Certified Randomized Model Order Reduction Methods for High-Dimensional Approximation

Kathrin Smetana (Stevens Institute of Technology)

June 24, 2026

Symposium "Mathematical and Applied Aspects of Complexity Reduction Methods"
in honor of Yvon Maday's chair at the Collège de France

Funding: National Science Foundation (NSF): Award Number: 2145364

Motivation

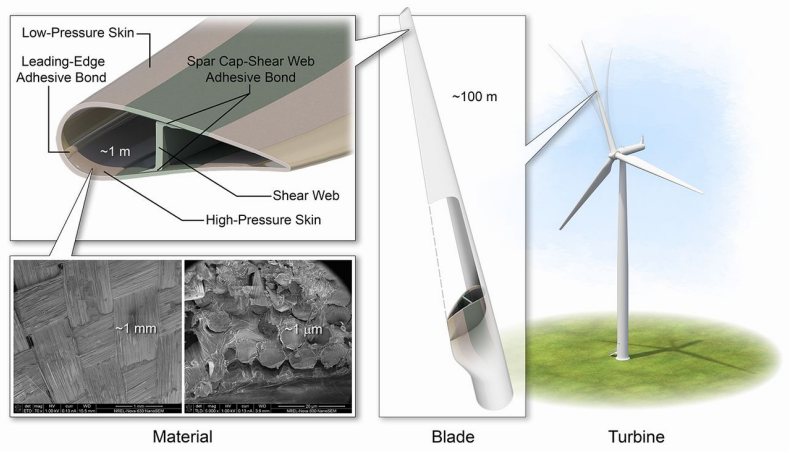


Figure: Picture by NREL, illustrations: Besiki Kazaishvili (NREL), published in [Veers et al., Science, 2019]

Large-scale and multiscale problems, Digital Twins

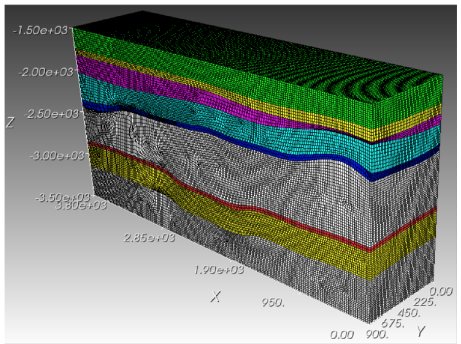


Figure: by J. Trampert

Inverse problems and Uncertainty Quantification

Goal: Discovering Low-Dimensional Structure

Goal: Approximate the set

$$\mathcal{M} := \{u(\mu) \in V : \ u(\mu) \text{ solves PDE dependent on } \mu \in \mathcal{P}\},$$

where

- $\mathcal{P}$: set of admissible parameters.
- V: Hilbert space.
- $\mathcal{P} \subset \mathbb{R}^P$: introduce probability space $(\mathcal{P}, \mathcal{F}, \rho)$.
- $\mathcal{P}$ potentially infinite dimensional: We write $u(Y)$, where $Y : \Omega \to \mathcal{P}$ random variable with values in $\mathcal{P}$ for probability space $(\Omega, \mathcal{F}, \rho)$.

Goal: Discovering Low-Dimensional Structure

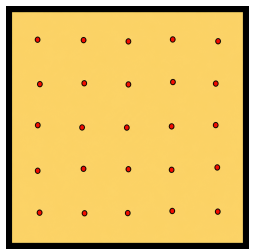
Goal: Approximate the set

$$\mathcal{M} := \{u(\mu) \in V : \ u(\mu) \text{ solves PDE dependent on } \mu \in \mathcal{P}\},$$

where

- $\mathcal{P}$: set of admissible parameters.
- V: Hilbert space.
- $\mathcal{P} \subset \mathbb{R}^P$: introduce probability space $(\mathcal{P}, \mathcal{F}, \rho)$.
- $\mathcal{P}$ potentially infinite dimensional: We write $u(Y)$, where $Y : \Omega \to \mathcal{P}$ random variable with values in $\mathcal{P}$ for probability space $(\Omega, \mathcal{F}, \rho)$.

Challenge: Curse of dimensionality [Bellman 57, 61]: The cardinality of the set $\mathcal{P}_{train}$, which is used for calculations typically scales exponentially with the dimension of $\mathcal{P}$.



Goal: Discovering Low-Dimensional Structure

Goal: Approximate the set

$$\mathcal{M} := \{u(\mu) \in V : \ u(\mu) \text{ solves PDE dependent on } \mu \in \mathcal{P}\},$$

where

- $\mathcal{P}$: set of admissible parameters.
- V: Hilbert space.
- $\mathcal{P} \subset \mathbb{R}^P$: introduce probability space $(\mathcal{P}, \mathcal{F}, \rho)$.
- $\mathcal{P}$ potentially infinite dimensional: We write $u(Y)$, where $Y : \Omega \to \mathcal{P}$ random variable with values in $\mathcal{P}$ for probability space $(\Omega, \mathcal{F}, \rho)$.

Idea: Exploit that many quantities we are interested in behave like **sub-Gaussian random variables - e.g., bounded random variables are sub-Gaussian random variables.** Yields **concentration properties that remain effective in high-dimensional settings.**

Outline

- Motivation
- **Non-asymptotic bounds for Proper Orthogonal Decomposition (POD)/ Principal Component Analysis (PCA)**
- Optimal sampling
- Randomized Greedy algorithm

Marissa Whitby

(PhD Student at Stevens)

Zhiyu Yin

(former Master student at Stevens)

Tommaso Taddei

(Sapienza Università di Roma)

POD/PCA: The Optimal Approximation Space

- We seek an n-dimensional space $V_n \subset V$ minimizing

$$\min_{\dim(V_n)=n} \mathbb{E}\Big[\|u(Y) - P_{V_n} u(Y)\|_V^2\Big].$$

- Define the covariance operator

$$\Sigma = \mathbb{E}\Big(u(Y) \otimes u(Y)\Big).$$

- Then $V_n = \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}$, where φ_j are eigenfunctions of

$$\Sigma\varphi_j = \lambda_j \varphi_j, \qquad \lambda_1 \geqslant \lambda_2 \geqslant \cdots.$$

- Optimal POD error:

$$\mathbb{E}\Big[\|u(Y) - P_{V_n} u(Y)\|_V^2\Big] = \sum_{j>n} \lambda_j.$$

References: [Pearson 1901], [Jolliffe 02], [Kahlbacher, Volkwein 07], [Holmes et al 12], ...

The Monte Carlo Approximation

- Challenge: Computing the covariance operator is usually infeasible due to expectation.

$\Longrightarrow$ Build Monte Carlo approximation from snapshots $u(Y_1), \ldots, u(Y_m)$:

$$\hat{\Sigma} = \frac{1}{m} \sum_{i=1}^{m} u(Y_i) \otimes u(Y_i).$$

- Then $\hat{V}_n = \operatorname{span}\{\hat{\varphi}_1, \ldots, \hat{\varphi}_n\}$, where $\hat{\varphi}_j$ are eigenfunctions of

$$\hat{\Sigma}\hat{\varphi}_j = \hat{\lambda}_j \hat{\varphi}_j, \qquad \hat{\lambda}_1 \geqslant \hat{\lambda}_2 \geqslant \cdots .$$

- Empirical POD error:

$$\mathbb{E}\Big[\|u(Y) - P_{\hat{V}_n} u(Y)\|_V^2\Big].$$

How many snapshots do we need to get a small empirical POD/PCA error?

Theorem 1 (Non-asymptotic bounds for POD/PCA in expectation)

Assume: $u(Y)$ *is bounded random variable with values in* V*. There exists* $c_1 > 0$ *such that* $c_1(\lambda_n - \lambda_{N_V}) \leqslant (\lambda_n - \lambda_{n+1})$ *and*

$$1 \geqslant \frac{4\|u(Y)\|_{L^\infty(\Omega,V)}\sqrt{\lambda_n}}{\sqrt{m}(\lambda_n - \lambda_{n+1})} + \frac{4\|u(Y)\|^2_{L^\infty(\Omega,V)}}{m(\lambda_n - \lambda_{n+1})}.$$

Then it follows that

$$\mathbb{E}\left[\mathbb{E}\left[\|u(Y) - P_{\widehat{V}_n}u(Y)\|_V^2\right]\right] \leqslant \left[\frac{16\|u(Y)\|^2_{L^\infty(\Omega,V)}}{m(\lambda_n - \lambda_{n+1})c_1} + 1\right] \sum_{j>n}\lambda_j + \textit{Concentration term.}$$

- Neglected eigenvalues are mulitplied with inverse of number of samples. $\Longrightarrow$ Can benefit from rapid eigenvalue decay.
- No dependence on dimension of parameter space!

Bound for fixed spectral gap or polynomial eigenvalue decay

- **Fixed spectral gap:** The bound simplifies to

$$\mathbb{E}\left[\mathbb{E}\left[\|u(Y) - P_{\hat{V}_n} u(Y)\|^2\right]\right] \lesssim C_{\text{gap}} \left(\left[1 + \frac{1}{m}\right] \sum_{j>n} \lambda_j + \left[\frac{n^2}{m} + n \sum_{k>n} \frac{\lambda_k}{\lambda_{n+1}}\right] e^{-c_{\text{gap}} m}\right)$$

- **Polynomial eigenvalue decay:** For $\lambda_n = n^{-\alpha}$ and $m \sim n^{\alpha}$, the bound simplifies to

$$\mathbb{E}\left[\mathbb{E}\left[\|u(Y) - P_{\hat{V}_n} u(Y)\|^2\right]\right] \lesssim C_n \left(\left[1 + \frac{1}{m}\right] \sum_{j>n} \lambda_j + \left[n + \frac{n^{2-\frac{\alpha}{2}}}{\sqrt{m}} + \frac{n^2}{m}\right] e^{-c_n m}\right).$$

Remark: n^2 and $n^{2-\frac{\alpha}{2}}$ terms are likely pessimistic

How do the results advance the state of the art

- Closest non-asymptotic result for bounded random variables [Holodnak, Ipsen, Smith 18]: Number of samples needs to go like $1/\varepsilon^2$, where
  - ε is the target relative error in the approximation of the covariance matrix.
  - The approximation error of the covariance matrix must be smaller than the spectral gap in order to reliably identify the dominant eigenspace.

- [Reiss, Wahl 20]: Paper requires that sub-Gaussian norm of random variable $(u(Y), v)_V$ is bounded by its variance for all $v \in V$:

$$\sup_{k \geqslant 1} \frac{1}{\sqrt{k}} \mathbb{E}\left[|(u(Y), v)_V|^k \right]^{1/k} \leqslant C_1 \mathbb{E}\left[(u(Y), v)_V^2 \right]^{1/2}.$$

This assumption is often difficult to verify and does not necessarily hold for bounded random variables (counter example!).

How do the results advance the state of the art

- Closest non-asymptotic result for bounded random variables [Holodnak, Ipsen, Smith 18]: Number of samples needs to go like $1/\varepsilon^2$, where
  - ε is the target relative error in the approximation of the covariance matrix.
  - The approximation error of the covariance matrix must be smaller than the spectral gap in order to reliably identify the dominant eigenspace.

- [Reiss, Wahl 20]: In numerical experiment we want to look at the POD eigenmodes (here variance goes to zero as eigenvalues go to zero):

$$\|(u(Y), v)_V\|_{L^\infty(\Omega)} \overset{?}{\leqslant} C_1 \mathbb{E}\left[(u(Y), v)_V^2\right]^{1/2}.$$

Potentially difficult for POD modes corresponding to small eigenvalues that correspond to solution features active on small subsets of $\mathcal{P}$.

Numerical results: Parametric nonlinear diffusion

$$-\operatorname{div}(d_\mu(u(x;\mu))\nabla u(x;\mu)) = 0 \quad \text{in } D,$$
$$u(x;\mu) = \mu_2 \quad \text{on } \partial D,$$
$$d_\mu(u(x;\mu)) := \frac{u(\mu)^2(1-u(\mu))^2}{(u(\mu)^3 + \frac{1}{3}(1-u(\mu))^3)^2} + 0.1 + 10^5 \mathbb{1}_{\{\mu_1 > 0.99\}} k(x),$$

where: $k(x)$ high conductivity channels, $\mu_1 \sim Unif(0,1)$, μ_2 interpolates with high probability function in H^1 on boundary

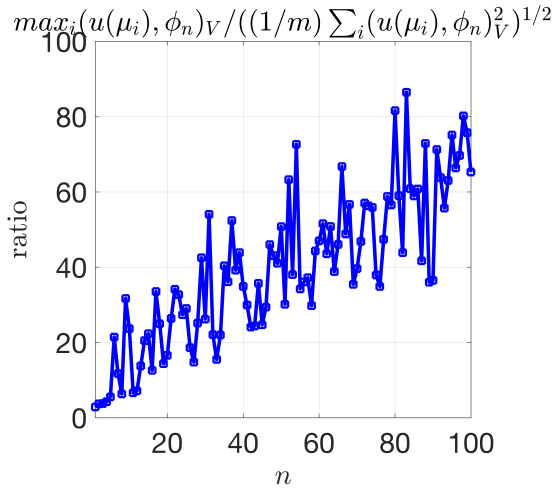
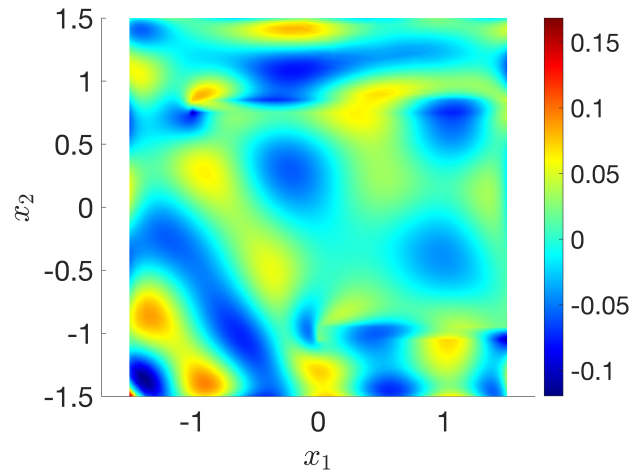


Figure: $\frac{\max_i(u(\mu_i),\widehat{\varphi}_n)_V}{\sqrt{(1/m)\sum_{i=1}^m(u(\mu_i),\widehat{\varphi}_n)_V^2}}$

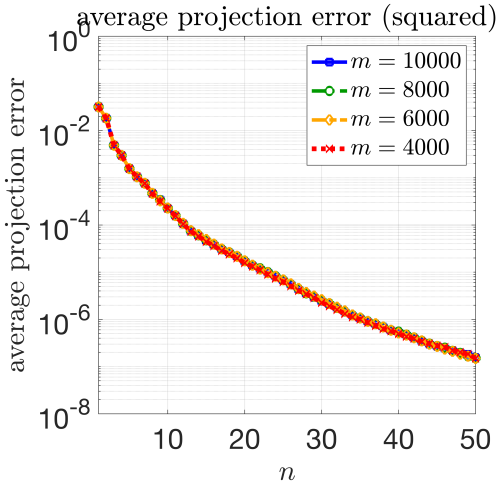


POD mode associated with *max*

Numerical results: Parametric nonlinear diffusion

$$-\operatorname{div}(d_\mu(u(x;\mu))\nabla u(x;\mu)) = 0 \quad \text{in } D,$$
$$u(x;\mu) = \mu_2 \quad \text{on } \partial D,$$
$$d_\mu(u(x;\mu)) := \frac{u(\mu)^2(1-u(\mu))^2}{(u(\mu)^3 + \frac{1}{3}(1-u(\mu))^3)^2} + 0.1 + 10^5 \mathbb{1}_{\{\mu_1 > 0.99\}} k(x),$$

where: $k(x)$ high conductivity channels, $\mu_1 \sim Unif(0,1)$, μ_2 interpolates with high probability function in H^1 on boundary



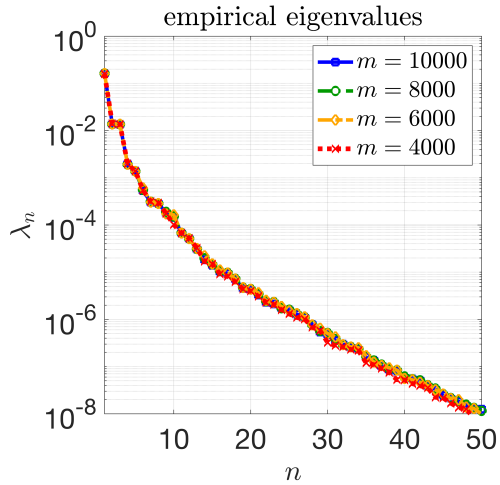


Figure: $\frac{1}{m}\sum_{i=1}^m \|u(\mu_i) - P_{\widehat{V}_n} u(\mu_i)\|_V^2$ $\widehat{\lambda}_n$

Where are we?

POD:

- For many problems a modest number of samples is sufficient to get an accurate approximation independently of the dimension of the parameter space.
- Numerical experiments suggest that assumptions in current non-asymptotic PCA/POD theory may be restrictive for parametric PDEs. Low-energy POD modes can correspond to solution features activated only on small subsets of parameter space, highlighting the relevance of our results.

Where are we?

POD:

- For many problems a modest number of samples is sufficient to get an accurate approximation independently of the dimension of the parameter space.
- Numerical experiments suggest that assumptions in current non-asymptotic PCA/POD theory may be restrictive for parametric PDEs. Low-energy POD modes can correspond to solution features activated only on small subsets of parameter space, highlighting the relevance of our results.

How can we improve:

- Exploit structure of problem
- Sampling strategy

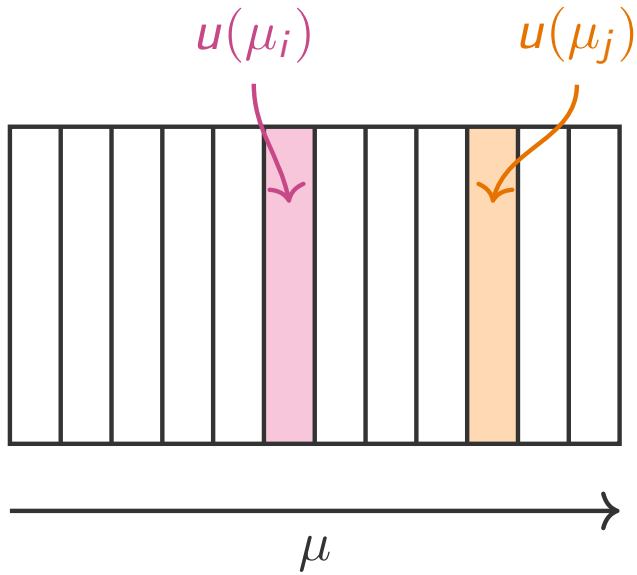
Outline

- Motivation
- Non-asymptotic bounds for POD/PCA
- **Optimal sampling**
- Randomized Greedy algorithm

Selecting Informative Parameters = Column Subset Selection

- $\mathcal{M} = \{u(\mu) : \mu \in \mathcal{P}\} \quad \rightsquigarrow \quad M = \left[u(\mu_1) \ \cdots \ u(\mu_N)\right]$
- Equivalent problem: Find subset of columns $C = [u(\mu_1) \ \cdots \ u(\mu_m)]$ such that $\|M - CC^\dagger M\|_F$ is small with high probability ($C^\dagger$: Moore-Penrose inverse).

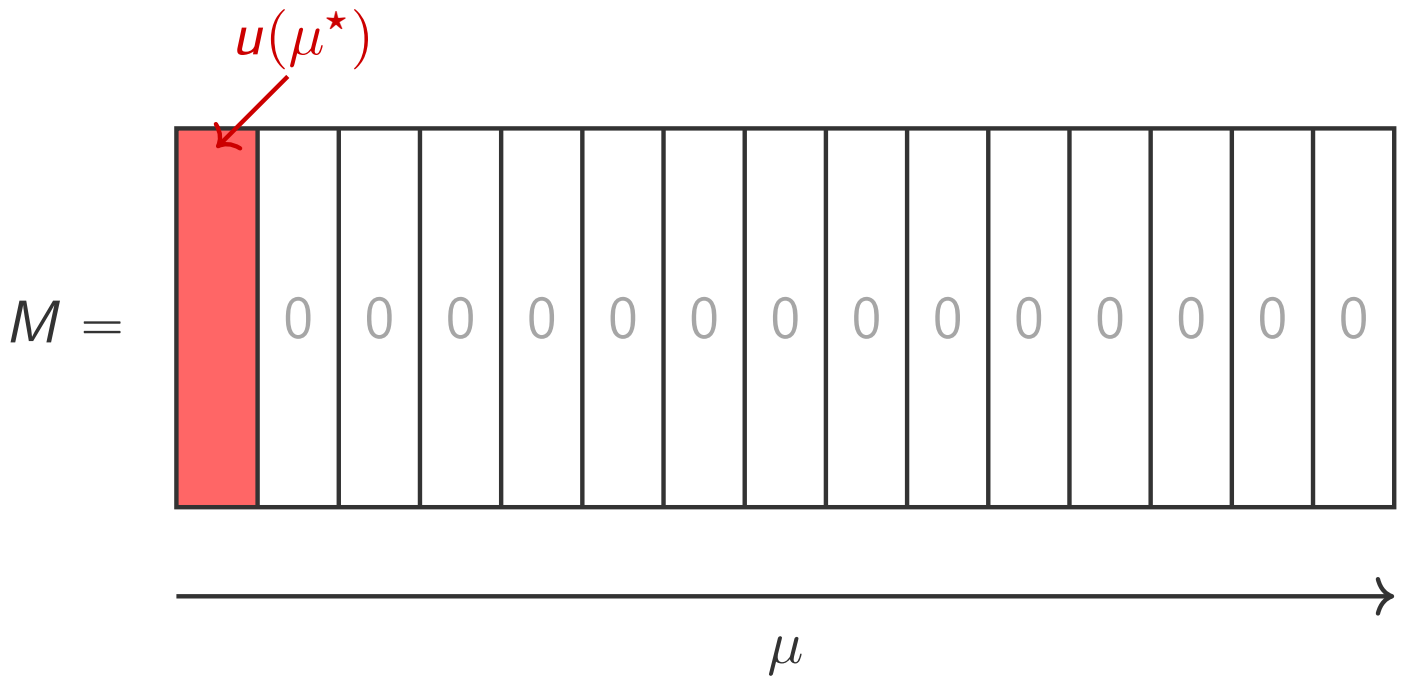
$M =$



- What are good strategies for column subset selection?

Why Uniform Sampling Can Fail

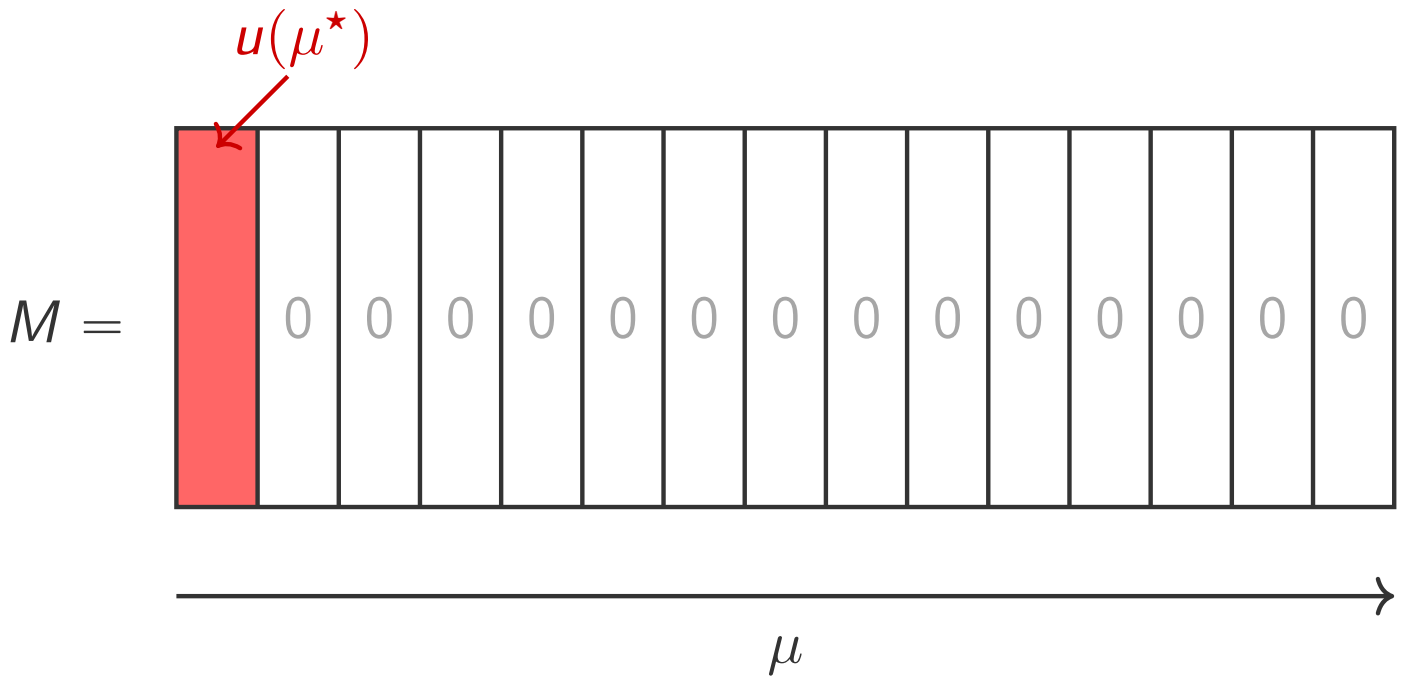
Uniform sampling: Choose parameters $\mu_1, \ldots, \mu_m$ independently and uniformly at random.



Worst-case scenario: The information relevant for approximation is highly localized: it is concentrated in a small subset of the parameter space and, equivalently, in only a few columns of the matrix.

Why Uniform Sampling Can Fail

Uniform sampling: Choose parameters $\mu_1, \ldots, \mu_m$ independently and uniformly at random.



How can we identify the informative regions of the parameter space?

Squared norm sampling

Squared norm sampling: Select the parameter/column with index i with probability $p_i = \|M[:,i]\|_2^2 / (\sum_{j=1}^{N} \|M[:,i]\|_2^2)$.

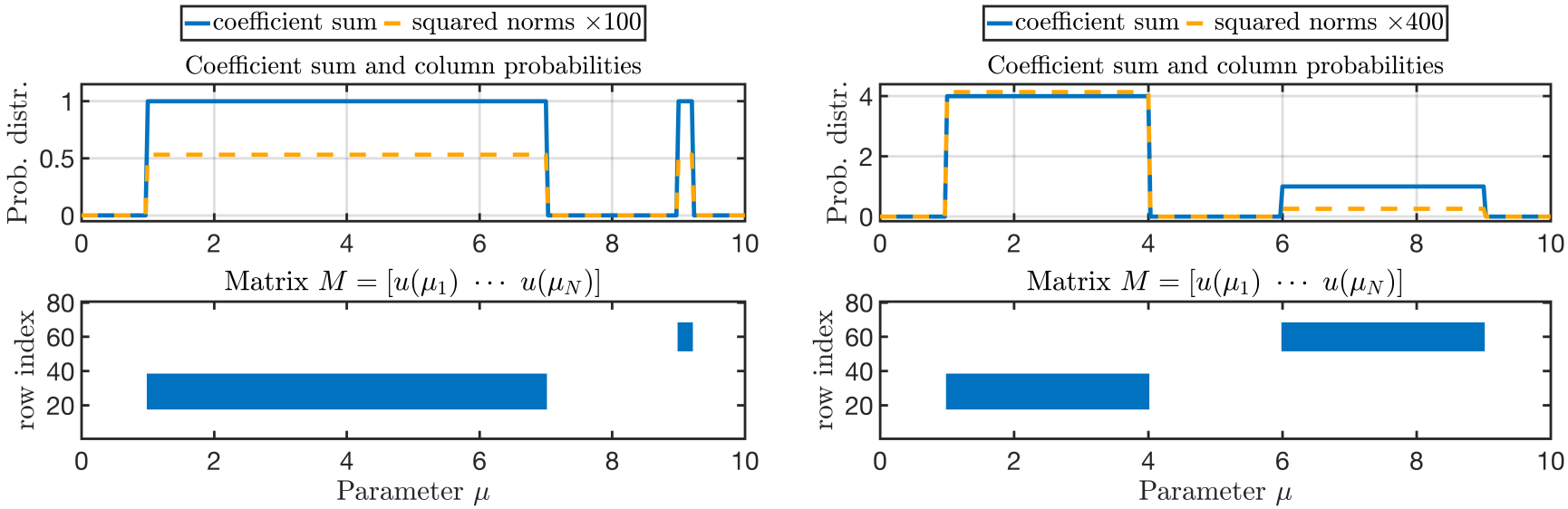


Figure: Examples adapted from [Schleuß, ter Maat, KS, SISC, 23]

Challenges: Data that has values on smaller scales (right example). Highly localized information can be still an issue.

Squared norm sampling

Squared norm sampling: Select the parameter/column with index i with probability $p_i = \|M[:,i]\|_2^2 / (\sum_{j=1}^{N} \|M[:,j]\|_2^2)$.

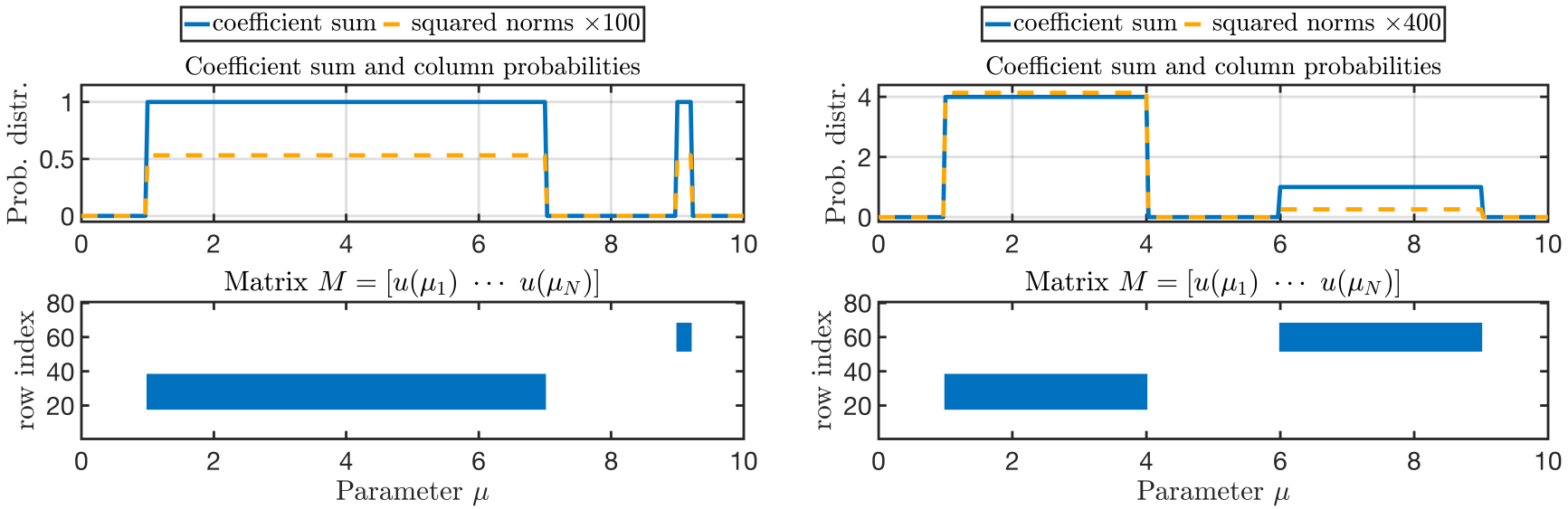


Figure: Examples adapted from [Schleuß, ter Maat, KS, SISC, 23]

How can we sample properly from areas that are important for approximation?

Leverage score sampling

Leverage score sampling: Select the parameter/column with index i with probability $p_i = (1/n)\sum_{j=1}^{n} v_j[i]$, where $v_1, \ldots, v_n$ are the leading n right singular vectors of M.

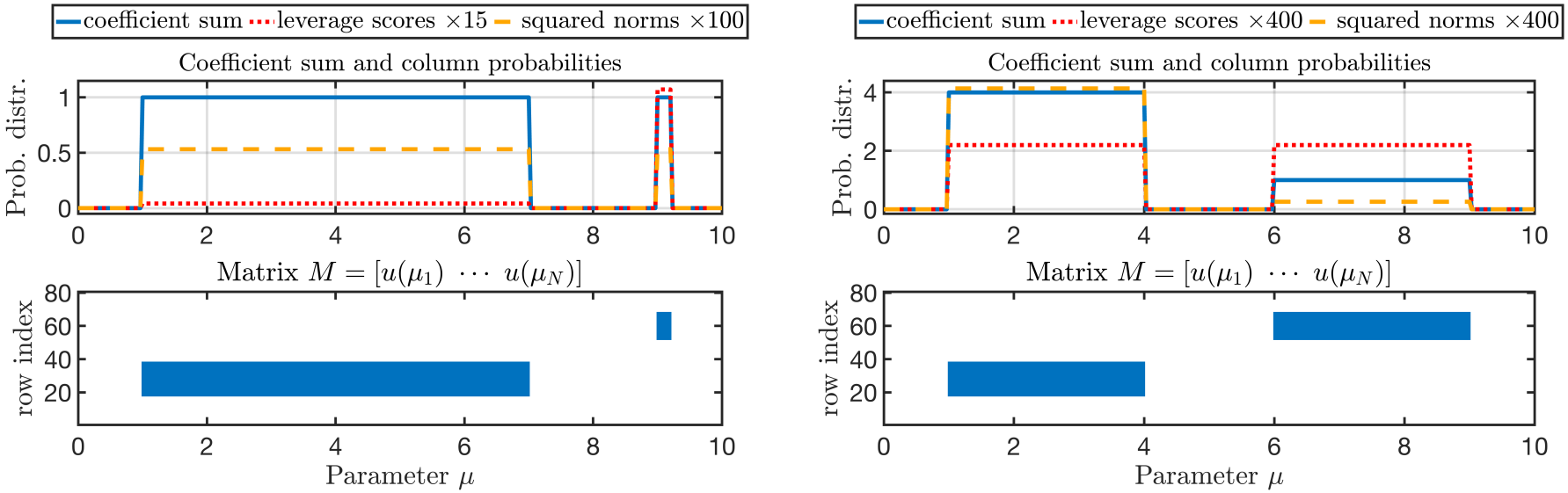


Figure: Examples adapted from [Schleuß, ter Maat, KS, SISC, 23]

Leverage score sampling chooses columns which have a large influence on the best rank-n fit of M.

Leverage score sampling

Leverage score sampling: Select the parameter/column with index i with probability $p_i = (1/n)\sum_{j=1}^{n} v_j[i]$, where $v_1, \ldots, v_n$ are the leading n right singular vectors of M.

Leverage score sampling yields an approximation which is, with high probability, within a factor $1+\varepsilon$ of the best rank-n approximation:

Theorem 2 ((Drineas, Mahoney, Muthukrishnan, SIMAX, 2018))

Let $\varepsilon \in (0,1]$ and $m \in \mathcal{O}(n^2/\varepsilon^2)$. Then with high probability there holds

$$\|M - CC^{\dagger}M\|_F \leqslant (1+\varepsilon)\|M - M_n\|_F,$$

where M_n: best rank-n approximation to M.

From Informative Columns to Informative Parameters

Discrete setting

- $M = \begin{bmatrix} u(\mu_1) & \cdots & u(\mu_M) \end{bmatrix}$

Continuous setting

- $\mathcal{M} = \{u(\mu) : \mu \in \mathcal{P}\}$

Approximation space: $U_n = \text{span}\{\phi_1, \ldots, \phi_n\}$ with $\{\phi_i\}_i$ orthogonal

- Leverage scores:

$$\ell_i = \frac{1}{n} \sum_{k=1}^{n} \frac{|\phi_k(\mu_i)|^2}{\|\phi_k\|_2^2}$$

- Leverage score distribution

$$p_i = \ell_i$$

- Christoffel function:

$$\mathcal{C}_n(\mu) = \sum_{k=1}^{n} \frac{|\phi_k(\mu)|^2}{\int_{\mathcal{P}} |\phi_k|^2 \, d\rho}$$

- Christoffel measure:

$$\frac{\mathrm{d}\nu_n}{\mathrm{d}\rho}(\mu) := \mathcal{C}_n(\mu)$$

Christoffel sampling is the natural continuous analogue of leverage-score sampling and thus quantifies approximation relevance.

From Informative Columns to Informative Parameters

Discrete setting

- $M = \begin{bmatrix} u(\mu_1) & \cdots & u(\mu_M) \end{bmatrix}$

Continuous setting

- $\mathcal{M} = \{u(\mu) : \mu \in \mathcal{P}\}$

Approximation space: $U_n = \text{span}\{\phi_1, \ldots, \phi_n\}$ with $\{\phi_i\}_i$ orthogonal

- Leverage scores:

$$\ell_i = \frac{1}{n} \sum_{k=1}^{n} \frac{|\phi_k(\mu_i)|^2}{\|\phi_k\|_2^2}$$

- Leverage score distribution

$$p_i = \ell_i$$

- Christoffel function:

$$\mathcal{C}_n(\mu) = \sum_{k=1}^{n} \frac{|\phi_k(\mu)|^2}{\int_{\mathcal{P}} |\phi_k|^2 \, d\rho}$$

- Christoffel measure:

$$\frac{\mathrm{d}\nu_n}{\mathrm{d}\rho}(\mu) := \mathcal{C}_n(\mu)$$

Results by Cohen, Dolbeault, Migliorati, Narayan and Xu show that Christoffel sampling performs well in detecting large values → Christoffel sampling is also great for certification.

Outline

- Motivation
- Non-asymptotic bounds for POD/PCA
- Optimal sampling
- **Randomized Greedy algorithm**

Charles Beall (Stevens)

A randomized Greedy algorithm

- Input: Finite probability measure ρ, Error tolerance *tol*
- In each iteration
  1. Draw m_n independent samples $\mu_1, \dots, \mu_{m_n}$ from the measure
  $$\frac{\mathrm{d}\nu_n}{\mathrm{d}\rho}(\mu) := \mathcal{C}_n(\mu)$$
  2. Enrich V_n with the solution $u(\mu^*)$, where
  $$\mu^* = \arg \max_{\mu_j \in \{\mu_1, \dots, \mu_{m_n}\}} \|u(\mu_j) - u_n(\mu_j)\|_V, \quad u_n(\mu) \in V_n \text{ approximates } u(\mu)$$
- Terminate if $\|u(\mu^*) - u_n(\mu^*)\|_V \leqslant tol$.

$\mathcal{C}_n : \mathcal{P} \to \mathbb{R}_{\geqslant 0}$ denotes the **modified Christoffel function**

$$\mathcal{C}_n(\mu) := \frac{1}{2}\left(1 + \frac{\|u(\mu) - u_n(\mu)\|_V^2}{\|u - u_n\|^2_{L^2_\rho(\mathcal{P}, V)}}\right), \quad \mu \in \mathcal{P}, \tag{1}$$

Assume that the PDE is stable for ρ-almost every parameter.

A randomized Greedy algorithm

- Input: Finite probability measure ρ, Error tolerance *tol*
- In each iteration
  1. Draw m_n independent samples $\mu_1, \ldots, \mu_{m_n}$ from the measure
  $$\frac{\mathrm{d}\nu_n}{\mathrm{d}\rho}(\mu) := \mathcal{C}_n(\mu)$$
  2. Enrich V_n with the solution $u(\mu^*)$, where
  $$\mu^* = \arg \max_{\mu_j \in \{\mu_1, \ldots, \mu_{m_n}\}} \|u(\mu_j) - u_n(\mu_j)\|_V, \quad u_n(\mu) \in V_n \text{ approximates } u(\mu)$$
- Terminate if $\|u(\mu^*) - u_n(\mu^*)\|_V \leqslant tol$.

Challenges:

- $L^2(\mathcal{P}, V)$−norm: Replace by estimator $\frac{1}{m_{L^2}} \sum_{j=1}^{m_{L^2}} \|u(\mu_j) - u_n(\mu_j)\|_V^2$ to get a practical algorithm. By Hoeffding's inequality, the practical algorithm enjoys essentially the same theoretical guarantees.
- Sampling: Slice sampling [Neal 03] worked well for us so far.

A randomized Greedy algorithm

- Input: Finite probability measure ρ, Error tolerance *tol*
- In each iteration
  1. Draw m_n independent samples $\mu_1, \ldots, \mu_{m_n}$ from the measure
  $$\frac{\mathrm{d}\nu_n}{\mathrm{d}\rho}(\mu) := \mathcal{C}_n(\mu)$$
  2. Enrich V_n with the solution $u(\mu^*)$, where
  $$\mu^* = \arg \max_{\mu_j \in \{\mu_1, \ldots, \mu_{m_n}\}} \|u(\mu_j) - u_n(\mu_j)\|_V, \quad u_n(\mu) \in V_n \text{ approximates } u(\mu)$$
- Terminate if $\|u(\mu^*) - u_n(\mu^*)\|_V \leqslant tol$.

The error $\|u(\mu) - u_n(\mu)\|_V$ can be replaced an a posteriori error estimator (often more cost efficient).

Theorem 3 (Per-iteration certification)

$(\mathcal{P}, \mathcal{F}, \rho)$: probability space, with $\mathcal{P} \subseteq \mathbb{R}^P$, $P \geqslant 1$.
ρ and Lebesgue measure are equivalent, i.e. every set of positive ρ-measure also has positive Lebesgue measure.
Draw $m_n \in \mathbb{N}$ i.i.d. samples $\mu_1, \ldots, \mu_{m_n} \in \mathcal{P}$ according to ν_n, where

$$m_n \geqslant \frac{2}{\alpha_n^2} \ln\left(\frac{2}{\delta_n}\right), \qquad \text{for given } \alpha_n, \delta_n \in (0,1). \tag{1}$$

Then, with probability at least $1 - \delta_n$, we have

$$\|u - u_n\|_{L^\infty_\rho(\mathcal{P},V)} \leqslant \sqrt{\frac{4\,\|\mathcal{C}_n\|_{L^\infty_\rho(\mathcal{P})}}{1-\alpha_n}} \max_{1\leqslant i\leqslant m_n} \|u(\mu_i) - u_n(\mu_i)\|_V. \tag{2}$$

We can bound the error over the whole parameter set by the max over few samples!

Theorem 3 (Per-iteration certification)

$(\mathcal{P}, \mathcal{F}, \rho)$: probability space, with $\mathcal{P} \subseteq \mathbb{R}^P$, $P \geqslant 1$.
ρ and Lebesgue measure are equivalent, i.e. every set of positive ρ-measure also has positive Lebesgue measure.
Draw $m_n \in \mathbb{N}$ i.i.d. samples $\mu_1, \ldots, \mu_{m_n} \in \mathcal{P}$ according to ν_n, where

$$m_n \geqslant \frac{2}{\alpha_n^2} \ln\left(\frac{2}{\delta_n}\right), \qquad \text{for given } \alpha_n, \delta_n \in (0,1). \tag{1}$$

Then, with probability at least $1 - \delta_n$, we have

$$\|u - u_n\|_{L^\infty_\rho(\mathcal{P},V)} \leqslant \sqrt{\frac{4\,\|\mathcal{C}_n\|_{L^\infty_\rho(\mathcal{P})}}{1-\alpha_n}} \max_{1\leqslant i\leqslant m_n} \|u(\mu_i) - u_n(\mu_i)\|_V. \tag{2}$$

Union bound $\Rightarrow$ **If the algorithm reaches *tol* with output V_N, the approximation is certified for ρ-almost every parameter $\mu \in \mathcal{P}$ with high probability.**

Estimating $\|\mathcal{C}_n\|_{L^\infty_\rho(\mathcal{P})}$

- Observation: For considered numerical test cases $\|\mathcal{C}_n\|_{L^\infty_\rho(\mathcal{P})}$ did not contribute to the convergence rate.

$\Longrightarrow$ It might often not be necessary to get an accurate estimate.

- Exploit that we sample from areas with high $\mathcal{C}_n$ with high probability:
  - Consider set $G_n := \{\mu \in \mathcal{P} : S_n(\mathcal{C}_n(\mu) - \mathbb{E}_\rho[\mathcal{C}_n]) \geqslant \|\mathcal{C}_n\|_{L^\infty_\rho(\mathcal{P})}\}$, where S_n safety factor, $\mathbb{E}_\rho[\mathcal{C}_n] = 1$.
  - Use estimator $S_n\left(\max_{1\leqslant j\leqslant m_{est}} \mathcal{C}_n(\mu_j) - 1\right)$
  - We obtain for the Christoffel measure ν_n that

$$\nu_n(\text{no draw belongs to } G_n) \leqslant \left(\frac{1}{2} + \frac{(S_n^{-1}\|\mathcal{C}_n\|_{L^\infty_\rho(\mathcal{P})} + 1)^2}{\|\mathcal{C}_n\|_{L^\infty_\rho(\mathcal{P})}}\right)^{m_{est}}.$$

$\Longrightarrow$ Very high values will only appear on sets of small measure.

Randomized greedy yields a quasi-optimally converging approximation

- In each iteration the randomized Greedy selects a near maximizing parameter with high probability:

$$\|u - u_n\|_{L^\infty_\rho(\mathcal{P},V)} \leqslant \sqrt{\frac{4\,\|\mathcal{C}_n\|_{L^\infty_\rho(\mathcal{P})}}{1-\alpha_n}} \max_{1\leqslant i\leqslant m_n} \|u(\mu_i) - u_n(\mu_i)\|_V.$$

- The (deterministic) Greedy algorithm [Veroy, Prud'Homme, Rovas, Patera 03] selects the global maximizer in each iteration.

$\Longrightarrow$ Thanks to results in [DeVore, Petrova, Wojtaszczyk 13] for the (deterministic) Greedy algorithm, we conclude that the approximation constructed by the randomized Greedy algorithm converges at a quasi-optimal rate (slightly worse than the Kolmogorov n-width) with high probability.

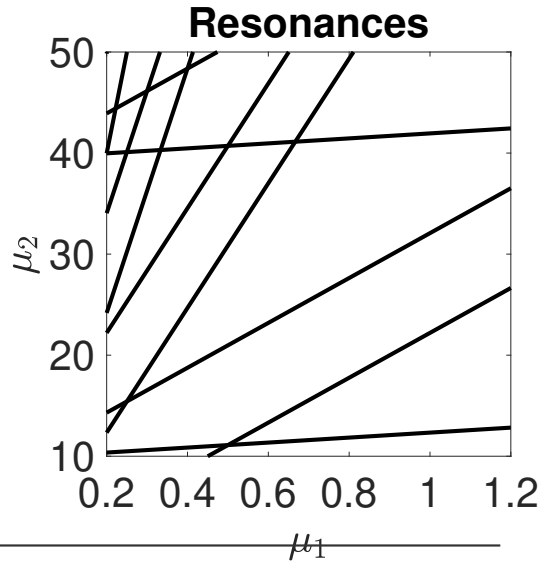
Relation to existing approaches

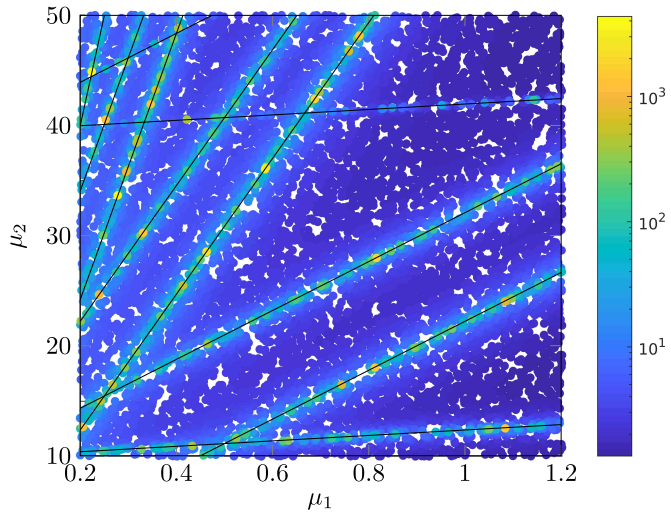
- **Randomized Greedy for model order reduction** [Cohen et al. ESAIM: M2AN '20], [Billaud-Friess et al. Adv. Comput. Math. '24], [Nielen, Tse 26]
- **Adaptive updates/partitioning of the training set:** [Sen Numer. Heat Transfr. B: Fundam. '08], [Eftang, Patera, Rønquist, SIAM J. Sci. Comput. '09], [Haasdonk, Dihlmann, Ohlberger Math. Comput. Model. Dyn. Syst. '11], [Eftang, Stamm Int. J. Numer. Methods Eng. '12], [Hesthaven, Stamm, Zhang ESAIM: M2AN '14], [Jiang, Chen, Narayan J. Sci. Comput. '17], [Jiang, Chen Int. J. Numer. Methods Eng. '20], [Nielen, Tse, Veroy-Grepl SISC 25]
- **Nonlinear optimization-based approach:** [Urban, Volkwein, Zeeb, ROM for Modeling & Comput. Reduction '14]
- **Low-rank tensor approaches:** [Khoromskij, Schwab SIAM J. Sci. Comput. '11], [Ballani, Kressner SIAM J. Sci. Comput. '16], work by R. Scheichl, A. Nouy, M. Olshanskii, ...

Multiparametric Helmholtz problem[1]

Given $\mu = (\mu_1, \mu_2) \in \mathcal{P} = [0.2, 1.2] \times [10, 50]$, find $u(\mu)$ such that

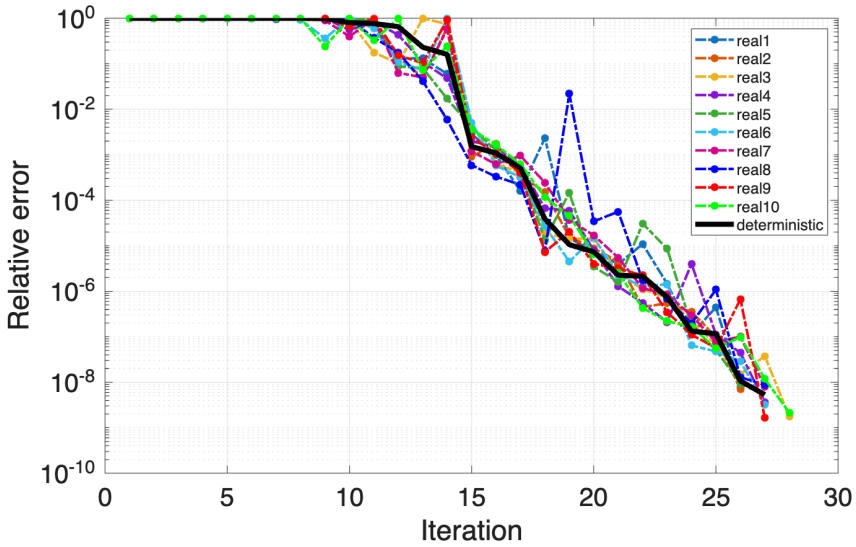
$$
\begin{aligned}
-\frac{\partial^2 u(\mu)}{\partial x^2} - \mu_1 \frac{\partial^2 u(\mu)}{\partial y^2} - \mu_2 u &= f(x,y) && \text{in } D = (0,1)^2,\\
u(\mu) &= 0 && \text{on } (0,1) \times \{0\},\\
\frac{\partial u(\mu)}{\partial y} &= \cos(\pi x) && \text{on } (0,1) \times \{1\},\\
\frac{\partial u(\mu)}{\partial x} &= 0 && \text{on } \{0,1\} \times (0,1).
\end{aligned}
$$





[1] see [KS, Zahm, Patera 19] and [Huynh et al. 10]

Convergence results



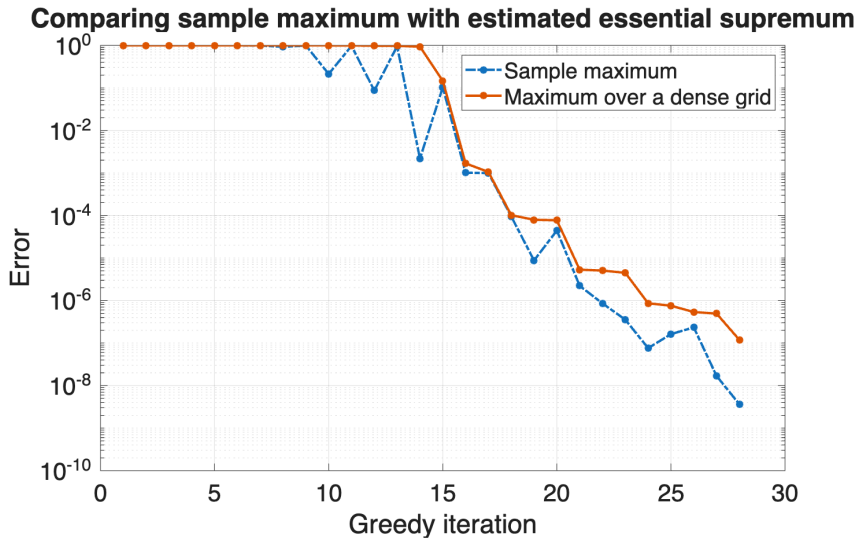


Figure: 40 samples per iteration, cardinality of training set for deterministic Greedy algorithm: 1200, maximum of error estimated on set of $22,500$ samples.

Summary

- Goal: Approximate $\mathcal{M} := \{u(\mu) : u(\mu) \text{ solves PDE dependent on } \mu \in \mathcal{P}\}$
- Idea: Exploit that under mild conditions many quantities we are interested in behave like sub-Gaussian random variables.
- POD: For many problems a modest number of samples is sufficient to get an accurate approximation independently of the dimension of the parameter space.
- Data dependent sampling strategies are well-suited to sample form areas of the parameter set that are relevant for approximation.
- Randomized Greedy algorithm provides certified and quasi-optimally converging reduced approximations for almost every $\mu \in \mathcal{P}$ by relying on a measure that ensures sampling from areas with high approximation error.

Thank you very much for your attention!

Appendix

Appendix

Theorem 4 (Non-asymptotic bounds for POD/PCA)

Assume: $u(Y)$ *is bounded random variable with values in Hilbert space* V *of dimension* N_V. *There exists* $c_1 > 0$ *such that* $c_1(\lambda_n - \lambda_{N_V}) \leqslant (\lambda_n - \lambda_{n+1})$ *and*

$$1 \geqslant \frac{4\|u(Y)\|_{L^\infty(\Omega,H)}\sqrt{\lambda_n}}{\sqrt{M}(\lambda_n - \lambda_{n+1})} + \frac{4\|u(Y)\|^2_{L^\infty(\Omega,H)}}{M(\lambda_n - \lambda_{n+1})}.$$

Then it follows that

$$\mathbb{E}\left[\mathbb{E}\left[\|u(Y) - P_{\hat{H}_n}u(Y)\|^2_H\right]\right] \leqslant c_1^{-1}\left(\frac{16\|u(Y)\|^2_{L^\infty}}{M(\lambda_n - \lambda_{n+1})}\sum_{j>n}\lambda_j\right.$$

$$+8\sum_{j\leqslant n}(\lambda_j - \lambda_{n+1})\sum_{k>n}\frac{\lambda_k}{\lambda_{n+1}}\exp\left(-\frac{M}{4}\min\left(\frac{(\lambda_j - \lambda_{n+1})^2}{4\|u(Y)\|^2_{L^\infty(\Omega,H)}\lambda_{n+1}}, \frac{(\lambda_j - \lambda_{n+1})}{2\|u(Y)\|^2_{L^\infty(\Omega,H)}}\right)\right)$$

$$+\frac{32\|u(Y)\|^2_{L^\infty}}{M}\sum_{i=1}^{n}\frac{\lambda_{n+1-i}(\lambda_n - \lambda_{n+1})}{(\lambda_{n+1-i} - \lambda_{n+1})\lambda_n}\left(\sqrt{6}\sum_{j\leqslant n}\frac{\lambda_j}{(\lambda_j - \lambda_{n+1})} + \sqrt{8}\sum_{j\leqslant n}\frac{\|u(Y)\|_{L^\infty}\sqrt{\lambda_j}}{\sqrt{M}(\lambda_j - \lambda_{n+1})}\right)$$

$$\left.\cdot\exp\left(-M\frac{(\lambda_n - \lambda_{n+1})^2}{64\lambda_n\|u(Y)\|^2_{L^\infty}}\right)\right) + \sum_{j>n}\lambda_j$$

Sketch of proof of theorem

Part I. Establish Marcinkiewicz-Zygmund (MZ) inequalities in $L^2_\nu(\mathcal{P})$:

1. We introduce the weighted estimator

$$\frac{1}{m}\sum_{j=1}^{m}\frac{1}{\mathcal{C}(\mu_j)}\phi^2(\mu_j), \quad \mu_j \overset{\text{i.i.d.}}{\sim} \nu, \quad \phi \in \text{span}\{\mathcal{E}\}, \quad \mathcal{E}(\mu) = \|u(\mu) - u_n(\mu)\|_V.$$

2. Hoeffding's inequality for bounded random variables (see e.g., [Vershynin 18]) yields:

$$\mathbb{P}_\nu\left(\left|\frac{1}{m}\sum_{j=1}^{m}\frac{1}{\mathcal{C}(\mu_j)}\phi^2(\mu_j) - \|\phi\|^2_{L^2_\rho(\mathcal{P})}\right| \geqslant \alpha\right) \leqslant 2\exp\left\{-\frac{M\alpha^2}{2}\right\}.$$

and thus MZ inequalities with probability at least $1-\delta$, $\delta = 2\exp\left\{-\frac{M\alpha^2}{2}\right\}$:

$$(1-\alpha)\|\phi\|^2_{L^2_\rho(\mathcal{P})} \leqslant \frac{1}{m}\sum_{j=1}^{m}\frac{1}{\mathcal{C}(\mu_j)}\phi^2(\mu_j) \leqslant (1+\alpha)\|\phi\|^2_{L^2_\rho(\mathcal{P})}. \tag{3}$$

Sketch of proof of theorem

Part II. Lift Marcinkiewicz-Zygmund inequalities from $L^2_\rho(\mathcal{P})$ to $L^\infty_\rho(\mathcal{P})$:

1. The definition of $\mathcal{C}$ gives the key inequality (see also e.g., [Xu, Narayan 23]):

$$\|\mathcal{E}\mathcal{C}^{-1/2}\|^2_{L^\infty_\rho(\mathcal{P})} \leqslant 2\|\mathcal{E}\mathcal{C}^{-1/2}\|^2_{L^2_\nu(\mathcal{P})}. \tag{4}$$

2. Applying Marcinkiewicz-Zygmund inequalities in $L^2_\nu(\mathcal{P})$ for $\mathcal{E}$ gives:

$$\frac{1}{\|\mathcal{C}\|_{L^\infty_\rho(\mathcal{P})}}\|\mathcal{E}\|^2_{L^\infty_\rho(\mathcal{P})} \leqslant \|\mathcal{E}\mathcal{C}^{-1/2}\|^2_{L^\infty_\rho(\mathcal{P})} \overset{(4)}{\leqslant} 2\|\mathcal{E}\mathcal{C}^{-1/2}\|^2_{L^2_\nu(\mathcal{P})} = 2\|\mathcal{E}\|^2_{L^2_\rho(\mathcal{P})}$$

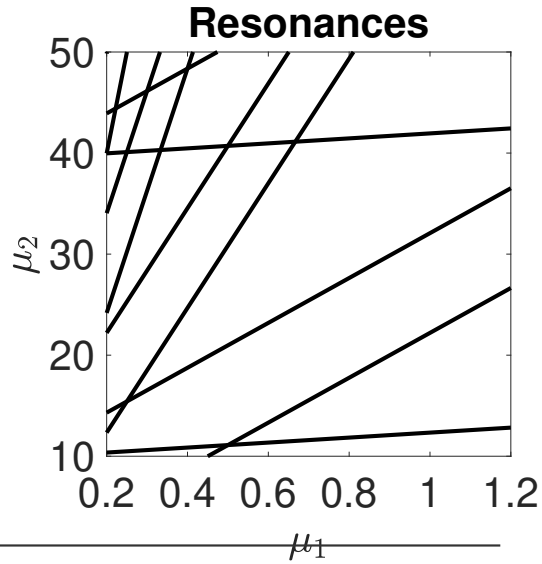
$$\overset{MZ}{\leqslant} \frac{2}{m(1-\alpha)}\sum_{j=1}^{m}|\mathcal{E}\mathcal{C}^{-1/2}(\mu_j)|^2 \leqslant \frac{4}{1-\alpha}\left\{\max_{1\leqslant j\leqslant m}\mathcal{E}(\mu_j)\right\}^2.$$

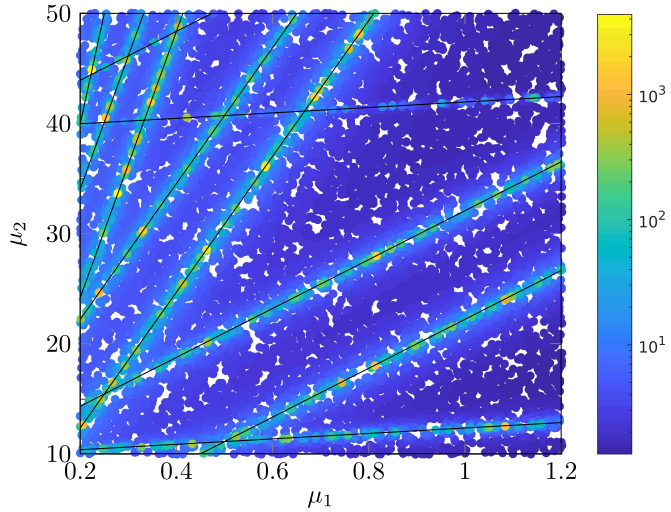
References: Part I makes use of classical arguments in sampling discretization theory (e.g., [Cohen, Migliorati 17])

Multiparametric Helmholtz problem[2]

Given $\mu = (\mu_1, \mu_2) \in \mathcal{P} = [0.2, 1.2] \times [10, 50]$, find $u(\mu)$ such that

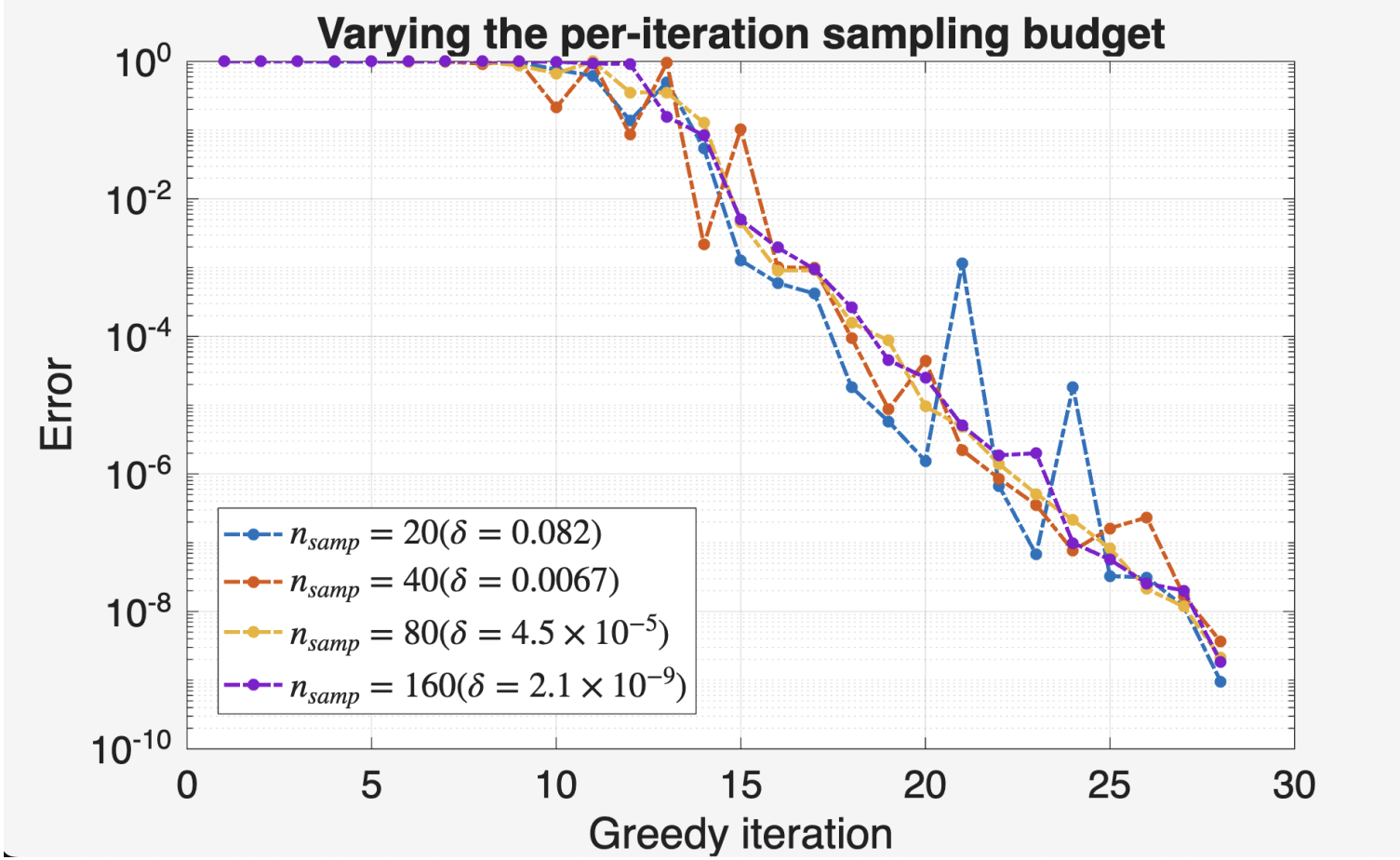
$$
\begin{aligned}
-\frac{\partial^2 u(\mu)}{\partial x^2} - \mu_1 \frac{\partial^2 u(\mu)}{\partial y^2} - \mu_2 u &= f(x,y) && \text{in } D = (0,1)^2, \\
u(\mu) &= 0 && \text{on } (0,1) \times \{0\}, \\
\frac{\partial u(\mu)}{\partial y} &= \cos(\pi x) && \text{on } (0,1) \times \{1\}, \\
\frac{\partial u(\mu)}{\partial x} &= 0 && \text{on } \{0,1\} \times (0,1).
\end{aligned}
$$





[2] see [KS, Zahm, Patera 19] and [Huynh et al. 10]

Convergence results



Parameters selected by the randomized Greedy

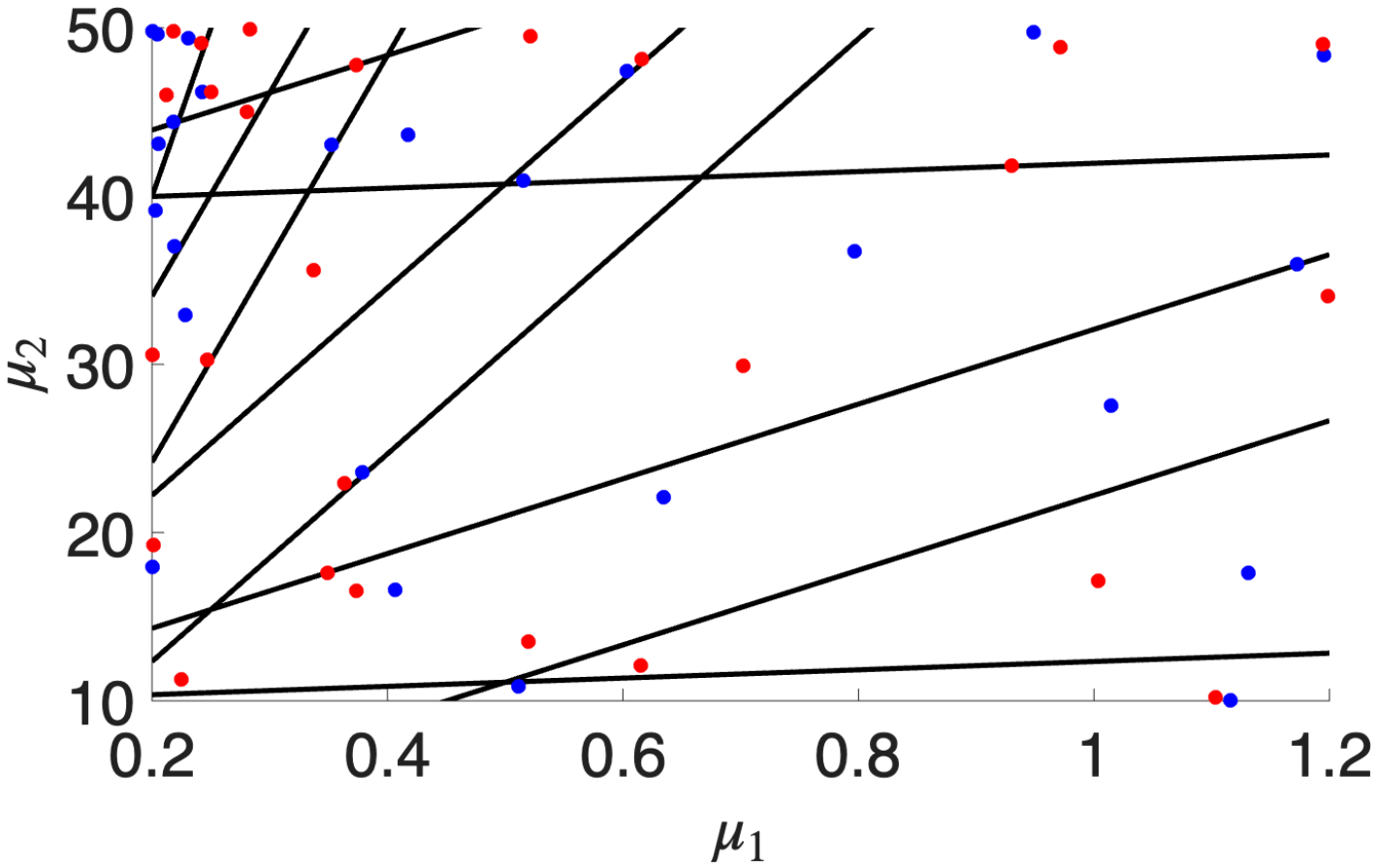


Figure: blue: randomized Greedy; red: deterministic Greedy.

Evolution of the Christoffel function

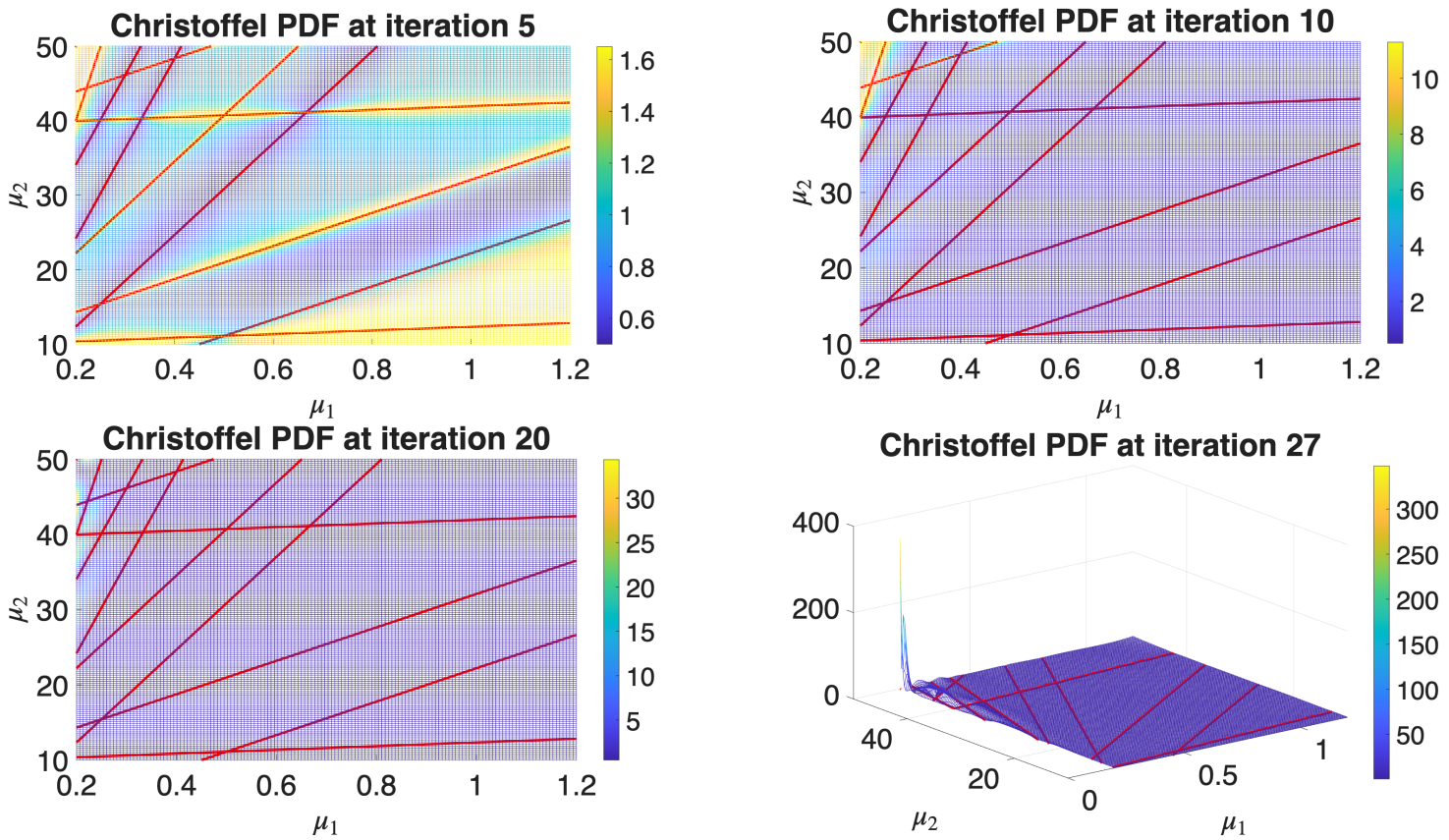


Figure: Christoffel function in iteration $5, 10, 20, 27$ (top-bottom; left-right) estimated on a uniform grid from $22,500$ samples