An Adaptive Reduced Basis Approach for PDE Constrained Optimization under Uncertainty

Wilkins Aquino

Professor

Department of Mechanical Engineering and Materials Science,

Duke University, Durham, NC

Duke
UNIVERSITY

Acknowledgements

- Zilong Zou, PhD Student (Duke), now at GS
- Sponsor: DARPA EQUiPS Grant Number: SNL 014150709

  Team members:

  Drew Kouri (PI), Sandia National Laboratories

  Denis Ridzal, Sandia National Laboratories

  Stan Uruasev, University of Florida

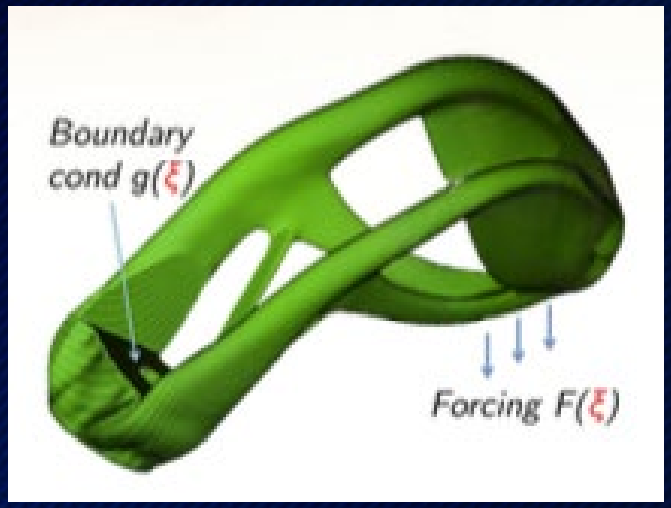
  R. Tyrrell Rockafellar, University of Washington

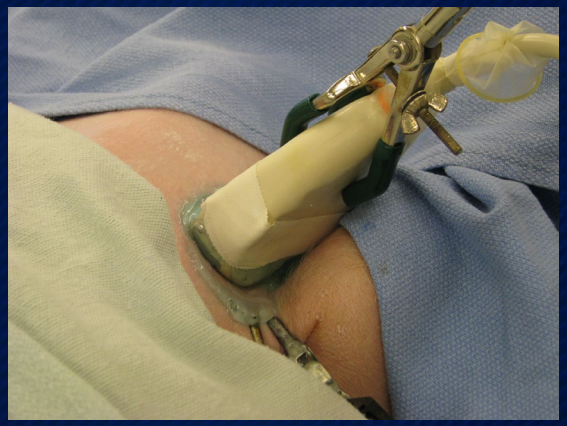
  Alexander Shapiro, Georgia Institute of Technology

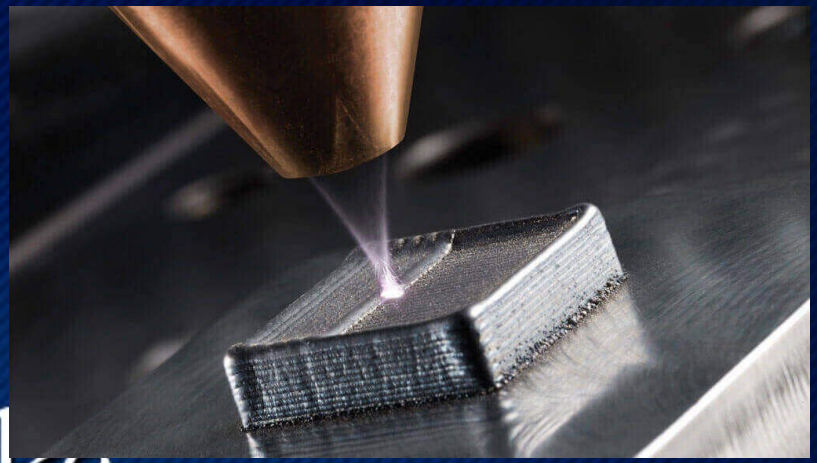
  Wilkins Aquino, Duke University

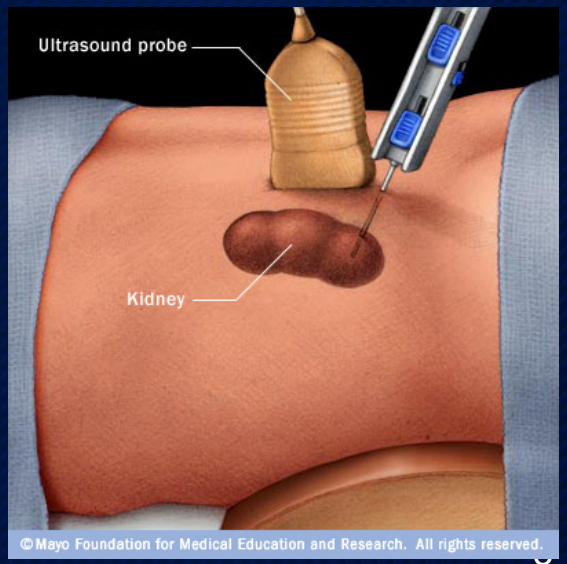
Duke UNIVERSITY

Motivation:
Design and Control under Uncertainty









Duke
UNIVERSITY

Stochastic Formulation of Governing Problem

- We will denote U as the state space and Z as the space of optimization or control variables.
- Let ξ be a random vector with known probability distribution.
- Consider the following parametrized, linear (PDE) problem: for fixed $z \in Z$, find $u : \Xi \to U$ such that

$$M(u, z; \xi) := \mathcal{L}(\xi)u(\xi) + \mathcal{B}(\xi)z + \ell(\xi) = 0 \quad \text{a.s.}$$

- Denote a solution to our (PDE) problem as $S(\xi; z)$ for given z such that

$$M(S(\xi; z), z; \xi) = 0 \quad \text{a.s.}$$

Duke UNIVERSITY

Stochastic Formulation of Governing Problem

- The differential operator satisfies the stability conditions: $\exists\, \kappa_1 > 0$ independent of $\xi \in \Xi$ such that for all $\xi \in \Xi$

$$\inf_{u \in U \setminus \{0\}} \sup_{v \in V \setminus \{0\}} \frac{|\langle \mathcal{L}(\xi) u, v \rangle_{V^*,V}|}{\|u\|_U \|v\|_V} =: \gamma(\xi) \geq \kappa_1, \quad \forall\, \xi \in \Xi$$

$$\langle \mathcal{L}(\xi)^* v, u \rangle_{U^*,U} = 0 \quad \forall\, u \in U \quad \Longrightarrow \quad v = 0$$

- In addition, there $\exists\, \kappa_2 > 0$ independent of $\xi \in \Xi$ such that for all $\xi \in \Xi$

$$|\langle \mathcal{L}(\xi) u, v \rangle_{V^*,V}| \leq \kappa_3 \|u\|_U \|v\|_V$$

Duke UNIVERSITY

Risk-Averse Optimization

- Let $\wp : Z \to \mathbb{R}$, $G : U \times \Xi \to \mathbb{R}$ and note that $G(S(\xi; z), \xi)$ is a random variable.
- We seek to minimize

$$J(z) = \mathcal{R}\big(G(S(\xi; z), \xi)\big) + \wp(z)$$

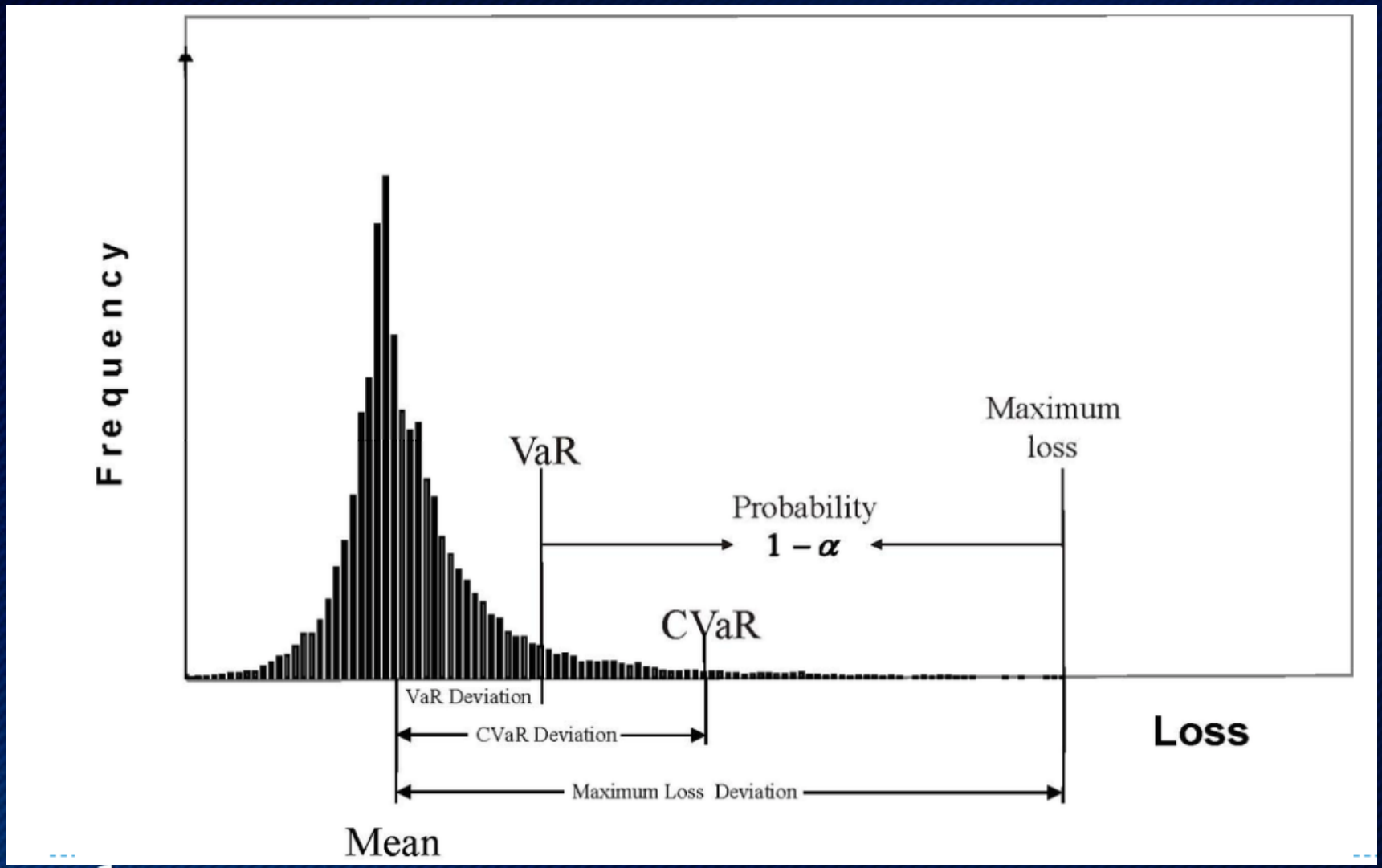
  where $\mathcal{R} : \mathcal{X} \to \mathbb{R}$ is a risk measure with the form

$$\mathcal{R}(X) = \inf_{t \in \mathbb{R}} \{t + \mathbb{E}[v(X - t)]\}.$$

- $v : \mathbb{R} \to \mathbb{R}$ is convex and satisfies $v(0) = 0$ and $v(x) > x$ for $x \neq 0$. For example, when $v(x) = (1 - \beta)^{-1}[x]^+$ with $\beta \in (0, 1)$ and $[\cdot]^+ = \max\{0, \cdot\}$, the risk measure $\mathcal{R}$ is the conditional value-at-risk (CVaR)

Duke UNIVERSITY

Conditional Value at Risk (CVaR)



Courtesy of Prof. Stan Uryasev, University of Florida.

Duke UNIVERSITY

Risk-Averse Optimization Formulation

- The expectation that appears in the $\mathcal{R}$ is approximated by sample average as

$$\mathbb{E}[v\big(G(S(\xi;z),\xi)-t\big)] \approx \frac{1}{N}\sum_{j=1}^{N} v\big(G(S(\xi;z),\xi)-t\big)$$

- We incorporate t from the definition of $\mathcal{R}$ as an optimization variable and solve

$$\min_{z\in Z, t\in\mathbb{R}} \ \widehat{J}(t,z) \quad \text{where} \quad \widehat{J}(t,z) := \left\{ t + \frac{1}{N}\sum_{j=1}^{N} v\big(G(S(\xi;z),\xi)-t\big) \right\}$$

Duke UNIVERSITY

Risk-Averse Optimization Formulation

- Under standard assumptions, the gradient of the objective functions $\widehat{J}$ with respect to the control z is computed as

$$\nabla_z \widehat{J}(t,z) = \frac{1}{N}\sum_{j=1}^{N} \partial v\big(G(S(\xi_j;z),\xi_j) - t\big)\mathcal{B}(\xi_j)^*\lambda_j + \nabla_{\!\wp}(z)$$

  where λ_j solves the adjoint equation

$$\mathcal{L}(\xi_j)^*\lambda_j = -\nabla_u G(S(\xi_j;z),\xi_j).$$

- Similarly, the derivative with respect to t is

$$\nabla_t \widehat{J}(t,z) = 1 - \frac{1}{N}\sum_{j=1}^{N} \partial v\big(G(S(\xi_j;z),\xi_j) - t\big).$$

Duke UNIVERSITY

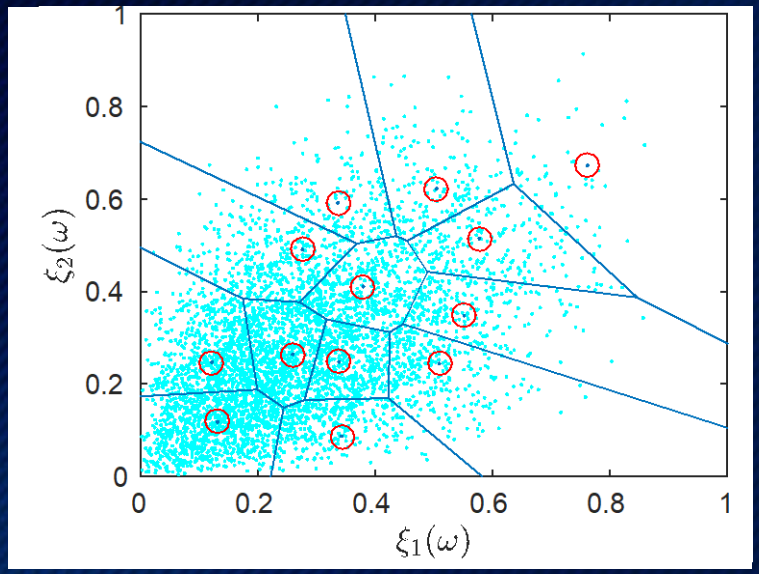
Computational Challenges

- Evaluation of the objective and its gradient requires N solutions of the state PDE and N solutions of the adjoint PDE.
- Our proposed approach is two-pronged:
  - Adaptive sample-based reduced basis approach for approximating the PDE solution (provably convergent).
  - Inexact trust region framework that allows for inexact evaluations of objective and gradient (provably convergent).

Duke UNIVERSITY

Sample-Based Approximations

- Take a finite set of atoms $\{\bar{\boldsymbol{\xi}}_k, k = 1, 2, \ldots, \bar{m}\}$ with probability $p_k \geq 0$ such that $\sum_{k=1}^{\bar{m}} p_k = 1$.
- These atoms and probabilities are chosen as to approximate the original random variable in some sense.
- $\{\bar{\boldsymbol{\xi}}_k, k = 1, 2, \ldots, \bar{m}\}$ generate a Voronoi-cell based partition of the sample space $\Xi = \cup_{k=1}^{m} \bar{\Xi}_k$ ($\bar{\Xi}_k$ is centered at $\bar{\boldsymbol{\xi}}_k$).



Duke UNIVERSITY

Local reduced basis approximation

- Given z, we construct approximations of the state as

$$\bar{S}(\xi, z) := \sum_{k=1}^{m} \mathbf{1}(\xi \in \Xi_k) S_k(\xi, z)$$

- $S_k(\xi, z)$ solves the following reduced local problem: find $S_k(\xi, z) \in U_k$ such that

$$\langle M(S_k(\xi, z), z; \xi), v \rangle_{V^*, V} = 0 \quad \forall\, v \in V_k, \quad \forall\, \xi \in \Xi_k.$$

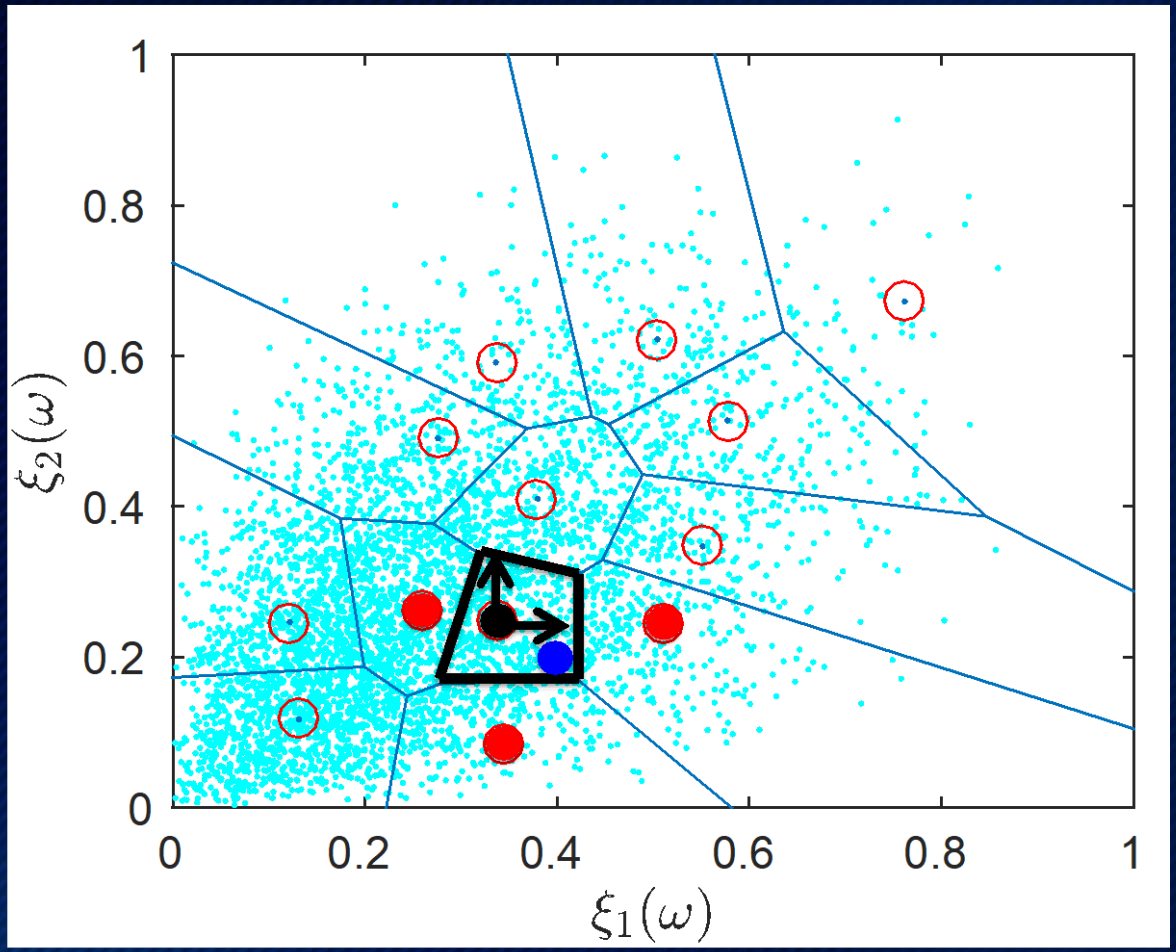
- $U_k = \text{span}\{\Phi_k\}$ is the reduced approximation space and

$$\Phi_k = [\nabla_\xi S(\xi_k; z), S(\xi_k; z), S(\xi_{k_1}; z), S(\xi_{k_2}; z), \ldots, S(\xi_{k_N}; z)]$$

where $S(\xi_{k_1}; z), S(\xi_{k_2}; z), \ldots, S(\xi_{k_N}; z)$ are solutions at the N atoms closest to ξ_k.

Duke UNIVERSITY

Sketch of Enriched Local Reduced Basis



Duke UNIVERSITY

Adaptive Selection of Atoms

- Multiple techniques for a posteriori error estimation and atom selection: residual based, QoI based on adjoints, etc.
- We used a residual-based error indicator.

$$\|e\|_U \leq \frac{\|r(\bar{u}, \cdot; \xi)\|_{V^*}}{\gamma(\xi)}$$

- Error decomposition thanks to the Voronoi partion,

$$\mathbb{E}\left[\|e(\xi)\|_U\right] = \sum_{k=1}^{m} \mathbb{E}\left[\|e(\xi)\|_U \mathbf{1}(\xi \in \Xi_k)\right] \leq \sum_{k=1}^{m} \mathbb{E}\left[\epsilon_u(\xi)\mathbf{1}(\xi \in \Xi_k)\right] = \sum_{k=1}^{m} \eta_k$$

where $\eta_k := \mathbb{E}\left[\epsilon_u(\xi)\mathbf{1}(\xi \in \Xi_k)\right]$ is the *local* error indicator. In a greedy method, the next atom is selected from the cell with the largest η_k.

Duke UNIVERSITY

1D Helmholtz Problem with Two Stochastic Dimensions

- Helmholtz Problem: Damped vibration

$$-\frac{d}{dx}\left(\nu(x,\omega)\frac{du}{dx}\right) - ic\tau u - \tau^2 u + f = 0, \quad (x,\omega) \in D \times \Omega \text{ a.s.}$$

$$u(0) = u(1) = 0, \quad \text{a.s.}$$

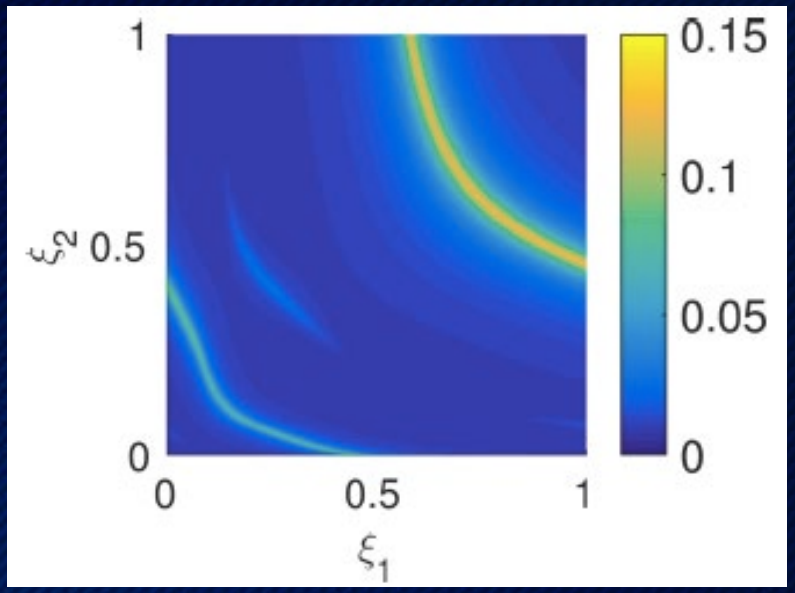
- Random Variables

$$\nu(x,\omega) = [1 + 3\xi_1(\omega)]\,\mathbf{1}(x \in [0, 0.5)) + [2.5 + 3\xi_2(\omega)]\,\mathbf{1}(x \in [0.5, 1]).$$

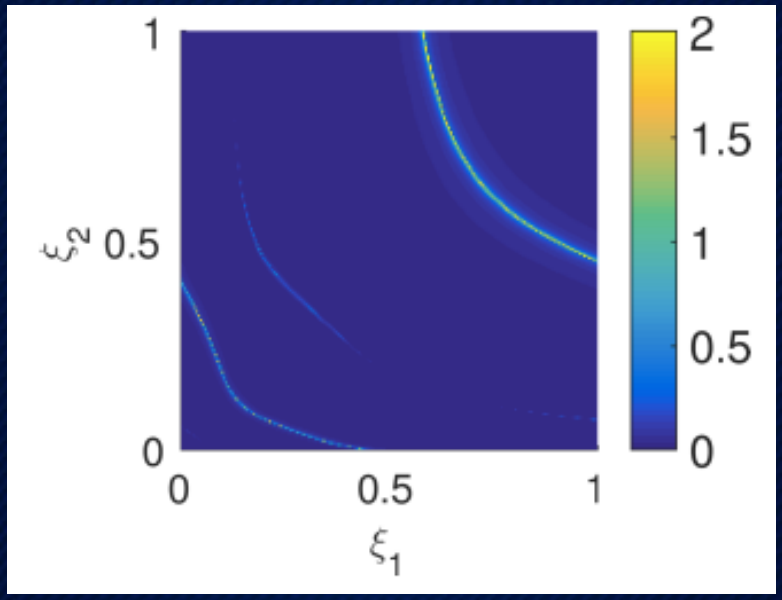
$\xi_1 \sim \text{Beta}(1,3)$ and $\xi_2 \sim \text{Beta}(3,2)$ with a correlation $\rho(\xi_1, \xi_2) = 0.5$.

Duke UNIVERSITY

1D Helmholtz Equation
Norm of the Solution in Parameter Space



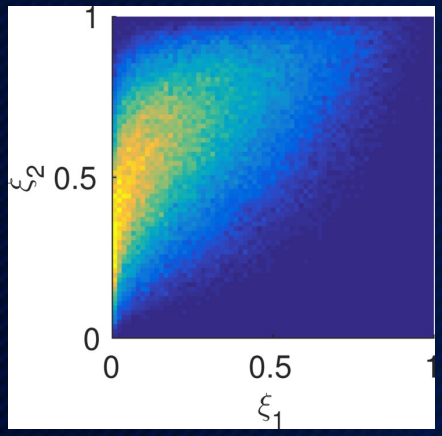
c= 0.2



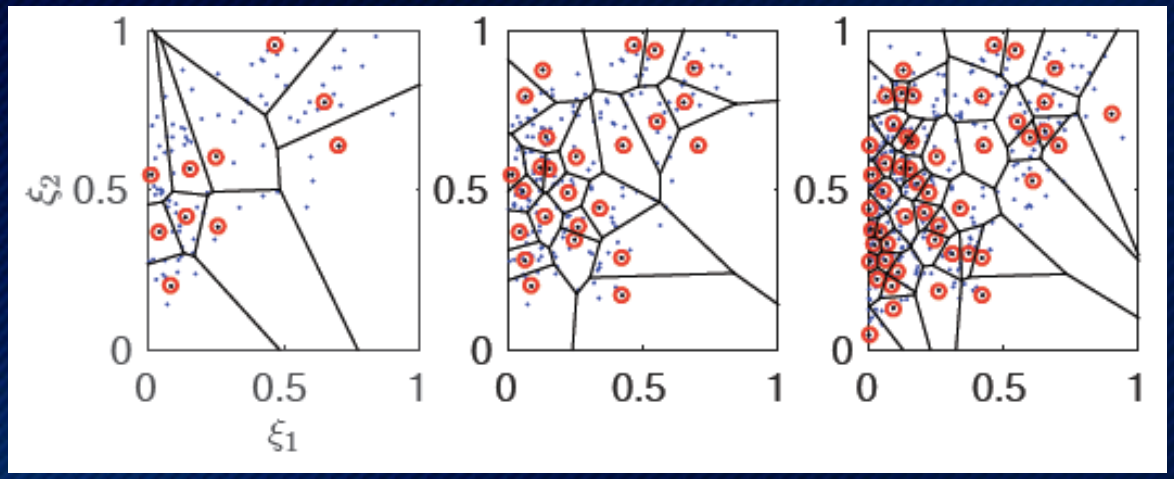
c= 0.01

Duke UNIVERSITY

1D Helmholtz Equation:partition generated



Joint parameter distribution

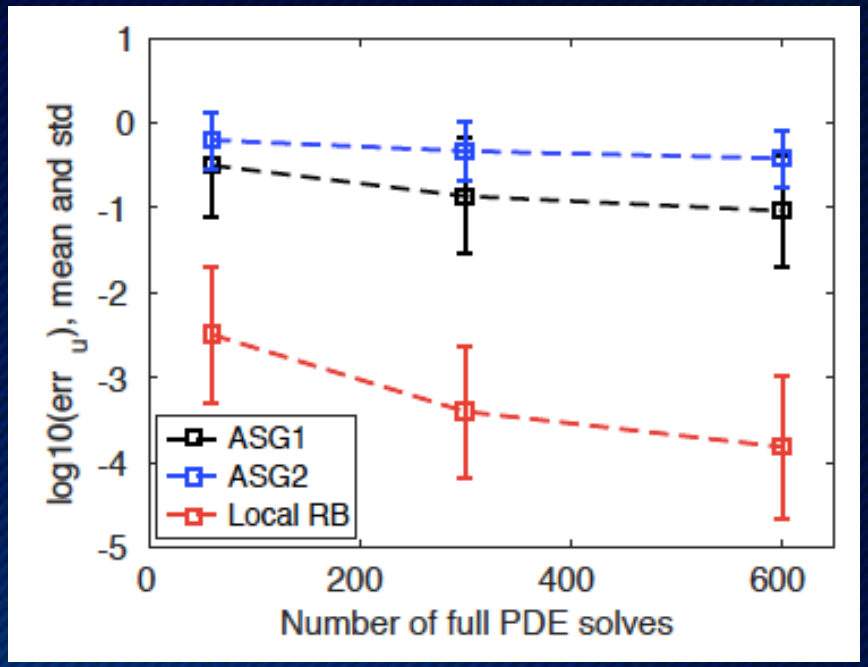


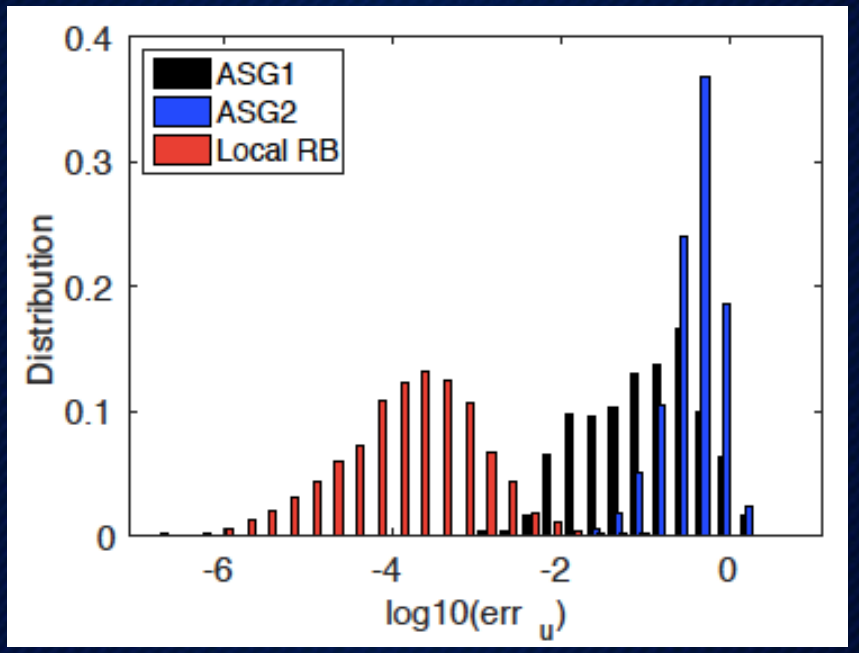
Partitions generated with 10, 25, and 50 atoms and c=0.2

Duke UNIVERSITY

1D Helmholtz Equation
Comparison with Sparse Grid Methods

$$e_u(\xi) = \frac{\|\bar{u}(\xi) - u(\xi)\|_{L_2(\mathcal{D})}}{\|u(\xi)\|_{L_2(D)}}$$





600 PDE solves

Duke UNIVERSITY