A Randomized Greedy Algorithm with Certification over the Entire Parameter Set

Charles Beall
Joint work with Kathrin Smetana

Stevens Institute of Technology

Reduced Order & Surrogate Modeling for Digital Twins Workshop
IMSI, 14 November 2025

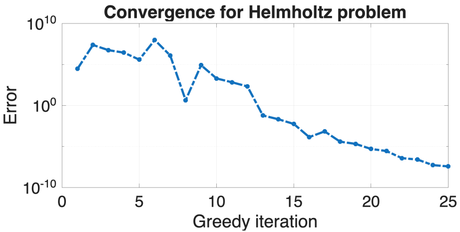
Supported by NSF Award # 2145364

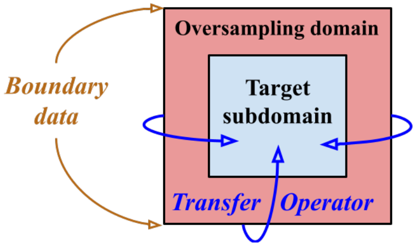
Outline

Main ideas

Contributions:

- A randomized Greedy algorithm with **offline certification at high probability.**
- Applicability to parametrized PDE problems with **high-dimensional parameter sets** – e.g. boundary data in localized model reduction, inverse problems, UQ, . . .
- Capabilities to **localize the construction of local reduced ansatz spaces**: does not rely on simulations of global computational domain.





The hope for certification

- **Goal:** Approximate set of PDE solutions

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

and certify this approximation for (almost) all parameters.

The hope for certification

- **Goal:** Approximate set of PDE solutions

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

and certify this approximation for (almost) all parameters.

- **Relevance:**
  - **Bayesian Inverse Problems & UQ:** Reduced models must be evaluated for many parameters to estimate statistics via Monte Carlo integration: certification can ensure accurate approximation of parameter distributions/quantified uncertainties.

The hope for certification

- **Goal:** Approximate set of PDE solutions

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

and certify this approximation for (almost) all parameters.

- **Relevance:**
  - **Bayesian Inverse Problems & UQ:** Reduced models must be evaluated for many parameters to estimate statistics via Monte Carlo integration: certification can ensure accurate approximation of parameter distributions/quantified uncertainties.
  - **Real-time contexts:** Reduced models in action/use (by e.g. autonomous systems, doctors in an operating room) must provide efficient results accurately for parameters outside of training set: certification can help ensure trust in the real-time outputs.

The hope for certification and Digital Twins

- **Goal:** Approximate set of PDE solutions

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

and certify this approximation for (almost) all parameters.

- **Relevance for Digital Twins:** High-dimensionality of the parameter set!
  - "For example, a surrogate model of the structural health of an engineering structure (e.g., building, bridge, airplane wing) would need to be representative over many thousands of material and structural properties that capture variation over space and time ...
  - ... Similarly, a surrogate model of tumor evolution in a cancer patient digital twin would potentially have thousands of parameters representing patient anatomy, physiology, and mechanical properties." [NASEM Digital Twins report, '24]

The challenges of certification

- **Proper Orthogonal Decomposition (POD)**[1]

$$\mathbf{S} = [\mathbf{u}(\mu^1) \cdots \mathbf{u}(\mu^{n_{\mathtt{samp}}})] = \mathbf{U\Sigma V}^\top$$

[1] [Sirovich Q. Appl. Math. '87], [Berkooz, Holmes, Lumley Annu. Rev. Fluid Mech. '93], ...

The challenges of certification

- **Proper Orthogonal Decomposition (POD)**[1]

$$\mathbf{S} = [\mathbf{u}(\mu^1) \cdots \mathbf{u}(\mu^{n_{\texttt{samp}}})] = \mathbf{U\Sigma V}^\top \implies \operatorname{span}\{\mathbf{U}_r\} \approx \mathcal{M}?$$

[1] [Sirovich Q. Appl. Math. '87], [Berkooz, Holmes, Lumley Annu. Rev. Fluid Mech. '93], . . .

The challenges of certification

- **Proper Orthogonal Decomposition (POD)**[1]

$$\mathbf{S} = [\mathbf{u}(\mu^1)\cdots\mathbf{u}(\mu^{n_{\text{samp}}})] = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top \implies \text{span}\{\mathbf{U}_r\} \approx \mathcal{M}?$$

- **Pro:**

$$\sum_{i=1}^{n_{\text{samp}}} \|\mathbf{u}(\mu^i) - \mathbf{U}_r\mathbf{U}_r^\top\mathbf{u}(\mu^i)\|_2 = \min_{\mathbf{C}_r} \sum_{i=1}^{n_{\text{samp}}} \|\mathbf{u}(\mu^i) - \mathbf{C}_r\mathbf{C}_r^\top\mathbf{u}(\mu^i)\|_2 = \sum_{k=r+1}^{n_{samp}} \sigma_k^2$$

[1] [Sirovich Q. Appl. Math. '87], [Berkooz, Holmes, Lumley Annu. Rev. Fluid Mech. '93], ...

The challenges of certification

- **Proper Orthogonal Decomposition (POD)**[1]

$$\mathbf{S} = [\mathbf{u}(\mu^1) \cdots \mathbf{u}(\mu^{n_{\text{samp}}})] = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top \implies \text{span}\{\mathbf{U}_r\} \approx \mathcal{M}?$$

- **Pro:**

$$\sum_{i=1}^{n_{\text{samp}}} \|\mathbf{u}(\mu^i) - \mathbf{U}_r\mathbf{U}_r^\top \mathbf{u}(\mu^i)\|_2 = \min_{\mathbf{C}_r} \sum_{i=1}^{n_{\text{samp}}} \|\mathbf{u}(\mu^i) - \mathbf{C}_r\mathbf{C}_r^\top \mathbf{u}(\mu^i)\|_2 = \sum_{k=r+1}^{n_{samp}} \sigma_k^2$$

- **Difficulty:** This ℓ^2-optimality provides average-case error analysis, but we are interested in certifying for worst-case (i.e., L^∞) scenarios too!

[1] [Sirovich Q. Appl. Math. '87], [Berkooz, Holmes, Lumley Annu. Rev. Fluid Mech. '93], ...

The challenges of certification

- **(Deterministic) Greedy algorithm**[1]
- Select a training set $\mathcal{P}_{\texttt{train}} \subset \mathcal{P}$ and error tolerance $\varepsilon_{\texttt{tol}}$;
  for $n = 0, 1, 2, \ldots,$ **do**
  - For each $\mu \in \mathcal{P}_{\texttt{train}}$, assess error† $\mathcal{E}^{(n)}(\mu)$ between $\mathbf{u}(\mu)$ and approximation in current reduced approximation space $\mathcal{X}_n$.

[1] [Veroy et al., *16th AIAA CFD Proceedings* '03]

The challenges of certification

- **(Deterministic) Greedy algorithm**[1]
- Select a training set $\mathcal{P}_{\texttt{train}} \subset \mathcal{P}$ and error tolerance $\varepsilon_{\texttt{tol}}$;
  for $n = 0, 1, 2, \ldots,$ **do**
  - For each $\mu \in \mathcal{P}_{\texttt{train}}$, assess error† $\mathcal{E}^{(n)}(\mu)$ between $\mathbf{u}(\mu)$ and approximation in current reduced approximation space $\mathcal{X}_n$.
  - **if** $\max_{\mu \in \mathcal{P}_{\texttt{train}}} \mathcal{E}^{(n)}(\mu) \leq \varepsilon_{\texttt{tol}}$, **break, return** $\mathcal{X}_n$.
  - **else** select $\mu^* \in \arg\max_{\mu \in \mathcal{P}_{\texttt{train}}} \mathcal{E}^{(n)}(\mu)$, $\mathcal{X}_{n+1} \leftarrow \mathcal{X}_n \oplus \mathrm{span}\{\mathbf{u}(\mu^*)\}$.

[1] [Veroy et al., *16th AIAA CFD Proceedings* '03]

The challenges of certification

- **(Deterministic) Greedy algorithm**[1]
- Select a training set $\mathcal{P}_{\texttt{train}} \subset \mathcal{P}$ and error tolerance $\varepsilon_{\texttt{tol}}$;
  for $n = 0, 1, 2, \ldots,$ **do**
  - For each $\mu \in \mathcal{P}_{\texttt{train}}$, assess error[†] $\mathcal{E}^{(n)}(\mu)$ between $\mathbf{u}(\mu)$ and approximation in current reduced approximation space $\mathcal{X}_n$.
  - **if** $\max_{\mu \in \mathcal{P}_{\texttt{train}}} \mathcal{E}^{(n)}(\mu) \leq \varepsilon_{\texttt{tol}}$, **break, return** $\mathcal{X}_n$.
  - **else** select $\mu^* \in \arg\max_{\mu \in \mathcal{P}_{\texttt{train}}} \mathcal{E}^{(n)}(\mu)$, $\mathcal{X}_{n+1} \leftarrow \mathcal{X}_n \oplus \mathrm{span}\{\mathbf{u}(\mu^*)\}$.

Certification: Reduced solutions $\mathbf{u}_N(\mu)$ approximate $\mathbf{u}(\mu)$ within $\varepsilon_{\texttt{tol}}$ for every $\mu \in \mathcal{P}_{\texttt{train}}$.

[1] [Veroy et al., *16th AIAA CFD Proceedings* '03]

The challenges of certification

- **(Deterministic) Greedy algorithm**[1]
- Select a training set $\mathcal{P}_{\texttt{train}} \subset \mathcal{P}$ and error tolerance $\varepsilon_{\texttt{tol}}$;
  for $n = 0, 1, 2, \ldots,$ **do**
  - For each $\mu \in \mathcal{P}_{\texttt{train}}$, assess error† $\mathcal{E}^{(n)}(\mu)$ between $\mathbf{u}(\mu)$ and approximation in current reduced approximation space $\mathcal{X}_n$.
  - **if** $\max_{\mu \in \mathcal{P}_{\texttt{train}}} \mathcal{E}^{(n)}(\mu) \le \varepsilon_{\texttt{tol}}$, **break, return** $\mathcal{X}_n$.
  - **else** select $\mu^* \in \arg\max_{\mu \in \mathcal{P}_{\texttt{train}}} \mathcal{E}^{(n)}(\mu)$, $\mathcal{X}_{n+1} \leftarrow \mathcal{X}_n \oplus \mathrm{span}\{\mathbf{u}(\mu^*)\}$.

Certification: Reduced solutions $\mathbf{u}_N(\mu)$ approximate $\mathbf{u}(\mu)$ within $\varepsilon_{\texttt{tol}}$ for every $\mu \in \mathcal{P}_{\texttt{train}}$.

- †Possible measures of error
  1. Best approximation error: $\mathcal{E}^{(n)}(\mu) := \inf_{v \in \mathcal{X}_n} \|\mathbf{u}(\mu) - v\|$.
  2. Galerkin approximation error: $\mathcal{E}^{(n)}(\mu) := \|\mathbf{u}(\mu) - \mathbf{u}_n(\mu)\|$.
  3. Error estimator or indicator: $\Delta^{(n)}(\mu)$.

[1] [Veroy et al., *16th AIAA CFD Proceedings* '03]

Is deterministic Greedy certification sufficient?

- **Idea:** If $\mathcal{E}^{(n)}$ can be evaluated cheaply, $\mu^* \in \arg\max_{\mu \in \mathcal{P}_{\mathtt{train}}} \mathcal{E}^{(n)}(\mu)$ may still be feasible to solve for large $\mathcal{P}_{\mathtt{train}}$. We could try creating a very fine training set, i.e., seek a $\varepsilon_{\mathtt{tol}}$-net for $\mathcal{P}$, so that $\mathcal{P}_{\mathtt{train}} \approx \mathcal{P}$.

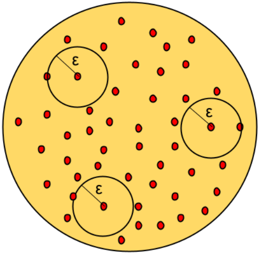


Figure: ε-net visualization

Is deterministic Greedy certification sufficient?

- **Idea:** If $\mathcal{E}^{(n)}$ can be evaluated cheaply, $\mu^* \in \arg\max_{\mu \in \mathcal{P}_{\texttt{train}}} \mathcal{E}^{(n)}(\mu)$ may still be feasible to solve for large $\mathcal{P}_{\texttt{train}}$. We could try creating a very fine training set, i.e., seek a $\varepsilon_{\texttt{tol}}$-net for $\mathcal{P}$, so that $\mathcal{P}_{\texttt{train}} \approx \mathcal{P}$.
- **Possible approach:** If $\mathbf{u}(\mu)$ is $L_{\mathcal{M}}$-Lipschitz with respect to the parametrs, then for $\mu \in \mathcal{P} \setminus \mathcal{P}_{\texttt{train}}$, $\tilde{\mu} \in \mathcal{P}_{\texttt{train}}$

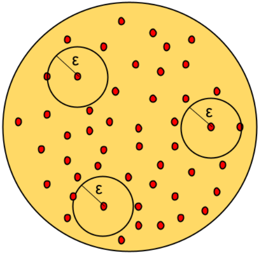


Figure: ε-net visualization

Is deterministic Greedy certification sufficient?

- **Idea:** If $\mathcal{E}^{(n)}$ can be evaluated cheaply, $\mu^* \in \arg\max_{\mu \in \mathcal{P}_{\texttt{train}}} \mathcal{E}^{(n)}(\mu)$ may still be feasible to solve for large $\mathcal{P}_{\texttt{train}}$. We could try creating a very fine training set, i.e., seek a $\varepsilon_{\texttt{tol}}$-net for $\mathcal{P}$, so that $\mathcal{P}_{\texttt{train}} \approx \mathcal{P}$.
- **Possible approach:** If $\mathbf{u}(\mu)$ is $L_{\mathcal{M}}$-Lipschitz with respect to the parametrs, then for $\mu \in \mathcal{P} \setminus \mathcal{P}_{\texttt{train}}$, $\tilde{\mu} \in \mathcal{P}_{\texttt{train}}$

$$\begin{aligned}\|\mathbf{u}(\mu) - \mathbf{u}_N(\tilde{\mu})\| &\leq \|\mathbf{u}(\mu) - \mathbf{u}(\tilde{\mu})\| + \|\mathbf{u}(\tilde{\mu}) - \mathbf{u}_N(\tilde{\mu})\| \\ &\leq L_{\mathcal{M}}|\mu - \tilde{\mu}| + \varepsilon_{\texttt{tol}}.\end{aligned}$$

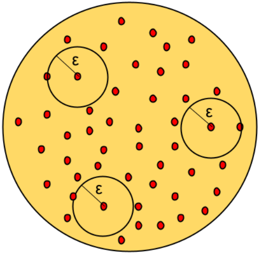


Figure: ε-net visualization

Is deterministic Greedy certification sufficient?

- **Idea:** If $\mathcal{E}^{(n)}$ can be evaluated cheaply, $\mu^* \in \arg\max_{\mu \in \mathcal{P}_{\texttt{train}}} \mathcal{E}^{(n)}(\mu)$ may still be feasible to solve for large $\mathcal{P}_{\texttt{train}}$. We could try creating a very fine training set, i.e., seek a $\varepsilon_{\texttt{tol}}$-net for $\mathcal{P}$, so that $\mathcal{P}_{\texttt{train}} \approx \mathcal{P}$.
- **Possible approach:** If $\mathbf{u}(\mu)$ is $L_{\mathcal{M}}$-Lipschitz with respect to the parametrs, then for $\mu \in \mathcal{P} \setminus \mathcal{P}_{\texttt{train}}$, $\tilde{\mu} \in \mathcal{P}_{\texttt{train}}$

$$\begin{aligned}\|\mathbf{u}(\mu) - \mathbf{u}_N(\tilde{\mu})\| &\leq \|\mathbf{u}(\mu) - \mathbf{u}(\tilde{\mu})\| + \|\mathbf{u}(\tilde{\mu}) - \mathbf{u}_N(\tilde{\mu})\| \\ &\leq L_{\mathcal{M}}|\mu - \tilde{\mu}| + \varepsilon_{\texttt{tol}}.\end{aligned}$$

- **Key challenge:** The cardinality of $\mathcal{P}_{\texttt{train}}$ for $\varepsilon_{\texttt{tol}}$-net scales *exponentially* with $\dim(\mathcal{P})$:

$$|\mathcal{P}_{\texttt{train}}| \sim \left(\frac{L_{\mathcal{M}}}{\varepsilon_{\texttt{tol}}}\right)^{c \dim(\mathcal{P})}$$

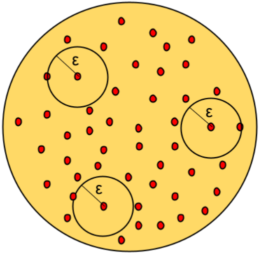


Figure: ε-net visualization

Is deterministic Greedy certification sufficient?

- **Idea:** If $\mathcal{E}^{(n)}$ can be evaluated cheaply, $\mu^* \in \arg\max_{\mu \in \mathcal{P}_{\texttt{train}}} \mathcal{E}^{(n)}(\mu)$ may still be feasible to solve for large $\mathcal{P}_{\texttt{train}}$. We could try creating a very fine training set, i.e., seek a $\varepsilon_{\texttt{tol}}$-net for $\mathcal{P}$, so that $\mathcal{P}_{\texttt{train}} \approx \mathcal{P}$.
- **Possible approach:** If $\mathbf{u}(\mu)$ is $L_{\mathcal{M}}$-Lipschitz with respect to the parametrs, then for $\mu \in \mathcal{P} \setminus \mathcal{P}_{\texttt{train}}$, $\tilde{\mu} \in \mathcal{P}_{\texttt{train}}$

$$\begin{aligned}\|\mathbf{u}(\mu) - \mathbf{u}_N(\tilde{\mu})\| &\leq \|\mathbf{u}(\mu) - \mathbf{u}(\tilde{\mu})\| + \|\mathbf{u}(\tilde{\mu}) - \mathbf{u}_N(\tilde{\mu})\| \\ &\leq L_{\mathcal{M}}|\mu - \tilde{\mu}| + \varepsilon_{\texttt{tol}}.\end{aligned}$$

- **Key challenge:** The cardinality of $\mathcal{P}_{\texttt{train}}$ for $\varepsilon_{\texttt{tol}}$-net scales *exponentially* with $\dim(\mathcal{P})$:

$$|\mathcal{P}_{\texttt{train}}| \sim \left(\frac{L_{\mathcal{M}}}{\varepsilon_{\texttt{tol}}}\right)^{c \dim(\mathcal{P})}$$

$\implies$ This approach suffers from the **curse of dimensionality!** [Bellman '57]

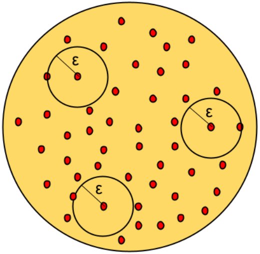


Figure: ε-net visualization

Existing works to address these challenges

- **Randomized Greedy for model order reduction** [Cohen et al. ESAIM: M2AN '20], [Billaud-Friess et al. Adv. Comput. Math. '24]
- **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 arXiv '24]
- **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, . . .

Outline

1. Introduction, main ideas & challenges
2. The algorithm
   - Design
   - Theory
3. Numerical results
   - Known nonlinear parameter-to-solution map
   - Multiparametric Helmholtz equation
4. Potential applications to Digital Twins

The randomized Greedy algorithm

- **Inputs:**
  - Parameter set $\mathcal{P}$
  - reference probability measure ρ supported on $\mathcal{P}$
  - per-iteration sampling budget $n_{\mathtt{samp}}$
  - measure of error $\mathcal{E}^{(n)} : \mathcal{P} \to \mathbb{R}_{\geq 0}$
  - error tolerance $\varepsilon_{\mathtt{tol}}$
  - Initialization: $n = 0$, $\mathcal{X}_0 := \emptyset$

The randomized Greedy algorithm

- **Inputs:**
  - Parameter set $\mathcal{P}$
  - reference probability measure ρ supported on $\mathcal{P}$
  - per-iteration sampling budget $n_{\mathtt{samp}}$
  - measure of error $\mathcal{E}^{(n)} : \mathcal{P} \to \mathbb{R}_{\geq 0}$
  - error tolerance $\varepsilon_{\mathtt{tol}}$
  - Initialization: $n = 0$, $\mathcal{X}_0 := \emptyset$
- **The algorithm: for** $n = 0, 1, 2, \ldots,$ **do**
  - Draw $n_{\mathtt{samp}}$ i.i.d. samples $\mu^1, \ldots, \mu^{n_{\mathtt{samp}}} \in \mathcal{P}$ according to the (modified) Christoffel measure ν, whose density is:

$$\frac{\mathrm{d}\nu}{\mathrm{d}\rho}(\mu) := K(\mu) := \frac{1}{2}\left(1 + \frac{|\mathcal{E}^{(n)}(\mu)|^2}{\int_{\mathcal{P}} |\mathcal{E}^{(n)}(\mu)|^2 \mathrm{d}\rho(\mu)}\right).$$

  - **if** $\max_{1 \leq i \leq n_{\mathtt{samp}}} \mathcal{E}^{(n)}(\mu^i) \leq \varepsilon_{\mathtt{tol}}$, then **break**, **return** $\mathcal{X}_n$.
  - **else** select μ_*^i that maximizes $\mathcal{E}^{(n)}$ over the sample set, update $\mathcal{X}_{n+1} \leftarrow \mathcal{X}_n \oplus \mathrm{span}\{\mathbf{u}(\mu_*^i)\}$.

What kind of distribution is K?

- **Question:** The randomized Greedy algorithm draws samples from a distribution with density K:

$$\frac{d\nu}{d\rho}(\mu) := K(\mu) := \frac{1}{2}\left(1 + \frac{|\mathcal{E}^{(n)}(\mu)|^2}{\int_{\mathcal{P}} |\mathcal{E}^{(n)}(\mu)|^2 d\rho(\mu)}\right).$$

What properties does this distribution have?

What kind of distribution is K?

- **Question:** The randomized Greedy algorithm draws samples from a distribution with density K:

$$\frac{\mathrm{d}\nu}{\mathrm{d}\rho}(\mu) := K(\mu) := \frac{1}{2}\left(1 + \frac{|\mathcal{E}^{(n)}(\mu)|^2}{\int_{\mathcal{P}} |\mathcal{E}^{(n)}(\mu)|^2 \mathrm{d}\rho(\mu)}\right).$$

  What properties does this distribution have?

- **Idea:** Assume that solutions $\mathbf{u}(\mu)$ are **stable** for ρ-almost every $\mu \in \mathcal{P} \implies \mathcal{E}^{(n)}$ is **bounded ρ-almost surely.**

What kind of distribution is K?

- **Question:** The randomized Greedy algorithm draws samples from a distribution with density K:

$$\frac{d\nu}{d\rho}(\mu) := K(\mu) := \frac{1}{2}\left(1 + \frac{|\mathcal{E}^{(n)}(\mu)|^2}{\int_{\mathcal{P}} |\mathcal{E}^{(n)}(\mu)|^2 d\rho(\mu)}\right).$$

  What properties does this distribution have?

- **Idea:** Assume that solutions $\mathbf{u}(\mu)$ are **stable** for ρ-almost every $\mu \in \mathcal{P} \implies \mathcal{E}^{(n)}$ is **bounded ρ-almost surely.**

Concentration for bounded random variables [version from Vershynin '18]

Let $X = (X^1, \ldots, X^m)$ be independent random variables such that $X^i \in [a_i, b_i]$, $1 \leq i \leq m$. Then, for any $t > 0$,

$$\mathbb{P}\left(\sum_{i=1}^{m} X_i - \mathbb{E}X_i \geq t\right) \leq \exp\left\{-\frac{2t^2}{\sum_{i=1}^{m}(b_i - a_i)^2}\right\}.$$

Sub-Gaussian distributions in high dimensions

- **Advantage:** The **sub-Gaussian** tail decay behavior of $\mathcal{E}^{(n)}$ and thus K ensures concentration around the mean, especially for high-dimensional $\mathcal{P}$ $\implies$ this is sometimes referred to as the "blessing of dimensionality."

Sub-Gaussian distributions in high dimensions

- **Advantage:** The **sub-Gaussian** tail decay behavior of $\mathcal{E}^{(n)}$ and thus K ensures concentration around the mean, especially for high-dimensional $\mathcal{P} \implies$ this is sometimes referred to as the "blessing of dimensionality."

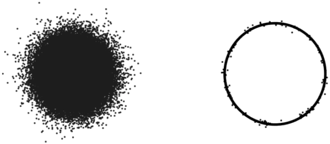


Figure: (Left) a Gaussian point cloud in two dimensions, and (right) its visualization in high dimensions [Vershynin '18].

Sub-Gaussian distributions in high dimensions

- **Advantage:** The **sub-Gaussian** tail decay behavior of $\mathcal{E}^{(n)}$ and thus K ensures concentration around the mean, especially for high-dimensional $\mathcal{P} \implies$ this is sometimes referred to as the "blessing of dimensionality."

Figure: (Left) a Gaussian point cloud in two dimensions, and (right) its visualization in high dimensions [Vershynin '18].

Implication: This favorable concentration of measure will ensure high-probability certification bounds are available!

Why the Christoffel sampling strategy?

- **Goal:** Obtain samples such that the following error bound holds with high probability

$$\sup_{\mu \in \mathcal{P}} \mathcal{E}(\mu) \leq C \max_{1 \leq i \leq n_{\mathtt{samp}}} \mathcal{E}(\mu^i).$$

Why the Christoffel sampling strategy?

- **Goal:** Obtain samples such that the following error bound holds with high probability

$$\sup_{\mu \in \mathcal{P}} \mathcal{E}(\mu) \leq C \max_{1 \leq i \leq n_{\mathtt{samp}}} \mathcal{E}(\mu^i).$$

- **Observation:** This is a problem of trying to bound a continuous norm (here L^∞) with a discrete norm (here ℓ^∞).

Why the Christoffel sampling strategy?

- **Goal:** Obtain samples such that the following error bound holds with high probability
$$\sup_{\mu \in \mathcal{P}} \mathcal{E}(\mu) \leq C \max_{1 \leq i \leq n_{\mathrm{samp}}} \mathcal{E}(\mu^i).$$
- **Observation:** This is a problem of trying to bound a continuous norm (here L^∞) with a discrete norm (here ℓ^∞).
- **Connection:** Sampling discretization seeks to develop strategies to effectively construct equivalent discrete norms, satisfying so-called Marcinkiewicz-Zygmund inequalities of the form
$$c_1 \|f(x)\|_{L^q}^q \leq \frac{1}{n_{\mathrm{samp}}} \sum_{k=1}^{n_{\mathrm{samp}}} |f(x^k)|^q \leq c_2 \|f(x)\|_{L^q}^q,$$
for specific q and functions f in a dictionary [e.g., Kashin et al. J. Complex. '22].

Why the Christoffel sampling strategy?

- **Goal:** Obtain samples such that the following error bound holds with high probability

$$\sup_{\mu\in\mathcal{P}} \mathcal{E}(\mu) \leq C \max_{1\leq i\leq n_{\mathtt{samp}}} \mathcal{E}(\mu^i).$$

Sampling complexity: The Christoffel measure yields a (quasi-)optimal sampling strategy, in the sense that the lower bound on $n_{\mathtt{samp}}$ is minimized [Cohen, Migliorati SMAI J. Comput. Math. '17].

- **Practical sampling algorithms (in a least-squares/L^2 context)**: [Bartel et al. Appl. Comput. Harmon. Anal. '23], [Dolbeault, Chkifa arXiv '24], [Trunschke, Nouy arXiv '24].
- **"Lifting" approximations from L^2 to other spaces:** [Xu, Narayan J. Approx. Theory '21], [Krieg et al. arXiv '23], many works by V. Temlyakov.

Why the Christoffel sampling strategy?

- **Goal:** Obtain samples such that the following error bound holds with high probability

$$\sup_{\mu \in \mathcal{P}} \mathcal{E}(\mu) \leq C \max_{1 \leq i \leq n_{\mathtt{samp}}} \mathcal{E}(\mu^i).$$

> **Sampling complexity:** The Christoffel measure yields a (quasi-)optimal sampling strategy, in the sense that the lower bound on $n_{\mathtt{samp}}$ is minimized [Cohen, Migliorati SMAI J. Comput. Math. '17].

- **Practical sampling algorithms (in a least-squares/L^2 context)**: [Bartel et al. Appl. Comput. Harmon. Anal. '23], [Dolbeault, Chkifa arXiv '24], [Trunschke, Nouy arXiv '24].
- **"Lifting" approximations from L^2 to other spaces:** [Xu, Narayan J. Approx. Theory '21], [Krieg et al. arXiv '23], many works by V. Temlyakov.

A new certification result

Offline certification with high probability

Let $\mathcal{M} := \{\mathbf{u}(\mu) : \mu \in \mathcal{P}\}$ consist of stable solutions for ρ-almost every parameter. Prescribe a first failure probability $\delta_{\mathtt{MZ}}$ and sampling budget $n_{\mathtt{samp}}$ satisfying

$$n_{\mathtt{samp}} \geq \frac{2}{\theta^2} \ln\left(\frac{2}{\delta_{\mathtt{MZ}}}\right), \quad \theta \in (0,1).$$

Moreover, because K is sub-Gaussian, prescribe a tail constant $C_{\mathtt{ub}}$, which yields a second failure probability $\delta_{\mathtt{ub}}$:

$$\delta_{\mathtt{ub}} = \mathbb{P}_\rho(K - \mathbb{E}_\rho K \geq C_{\mathtt{ub}}) \leq \exp\left(-\frac{2(C_{\mathtt{ub}})^2}{L^2 - (\mathbb{E}_\rho K)^2}\right), L = \frac{1}{\sqrt{\ln 2}} + \frac{\|\mathcal{E}^2\|_{\psi_2}}{\|\mathcal{E}\|^2_{L^2_\rho(\mathcal{P})}},$$

where ψ_2 denotes the sub-Gaussian norm[a]. Then, with probability at least $1 - \delta_{\mathtt{ub}} - \delta_{\mathtt{MZ}}$,

$$\sup_{\mu \in \mathcal{P}} \mathcal{E}(\mu) \leq \sqrt{\frac{8 C_{\mathtt{ub}}}{1-\theta}} \max_{1 \leq i \leq n_{\mathtt{samp}}} \mathcal{E}(\mu^i), \quad \mu^i \overset{\text{i.i.d.}}{\sim} \nu.$$

[a] $\|\mathcal{E}^2\|_{\psi_2} := \inf\{\beta > 0 : \mathbb{E}_\rho[\exp((\mathcal{E}^2/\beta)^2)] \leq 2\}$

Estimation of the norms

Key remark: Approximating the constant $C_{\mathtt{ub}}$, and normalizing $\mathcal{E}$, will involve estimating the sub-Gaussian and L^2_ρ-norms!

Estimation of the norms

Key remark: Approximating the constant $C_{\mathtt{ub}}$, and normalizing $\mathcal{E}$, will involve estimating the sub-Gaussian and L^2_ρ-norms!

- **Bounded random variables:** $\|X\|_{\psi_2} \leq \|X\|_{L^\infty_\rho(\mathcal{P})}/\sqrt{\ln 2}$
- ***L*-Lipschitz functions of Gaussians:** $\|f\|_{\psi_2} \sim L$

Estimation of the norms

Key remark: Approximating the constant C_{ub}, and normalizing $\mathcal{E}$, will involve estimating the sub-Gaussian and L^2_ρ-norms!

MC Estimator of the L^2_ρ-norm [Smetana, Taddei, Whitby, Yin]

Let $\mu^1, \dots, \mu^M \in \mathcal{P}$ be mutually independent samples drawn from ρ. The *a posteriori* error estimator

$$\Delta := \frac{1}{M} \sum_{i=1}^{M} \mathcal{E}^2(\mu^i)$$

satisfies the MZ inequalities $(1-\varepsilon)\|\mathcal{E}\|^2_{L^2_\rho(\mathcal{P})} \leq \Delta \leq (1+\varepsilon)\|\mathcal{E}\|^2_{L^2_\rho(\mathcal{P})}$, with probability at least $1-\delta$ for M satisfying

$$M \geq \frac{1}{2\varepsilon^2} \ln\left(\frac{1}{\delta}\right) \left(\frac{\|\mathcal{E}\|_{L^\infty_\rho(\mathcal{P})}}{\mathbb{E}_\rho[\mathcal{E}(\mu)]}\right)^2 .$$

Outline

Initial numerical test case: steady-state diffusion problem

Given $\mu = (\mu_1, \mu_2) \in [0.1, 10]^2$, $h(x, y) = \sin(\pi y)$, find $\mathbf{u}(\mu) \in H_0^1(\Omega)$ such that

$$\int_{\Omega_1} \mu_1 \nabla \mathbf{u}(\mu) \cdot \nabla v d\mathbf{x} + \int_{\Omega_2} \mu_2 \nabla \mathbf{u}(\mu) \cdot \nabla v d\mathbf{x} = \int_{\Omega} h v d\mathbf{x}, \forall v \in H_0^1(\Omega)$$

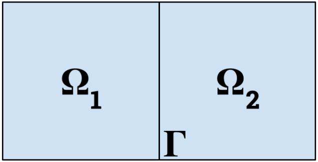


Figure: Domain of the diffusion problem. $\Omega_1 = (-1, 0) \times (0, 1)$, $\Omega_2 = (0, 1)^2$, $\Gamma = \{x = 0\} \times (0, 1)$.

Initial numerical test case: steady-state diffusion problem

Given $\mu = (\mu_1, \mu_2) \in [0.1, 10]^2$, $h(x, y) = \sin(\pi y)$, find $\mathbf{u}(\mu) \in H_0^1(\Omega)$ such that

$$\int_{\Omega_1} \mu_1 \nabla \mathbf{u}(\mu) \cdot \nabla v d\mathbf{x} + \int_{\Omega_2} \mu_2 \nabla \mathbf{u}(\mu) \cdot \nabla v d\mathbf{x} = \int_{\Omega} h v d\mathbf{x}, \forall v \in H_0^1(\Omega)$$

Exact computation of normalization constant

In [Autio, Hannukainen arXiv '24], a closed-form parameter-to-solution map is derived for this problem:

$$\mathbf{u}(\mu) = \frac{1}{\mu_1}\mathbf{w}_{\Omega_1} + \frac{1}{\mu_2}\mathbf{w}_{\Omega_2} + \frac{2}{\mu_1 + \mu_2}\mathbf{w}_{\Gamma},$$

where the **w** functions can be approximated via FE solutions. Knowledge of this map allows us to calculate the best approximation error, the normalization constant of the density of ν, etc., "exactly."

Convergence analysis across realizations

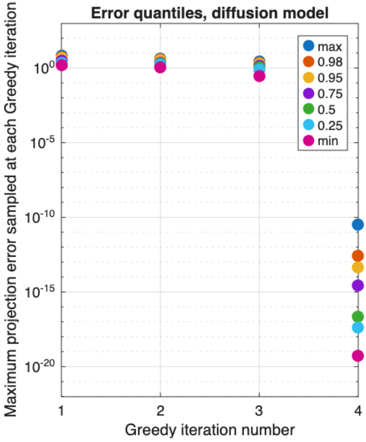
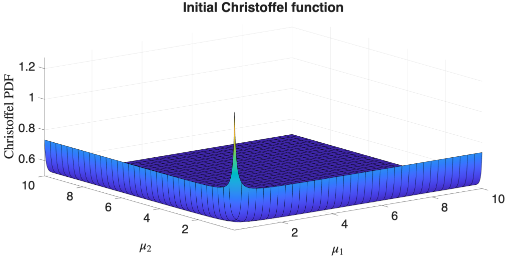
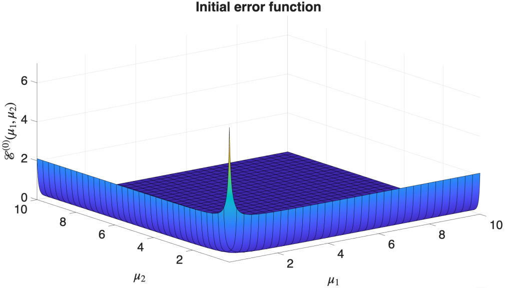


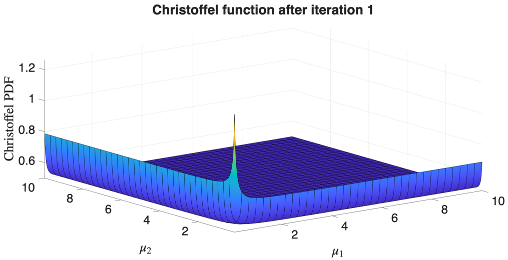
Figure: Error quantiles across 1,000 realizations of the randomized Greedy, $n_{\texttt{samp}} = 40$, $\varepsilon_{\texttt{tol}} = \texttt{1e-04}$.

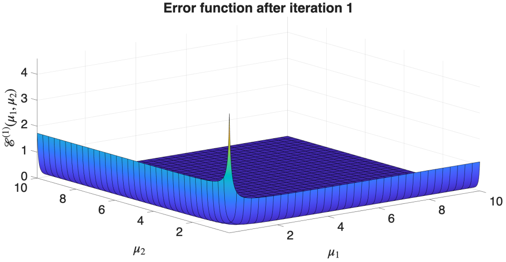
Evolution of the Christoffel pdfs and error measure



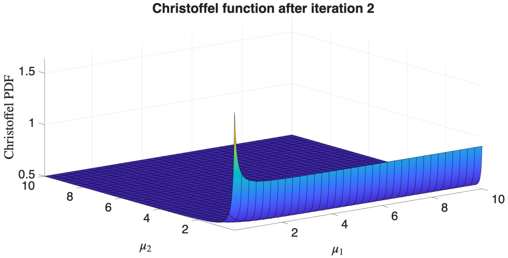


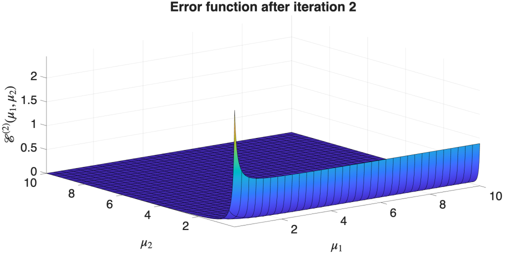
Evolution of the Christoffel pdfs and error measure



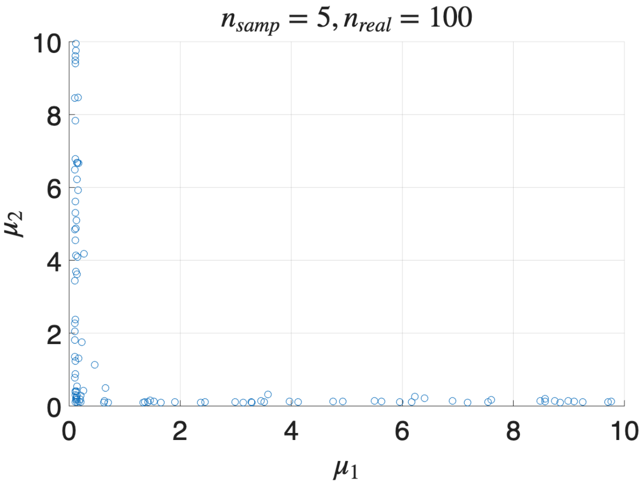


Evolution of the Christoffel pdfs and error measure

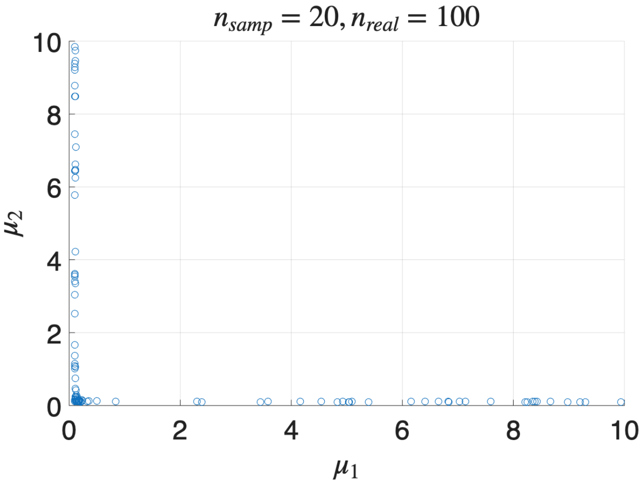




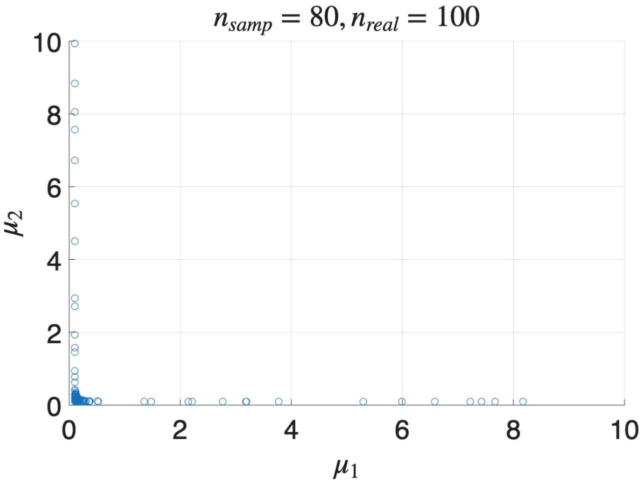
Varying the per-iteration budget



Varying the per-iteration budget



Varying the per-iteration budget



Second test case: Helmholtz problem

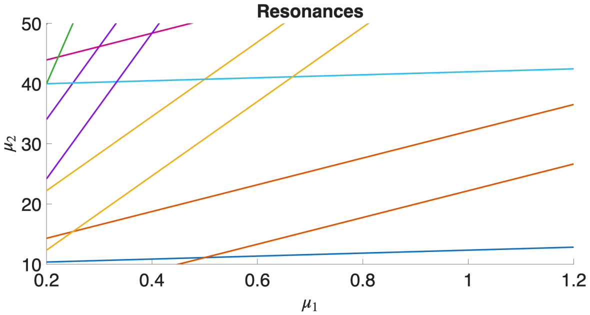
Given $\mu = (\mu_1, \mu_2) \in \mathcal{P} = [0.2, 1.2] \times [10, 50]$, find **u** such that

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

Second test case: Helmholtz problem

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

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



Parameters selected by the randomized Greedy

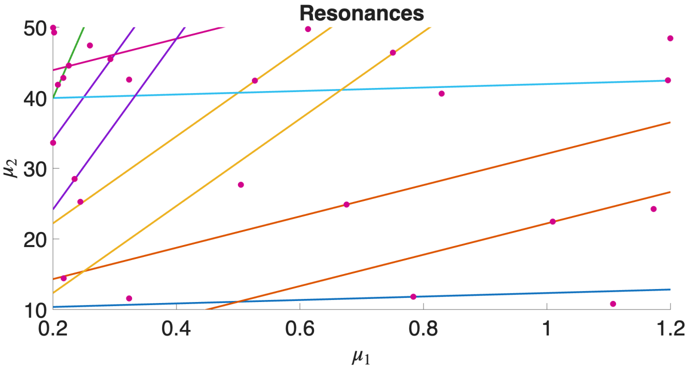


Figure: The parameters selected tend to be on or near resonance surfaces. Sampling here was done with a standard Metropolis-Hastings algorithm.

Convergence results (one realization)

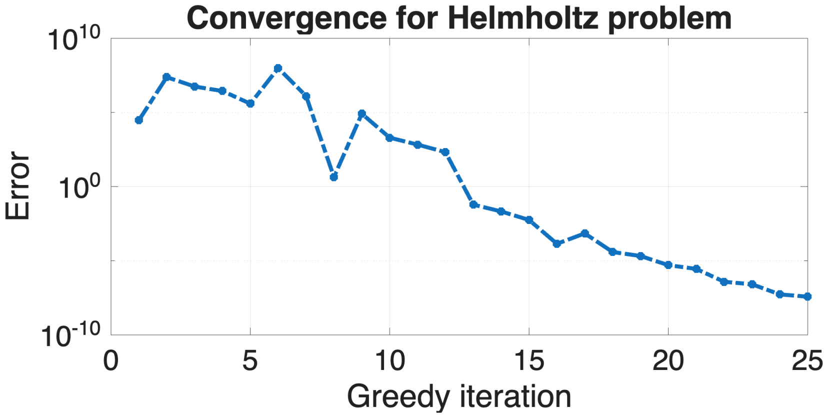


Figure: Tolerance here was set to $\varepsilon_{\mathtt{tol}} = \mathtt{1e-08}$.

Evolution of the Christoffel function and error

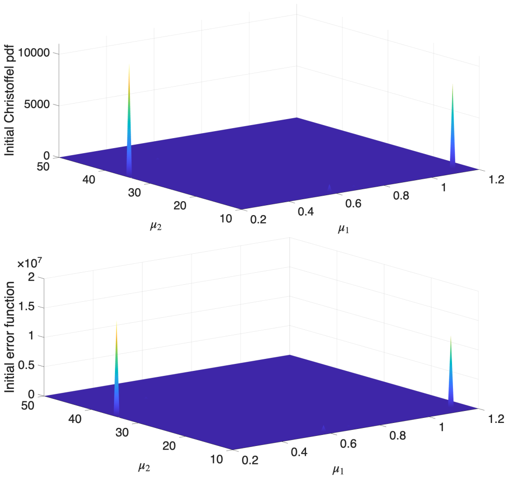


Figure: The error peaks initially are on the order of 10^7.

Evolution of the Christoffel function and error

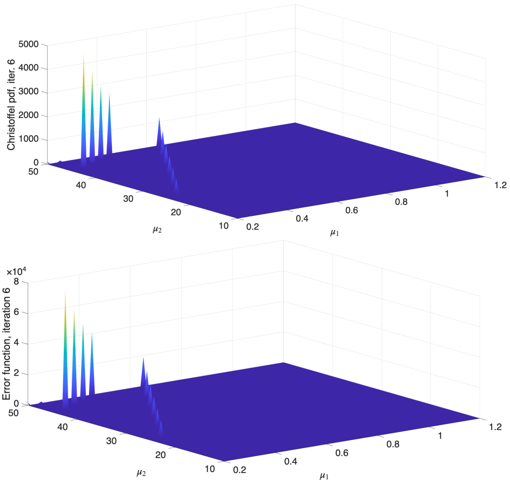


Figure: The error peaks at iteration 6 are on the order of 10^4.

Evolution of the Christoffel function and error

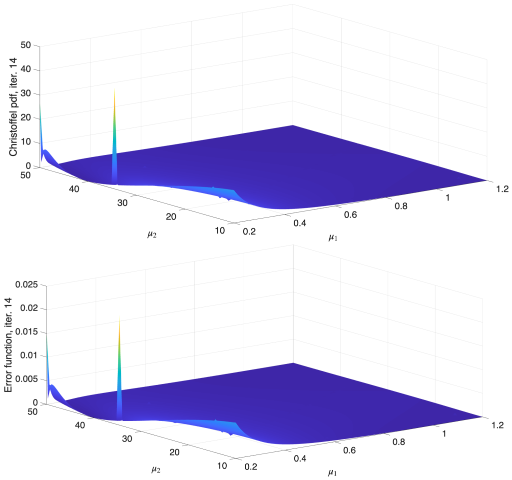


Figure: The error peaks at iteration 14 are on the order of 0.01.

Evolution of the Christoffel function and error

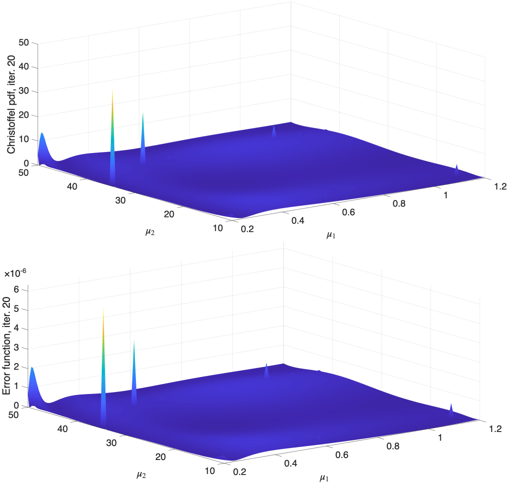


Figure: The error peaks at iteration 20 are on the order of 10^{-6}.

Outline

Digital Twins and localized model order reduction

- **Recall:** POD and deterministic Greedy have had many great successes, BUT will suffer from:
  - **curse of dimensionality** when the parameter set is high-dimensional;

Digital Twins and localized model order reduction

- **Recall:** POD and deterministic Greedy have had many great successes, BUT will suffer from:
  - **curse of dimensionality** when the parameter set is high-dimensional;
  - **inflexibility** when updating the (global) reduced model in response to local changes in PDE or domain geometry; and

Digital Twins and localized model order reduction

- **Recall:** POD and deterministic Greedy have had many great successes, BUT will suffer from:
  - **curse of dimensionality** when the parameter set is high-dimensional;
  - **inflexibility** when updating the (global) reduced model in response to local changes in PDE or domain geometry; and
  - questions about how to pose problem when **global PDE is inaccessible/unknown!**

Digital Twins and localized model order reduction

- **Recall:** POD and deterministic Greedy have had many great successes, BUT will suffer from:
  - **curse of dimensionality** when the parameter set is high-dimensional;
  - **inflexibility** when updating the (global) reduced model in response to local changes in PDE or domain geometry; and
  - questions about how to pose problem when **global PDE is inaccessible/unknown!**
- **Localized MOR** may effectively address these issues via, e.g.:
  1. Decompose the global domain into target subdomains ω^{in}, each with an associated oversampling domain $\omega^{out} \supset \omega^{in}$.
  2. Build reduced spaces on ω^{in} by solving local problems.
  3. Patch local spaces together to construct global reduced models.

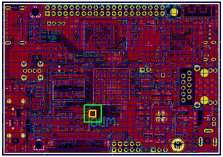
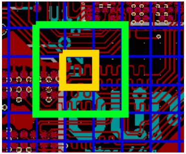
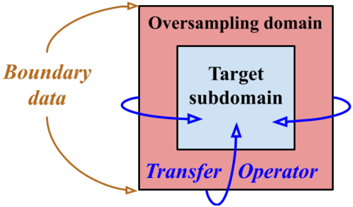



Image credit: A. Buhr

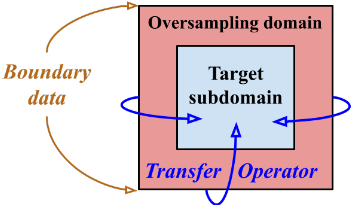
How optimal local approximation spaces are constructed

- **Idea:** Introduce a transfer operator $\mathcal{T}$ which maps arbitrary boundary data on $\partial\omega^{out}$ to a local solution on ω^{in}.



How optimal local approximation spaces are constructed

- **Idea:** Introduce a transfer operator $\mathcal{T}$ which maps arbitrary boundary data on $\partial\omega^{out}$ to a local solution on ω^{in}.
- **Key Observation:** The global solution $\mathbf{u}$ satisfies $\mathbf{u}|_{\omega^{in}} = \mathcal{T}\left(\mathbf{u}|_{\partial\omega^{out}}\right)$
$\implies$ Construct local reduced spaces that approximate range($\mathcal{T}$).

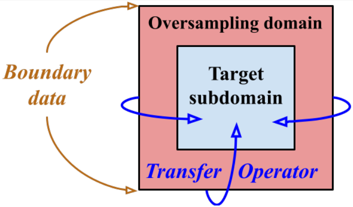


How optimal local approximation spaces are constructed

Optimal local approximation spaces [Babuska, Lipton SIAM MMS '11]

- ϕ_j: left singular vectors of $\mathcal{T}$; σ_j: singular values of $\mathcal{T}$.
- The reduced space $\mathcal{S}_{opt}^{(n)} := \text{span}\{\phi_1, \ldots, \phi_n\}$ is the optimal space in the sense of Kolmogorov [Kolmogoroff Annal Math. 1936].
- The error satisfies

$$\sup_{\mathbf{g}} \left\{ \frac{\|(\mathcal{T} - \mathrm{Proj}_{\mathcal{S}_{opt}^{(n)}} \mathcal{T})\mathbf{g}\|}{\|\mathbf{g}\|} \right\} = \sigma_{n+1}$$



Existing approaches for nonlinear problems

- Approaches utilizing randomization: [Chen, Li, Lu, Wright SIAM Multiscale Model. Simul. '22] (multiscale), [Smetana, Taddei SIAM J. Sci. Comput. '23] (local MOR)
- Localized Orthogonal Decomposition (LOD): [Verfürth IMA J. Numer. Anal. '21], [Khrais, Verfürth arXiv '25]
- Rough polyharmonic splines: [Kambampati '16 (Master's Thesis)], [Liu, Chung, Zhang SIAM Multiscale Model. Simul. '21]
- Generalization of Gamblets: [Chen, Hosseini, Owhadi, Stuart J. Comput. Phys. '21], . . .

Challenges & opportunities for nonlinear local MOR

- **Goal:** Approximate the set

$$\mathcal{M} := \{\mathcal{T}(\mathbf{g}) : \|\mathbf{g}\| \leq 1\}$$

where the transfer operator $\mathcal{T}$ is now nonlinear.

Challenges & opportunities for nonlinear local MOR

- **Goal:** Approximate the set

$$\mathcal{M} := \{\mathcal{T}(\mathbf{g}) : \|\mathbf{g}\| \leq 1\}$$

where the transfer operator $\mathcal{T}$ is now nonlinear.

- **Connection to Digital Twins:** In a DT setting, global high-fidelity solutions or data from the global computational domain may be unknown/inaccessible $\implies$ we would need to rely on and trust our local reduced models!

Challenges & opportunities for nonlinear local MOR

- **Goal:** Approximate the set

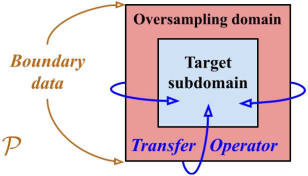
$$\mathcal{M} := \{\mathcal{T}(\mathbf{g}) : \|\mathbf{g}\| \leq 1\}$$

where the transfer operator $\mathcal{T}$ is now nonlinear.

- **Connection to Digital Twins:** In a DT setting, global high-fidelity solutions or data from the global computational domain may be unknown/inaccessible $\implies$ we would need to rely on and trust our local reduced models!

Opportunity for our approach: We can use the randomized Greedy to construct and certify local reduced spaces **in a local fashion**, while handling the high-dimensional parameter set of admissible local boundary data.

Randomized Greedy & localized model order reduction



- **Inputs:**
  - $\mathcal{P} := \{\text{boundary data } \mathbf{g}\}$
  - local solutions $\mathcal{T}(\mathbf{g})$ via transfer operator $\mathcal{T}$
  - reference probability measure ρ supported on $\mathcal{P}$
  - per-iteration sampling budget $n_{\mathtt{samp}}$
  - $\mathcal{E}^{(n)} : \mathcal{P} \to \mathbb{R}_{\geq 0}$, $\mathcal{E}^{(n)}(\mathbf{g}) := \|\mathcal{T}(\mathbf{g}) - \mathrm{Proj}_{\mathcal{X}_n}\mathcal{T}(\mathbf{g})\| / \|\mathbf{g}\|$
  - Initialize $\mathcal{X}_0 := \emptyset$, $n = 0$.
- **The algorithm: for** $n = 0, 1, 2, \ldots,$ **do**
  - Draw $n_{\mathtt{samp}}$ i.i.d. samples $\mu^1, \ldots, \mu^{n_{\mathtt{samp}}} \in \mathcal{P}$ according to the (modified) Christoffel measure ν, whose density is:

$$\frac{\mathrm{d}\nu}{\mathrm{d}\rho}(\mu) := \frac{1}{2}\left(1 + \frac{|\mathcal{E}^{(n)}(\mu)|^2}{\int_{\mathcal{P}} |\mathcal{E}^{(n)}(\mu)|^2 \mathrm{d}\rho(\mu)}\right).$$

  - **if** $\max_{1 \leq i \leq n_{\mathtt{samp}}} \mathcal{E}(\mu^i) \leq \varepsilon_{\mathtt{tol}}$, then **break**, **return** $\mathcal{X}_n$.
  - **else** select μ_*^i that maximizes $\mathcal{E}$ over the sample set, update $\mathcal{X}_{n+1} \leftarrow \mathcal{X}_n \oplus \mathrm{span}\{\mathbf{u}(\mu_*^i)\}$.

Summary & conclusion

- **Main question:** How to construct and certify reduced ansatz spaces, given high-dimensional parameter sets?
- **Idea:** Utilize randomization to explore the parameter set, relying on favorable concentration properties of sub-Gaussian distributions.
- **Approach:** Implement a randomized Greedy algorithm to sample parameters quasi-optimally via the Christoffel measure, and obtain offline certification with high probability.
- **Further advantages:** Flexibility in the algorithm design can allow for the use of different measures/indicators of error, localized approximation, and nonlinear PDEs.

Summary & conclusion

- **Main question:** How to construct and certify reduced ansatz spaces, given high-dimensional parameter sets?
- **Idea:** Utilize randomization to explore the parameter set, relying on favorable concentration properties of sub-Gaussian distributions.
- **Approach:** Implement a randomized Greedy algorithm to sample parameters quasi-optimally via the Christoffel measure, and obtain offline certification with high probability.
- **Further advantages:** Flexibility in the algorithm design can allow for the use of different measures/indicators of error, localized approximation, and nonlinear PDEs.

Thank you for your attention!

Questions?