$0 < \sigma_{\min} \leq \min \sigma(D\mathbf{V}(\mathbf{x}, \mathbf{y})) \leq \max \sigma(D\mathbf{V}(\mathbf{x}, \mathbf{y})) \leq \sigma_{\max} < \infty$$

for all $\mathbf{x} \in D$ and $\mathbf{y} \in U$, with $\sigma(M)$ denoting the set of singular values of matrix M .

The variational formulation of the model problem can be stated as follows: for $\mathbf{y} \in U$, find $u(\cdot, \mathbf{y}) \in H_0^1(D(\mathbf{y}))$ such that

$$\int_{D(\mathbf{y})} \nabla u(\mathbf{x}, \mathbf{y}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x} = \int_{D(\mathbf{y})} f(\mathbf{x}) v(\mathbf{x}) \, d\mathbf{x} \quad \forall v \in H_0^1(D(\mathbf{y})), \quad (2)$$

where $f: \mathcal{D} \rightarrow \mathbb{R}$ is assumed to satisfy $\|\partial^\nu f\|_{L^\infty(\mathcal{D})} \leq C_V(|\nu|!)^\beta \rho^\nu$.

We can transport the variational formulation (2) to the reference domain by a change of variables. Let us define

$$\begin{aligned} A(\mathbf{x}, \mathbf{y}) &:= (J(\mathbf{x}, \mathbf{y})^T J(\mathbf{x}, \mathbf{y}))^{-1} \det J(\mathbf{x}, \mathbf{y}) \\ f_{\text{ref}}(\mathbf{x}, \mathbf{y}) &:= f(\mathbf{V}(\mathbf{x}, \mathbf{y})) \det J(\mathbf{x}, \mathbf{y}), \end{aligned}$$

for $\mathbf{x} \in D_{\text{ref}}$, $\mathbf{y} \in U$. Then we can recast the problem (2) on the reference domain as follows: for $\mathbf{y} \in U$, find $\hat{u}(\cdot, \mathbf{y}) \in H_0^1(D_{\text{ref}})$ such that

$$\int_{D_{\text{ref}}} (A(\mathbf{x}, \mathbf{y}) \nabla \hat{u}(\mathbf{x}, \mathbf{y})) \cdot \nabla \hat{v}(\mathbf{x}) \, d\mathbf{x} = \int_{D_{\text{ref}}} f_{\text{ref}}(\mathbf{x}, \mathbf{y}) \hat{v}(\mathbf{x}) \, d\mathbf{x} \quad \forall \hat{v} \in H_0^1(D_{\text{ref}}). \quad (3)$$

The solutions to problems (2) and (3) are connected to one another via

$$u(\cdot, \mathbf{y}) = \hat{u}(\mathbf{V}^{-1}(\cdot, \mathbf{y}), \mathbf{y}) \quad \Leftrightarrow \quad \hat{u}(\cdot, \mathbf{y}) = u(\mathbf{V}(\cdot, \mathbf{y}), \mathbf{y}), \quad \mathbf{y} \in U.$$

Theorem

There holds for all $\mathbf{y} \in U$ and all multi-indices $\boldsymbol{\nu} \neq \mathbf{0}$ that

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} \hat{u}(\cdot, \mathbf{y})\|_{H_0^1(D_{\text{ref}})} \leq C_1 C_2^{|\boldsymbol{\nu}|} (|\boldsymbol{\nu}|!)^{\beta} \mathbf{b}^{\boldsymbol{\nu}},$$

where

$$C_1 = 1 + \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^d \frac{\sigma_{\max}^2 C_{D_{\text{ref}}} |D_{\text{ref}}|^{1/2} C_{\mathbf{V}}}{2^{\beta}}$$

$$C_2 = ((d^2)!)^{\beta} 2^{\beta(d^2+1)+1} \\ \times \max \left\{ \left(\frac{\sigma_{\max}^d}{\sigma_{\min}^{d+2} ((d^2)!)^{\beta}} \right), 2 \left(\frac{2^{\beta} C_{\mathbf{V}}}{\sigma_{\min}} \right)^3, \frac{C_1 - 1}{\sigma_{\max}^2 ((d^2)!)^{\beta}}, \frac{(2^{\beta} C_{\mathbf{V}})^2}{\sigma_{\min}} \max\{1, \|\boldsymbol{\rho}\|_{\ell^{1/\beta}}\} \right\}$$

with $C_{D_{\text{ref}}}$ denoting the Poincaré constant of D_{ref} and $|D_{\text{ref}}| = \int_{D_{\text{ref}}} d\mathbf{x}$.

This fits into our theoretical framework!

As the reconstruction of the unknown domain, we consider

$$\hat{\mathbf{V}}(\mathbf{x}) = \frac{\int_U \mathbf{V}(\mathbf{x}, \mathbf{y}) e^{-\frac{1}{2} \|\boldsymbol{\delta} - G(\mathbf{y})\|_{\Gamma^{-1}}^2} d\mathbf{y}}{\int_U e^{-\frac{1}{2} \|\boldsymbol{\delta} - G(\mathbf{y})\|_{\Gamma^{-1}}^2} d\mathbf{y}}.$$

Numerical experiments

Let $D_{\text{ref}} = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$ and consider the perturbation field

$$\mathbf{V}(\mathbf{x}, \mathbf{y}) = a(\mathbf{x}, \mathbf{y})\mathbf{x}, \quad \mathbf{x} \in D_{\text{ref}}, \quad \mathbf{y} \in \left[-\frac{1}{2}, \frac{1}{2}\right]^s,$$

where

$$a(\mathbf{x}, \mathbf{y}) := 1 + \sum_{j=1}^s j^{-2.1} \sin(3j \operatorname{atan2}(x_1, x_2) + \pi) e^{-(\frac{1}{2} + y_j)^{-1}}.$$

with $s = 100$. We fix $f(\mathbf{x}) = 10 \sin(x_1 x_2) - 5 \cos^2(x_1 + x_2)$.

Thus the reference domain in this case is the unit disk, and the uncertain boundary is the curve defined by the radial transformation

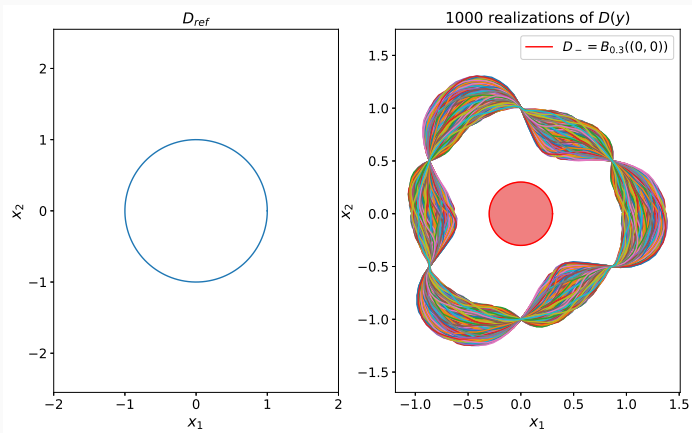
$r = 1 + \sum_{j=1}^s j^{-2.1} \sin(3j \theta + \pi) e^{-(\frac{1}{2} + y_j)^{-1}}$ for each realization of \mathbf{y} .

The reconstruction is obtained by solving the PDE problem over the transported domain using piecewise linear FEM with mesh size $h = 2^{-5}$.

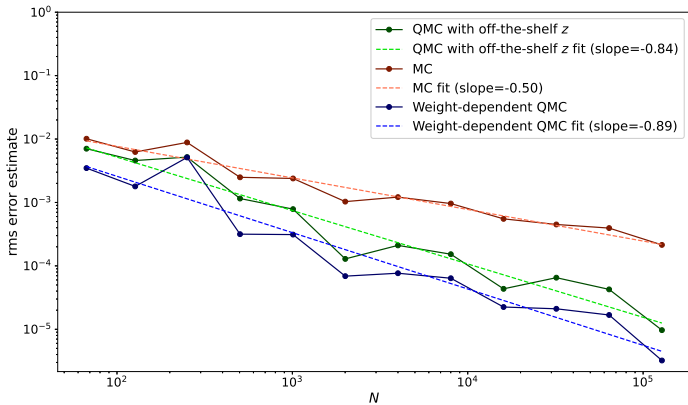
To assess the convergence behavior, we first consider the observation operator

$$G: H_0^1(\mathcal{D}) \rightarrow \mathbb{R}, \quad u \mapsto \int_{D_-} u(\mathbf{x}) \, d\mathbf{x},$$

where $D_- = B_{0.3}((0,0)) \subset \bigcap_{\mathbf{y} \in [-\frac{1}{2}, \frac{1}{2}]^{100}} D(\mathbf{y})$. The observations were generated using a finer FE mesh ($h = 2^{-6}$) and the measurements were contaminated with 10% relative Gaussian noise.



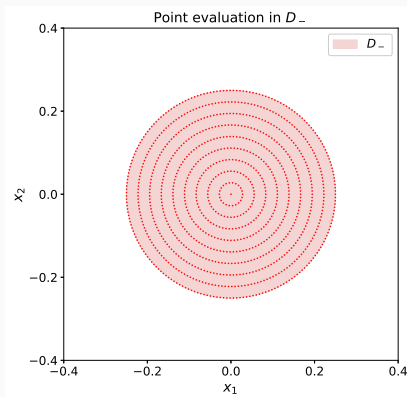
QMC convergence of \hat{V}



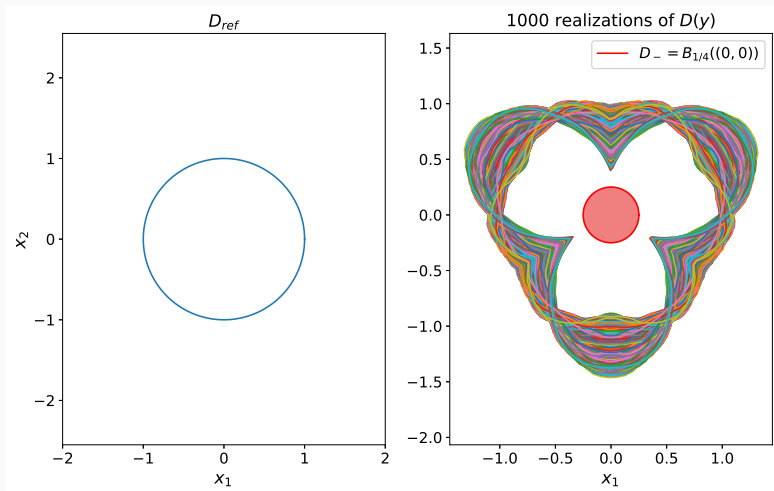
Next, we consider reconstructing the domain shape based on measurements. Here, we consider the observation operator

$$G(u(\cdot, \mathbf{y})) = [u(\mathbf{x}_1, \mathbf{y}), \dots, u(\mathbf{x}_k, \mathbf{y})]^\top,$$

where the PDE solution is sampled over a point set with $k = 996$ belonging to D_- (below). The observations were generated using a finer FE mesh ($h = 2^{-6}$) and contaminated with 5% relative Gaussian noise.

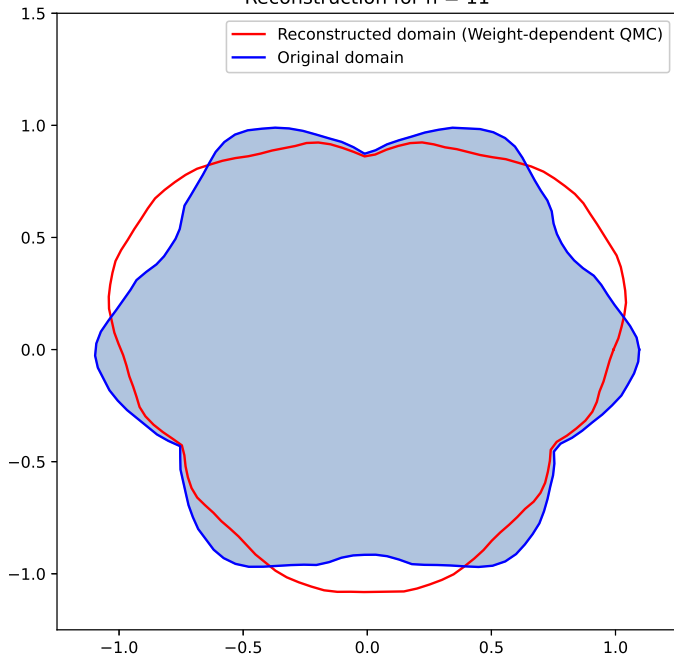


First, we generate the ground truth domain as a realization of the random field with $s = 200$ terms. Then we compute the reconstruction using $s = 20$ terms.

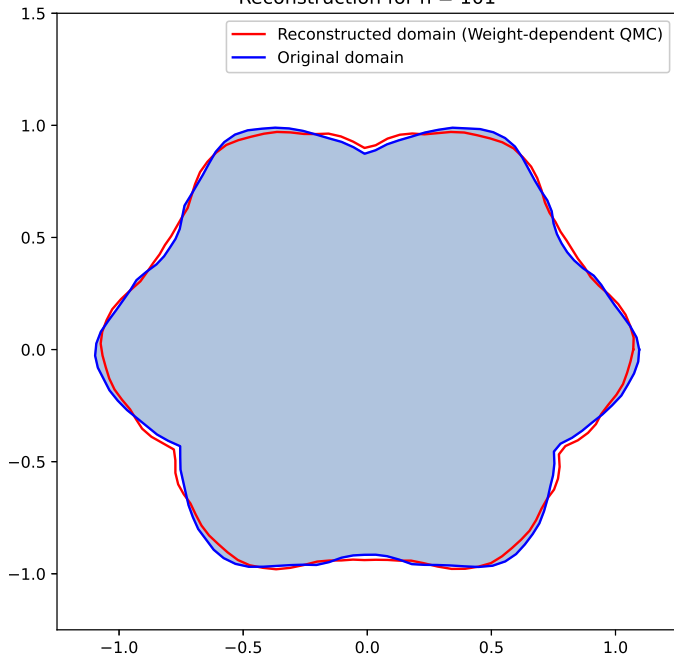


$$a(\mathbf{x}, \mathbf{y}) := 1 + 1.2 \sum_{j=1}^s j^{-2.1} \sin(3j \operatorname{atan2}(x_1, x_2) - \pi/2) e^{-(\frac{1}{2} + y_j)^{-1}}$$

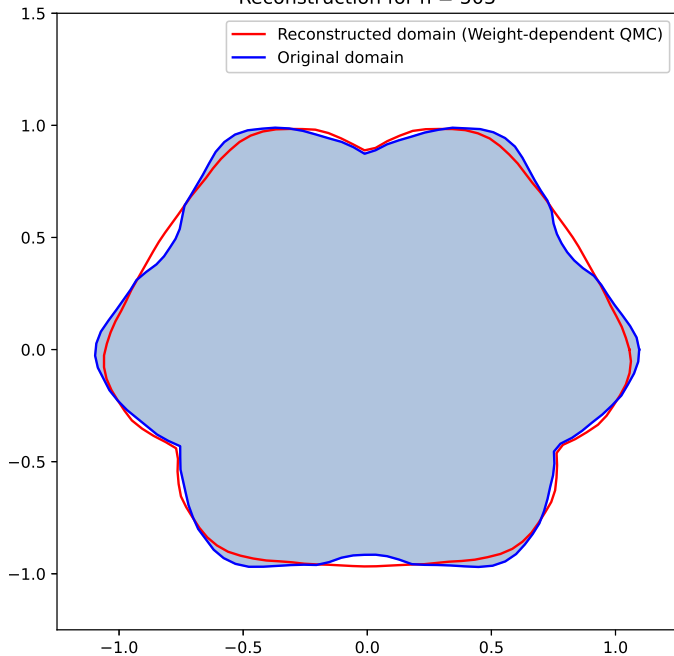
Reconstruction for $n = 11$



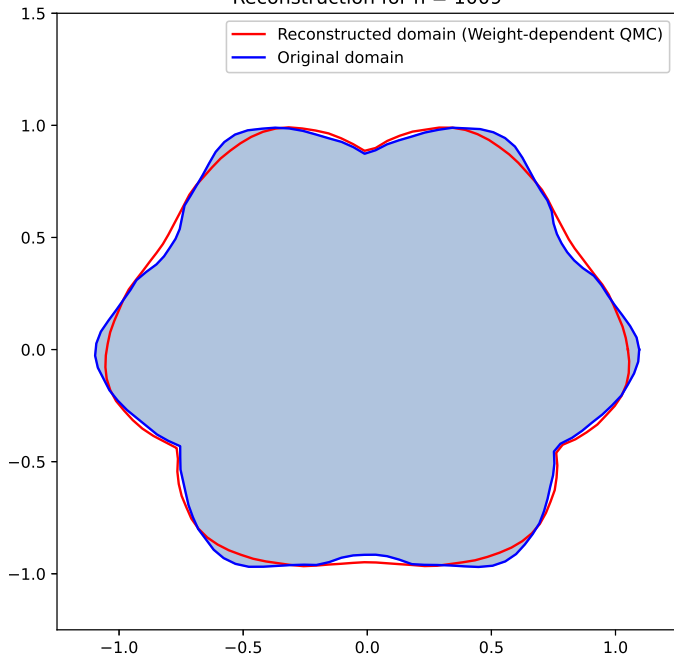
Reconstruction for $n = 101$



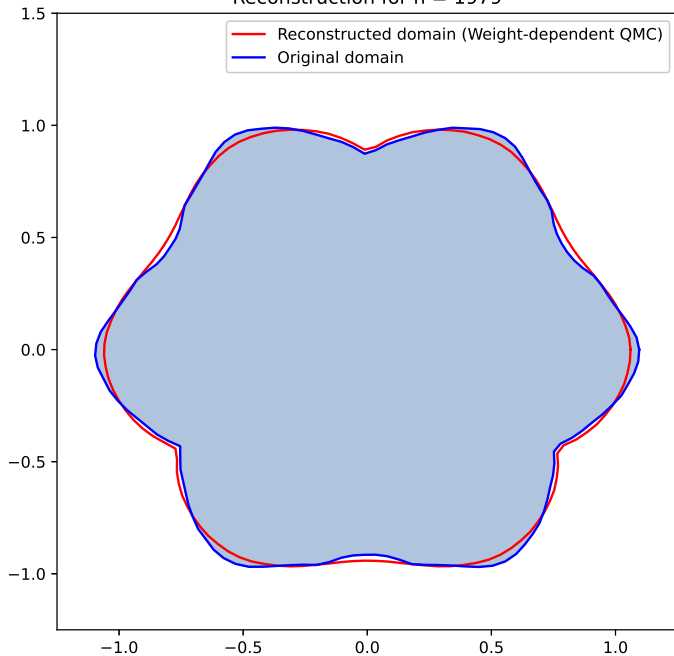
Reconstruction for $n = 503$



Reconstruction for $n = 1009$



Reconstruction for $n = 1979$



Application 3: Bayesian optimal experimental design

Let $G: U \times \Xi \rightarrow \mathbb{R}^k$ be a forward mapping depending on a true parameter $\mathbf{y} \in U$ and a design parameter $\xi \in \Xi$.

Measurement model:

$$\delta = G(\mathbf{y}, \xi) + \eta,$$

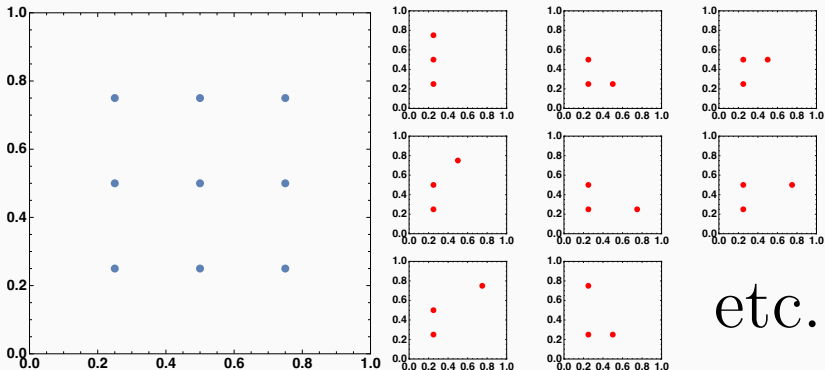
where $\delta \in \mathbb{R}^k$ is the measurement data and $\eta \in \mathbb{R}^k$ is Gaussian noise such that $\eta \sim \mathcal{N}(0, \Gamma)$ with positive definite covariance matrix $\Gamma \in \mathbb{R}^{k \times k}$.

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

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?

→ $\binom{9}{3} = 84$ 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 ξ and data δ is given by the Kullback–Leibler divergence

$$D_{\text{KL}}(\pi(\cdot|\delta, \xi) \parallel \pi(\cdot)) := \int_U \log \left(\frac{\pi(\mathbf{y}|\delta, \xi)}{\pi(\mathbf{y})} \right) \pi(\mathbf{y}|\delta, \xi) d\mathbf{y}. \quad (4)$$

We wish to maximize the expected utility (4) over the design space Ξ with respect to the data δ and model parameters \mathbf{y} :

$$\max_{\xi \in \Xi} \underbrace{\int_Y \int_U \log \left(\frac{\pi(\mathbf{y}|\delta, \xi)}{\pi(\mathbf{y})} \right) \pi(\mathbf{y}|\delta, \xi) \pi(\delta|\xi) d\mathbf{y} d\delta}_{=:\text{EIG}},$$

where $\pi(\mathbf{y}|\delta, \xi)$ corresponds to the posterior distribution of the parameter \mathbf{y} and $\pi(\delta|\xi) = \int_U \pi(\delta|\mathbf{y}, \xi) \pi(\mathbf{y}) d\mathbf{y}$ is the marginal distribution of the data δ .

The posterior is given by Bayes' theorem

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

which means that the expected utility can be written as

$$\begin{aligned}\text{EIG} &= \int_Y \int_U \log \left(\frac{\pi(\mathbf{y}|\boldsymbol{\delta}, \boldsymbol{\xi})}{\pi(\mathbf{y})} \right) \pi(\mathbf{y}|\boldsymbol{\delta}, \boldsymbol{\xi}) d\mathbf{y} \pi(\boldsymbol{\delta}|\boldsymbol{\xi}) d\boldsymbol{\delta} \\ &= \int_U \left[\int_Y \log \left(\frac{\pi(\boldsymbol{\delta}|\mathbf{y}, \boldsymbol{\xi})}{\pi(\boldsymbol{\delta}|\boldsymbol{\xi})} \right) \pi(\boldsymbol{\delta}|\mathbf{y}, \boldsymbol{\xi}) d\boldsymbol{\delta} \right] \pi(\mathbf{y}) d\mathbf{y}.\end{aligned}$$

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)

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

$$\pi(\boldsymbol{\delta}|\mathbf{y}, \boldsymbol{\xi}) = C e^{-\frac{1}{2}\|\boldsymbol{\delta}-G(\mathbf{y}, \boldsymbol{\xi})\|_{\Gamma^{-1}}^2}, \quad C = \frac{1}{(2\pi)^{k/2}\sqrt{\det \Gamma}}.$$

Under these conditions, it is easy to see that

$$\begin{aligned} \text{EIG} &= \int_Y \int_U \log \left(\frac{\pi(\boldsymbol{\delta}|\mathbf{y}, \boldsymbol{\xi})}{\pi(\boldsymbol{\delta}|\boldsymbol{\xi})} \right) \pi(\boldsymbol{\delta}|\mathbf{y}, \boldsymbol{\xi}) \pi(\mathbf{y}) \, d\mathbf{y} \, d\boldsymbol{\delta} \\ &= \log C - 1 - \int_Y \log \left(\int_U C e^{-\frac{1}{2}\|\boldsymbol{\delta}-G(\mathbf{y}, \boldsymbol{\xi})\|_{\Gamma^{-1}}^2} \, d\mathbf{y} \right) \int_U C e^{-\frac{1}{2}\|\boldsymbol{\delta}-G(\mathbf{y}, \boldsymbol{\xi})\|_{\Gamma^{-1}}^2} \, d\mathbf{y} \, d\boldsymbol{\delta}. \end{aligned}$$

Observations:

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

$$\int_Y \log \left(\int_U C e^{-\frac{1}{2} \|\delta - G(y, \xi)\|_{r-1}^2} dy \right) \int_U C e^{-\frac{1}{2} \|\delta - G(y, \xi)\|_{r-1}^2} dy d\delta$$

QMC weights for the **inner integral**:

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

with $\rho_j := \frac{2C_0}{\sqrt{\mu_{\min}}} b_j$, $j \in \{1, \dots, s\}$, 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.

For simplicity, assume that the data is in $Y = [-K, K]$.

$$\int_{[-K, K]^k} \log \left(\int_U C e^{-\frac{1}{2} \|\delta - G(y, \xi)\|_r^2} d\mathbf{y} \right) \int_U C e^{-\frac{1}{2} \|\delta - G(y, \xi)\|_r^2} d\mathbf{y} d\delta$$

QMC weights for the **outer integral**:

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

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

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

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

The nested integral

Goal of computation:

$$\mathcal{I}(f) = \int_Y g \left(\int_U f(\mathbf{y}, \boldsymbol{\delta}, \boldsymbol{\xi}) d\mathbf{y} \right) d\boldsymbol{\delta},$$

where $g(x) := x \log x$, $Y = [-K, K]^k$, and $f(\mathbf{y}, \boldsymbol{\delta}, \boldsymbol{\xi}) := C e^{-\frac{1}{2} \|\boldsymbol{\delta} - G(\mathbf{y}, \boldsymbol{\xi})\|_{\Gamma}^2}$.

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

$$I^{(1)}F := \int_Y F(\boldsymbol{\delta}) d\boldsymbol{\delta} \approx 2^{-\ell} \sum_{k=1}^{2^\ell} F(\boldsymbol{\delta}_k^{(\ell)}) =: Q_\ell^{(1)}F, \ell = \ell_0^{(1)}, \ell_0^{(1)}+1, \ell_0^{(1)}+2, \dots,$$

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

$$I^{(2)}F := \int_U F(\mathbf{y}) d\mathbf{y} \approx 2^{-\ell} \sum_{k=1}^{2^\ell} F(\mathbf{y}_k^{(\ell)}) =: Q_\ell^{(2)}F, \ell = \ell_0^{(2)}, \ell_0^{(2)}+1, \ell_0^{(2)}+2, \dots,$$

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

Why full tensor product cubature is a bad idea

Approximating the integral

$$\mathcal{I}(f) = \int_Y g \left(\int_U f(\mathbf{y}, \boldsymbol{\delta}, \boldsymbol{\xi}) \, \mathrm{d}\mathbf{y} \right) \mathrm{d}\boldsymbol{\delta},$$

by

$$\mathcal{I}(f) \approx Q_\ell^{(1)} g(Q_\ell^{(2)} f) \tag{5}$$

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 (5) takes $N = n^2$ function calls, but the cubature accuracy will not be better than $\mathcal{O}(n^{-\alpha}) = \mathcal{O}(N^{-\alpha/2})$
→ 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

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

under some additional technical assumptions.

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

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

equipped with the parametric diffusion coefficient

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

Numerical experiment

The goal is to find a design ξ^* from the set

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

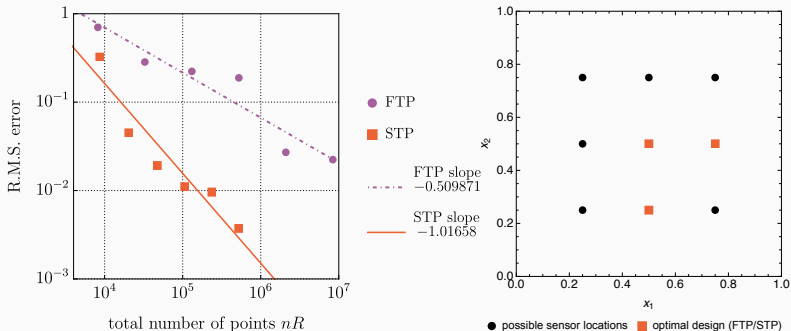
where

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

maximizing the expected information gain subject to the observation operator

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

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

- Modeling the uncertain inputs using Gevrey regular parameterizations leads to dimension-independent QMC convergence rates when computing high-dimensional integrals over the posterior.
- Gevrey regular random fields cover a wider range of potential parameterizations for the uncertain input than those covered by affine and uniform models.
- This approach could be extended to simultaneous recovery of the domain shape and diffusion coefficient a .
- QMC for DOE: sparse approach can recover almost the optimal rate. Future work: optimizing electrode positions, input current patterns or the measurement geometry for Bayesian inversion?

References



A. Djurdjevac, V. Kaarnioja, M. Orteu, and C. Schillings. *Quasi-Monte Carlo for Bayesian shape inversion governed by the Poisson problem subject to Gevrey regular domain deformations*. Preprint 2025, arXiv:2502.14661 [math.NA].



A. Djurdjevac, V. Kaarnioja, C. Schillings, and A. Zepernick. *Uncertainty quantification for stationary and time-dependent PDEs subject to Gevrey regular random domain deformations*. Preprint 2025, arXiv:2502.12345 [math.NA].



L. Bazahica, V. Kaarnioja, and L. Roininen. *Uncertainty quantification for electrical impedance tomography using quasi-Monte Carlo methods*. Preprint 2024, arXiv:2411.11538 [math.NA].



V. Kaarnioja and C. Schillings. *Quasi-Monte Carlo for Bayesian design of experiment problems governed by parametric PDEs*. Preprint 2024, arXiv:2405.03529 [math.NA].

Thank you for your attention!