A Data-Driven Approach to PDE-Constrair SAND2015-11107C Uncertainty

Drew Kouri

Optimization and Uncertainty Quantification
Sandia National Laboratories, Albuquerque, New Mexico
dpkouri@sandia.gov

January 6, 2016

Joint Mathematics Meeting, Washington State Convention Center, Seattle, WA.

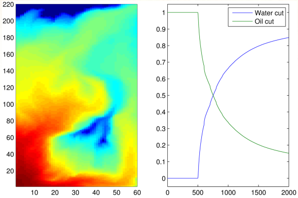
Funding support: ASC.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

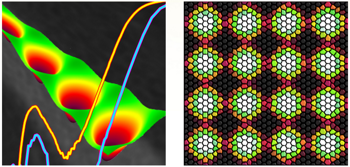
Motivation

Reservoir Optimization



$$v = -\mathbf{K}\lambda(s)\nabla p, \quad \nabla \cdot v = q$$
$$\phi\, \partial_t s + \nabla \cdot (\, f(s) v\,) = \widehat{q}$$

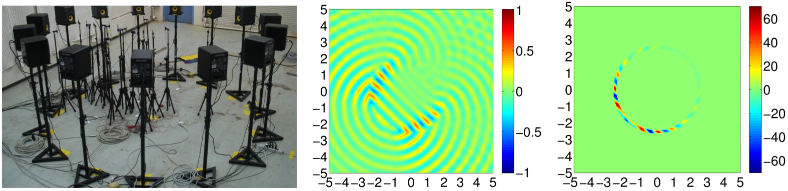
Superconductor Vortex Pinning



Courtesy Argonne National Laboratory

$$\gamma(\partial_t + i\mu)\psi = \epsilon\psi - |\psi|^2\psi + (\nabla - i\mathbf{A})^2\psi$$
$$\mathbf{J} = \mathrm{Im}(\bar{\psi}(\nabla - i\mathbf{A})\psi) - (\partial_t \mathbf{A} + \nabla\mu), \quad \nabla \cdot \mathbf{J} = 0$$

Direct Field Acoustic Testing

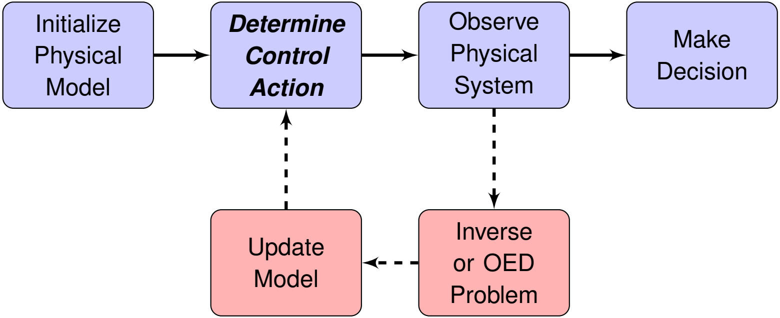


$$-\Delta u - \kappa^2(1 + \sigma\epsilon)^2 u = z$$

Sandia National Laboratories

Optimization Problem Formulation

Goal: Control uncertainty rather than quantify uncertainty.



We implement the control prior to observing the state.
Control is deterministic.

Sandia National Laboratories

Optimization of PDEs with Uncertain Inputs

Optimal Control: Given $\alpha > 0$, $\Omega_o \subseteq \Omega$, $\Omega_c \subseteq \Omega$, and $w \in L^2(\Omega_o)$.

$$\min_{z \in \mathcal{Z}} \; J(z) \equiv \frac{1}{2}\mathcal{R}\left[\int_{\Omega_o} ((U(z))(\xi, x) - w(x))^2 \, \mathrm{d}x\right] + \frac{\alpha}{2}\int_{\Omega_c} z^2(x) \, \mathrm{d}x$$

where $U(z) = u : \Xi \to H^1(\Omega)$ solves the **weak form** of

$$\begin{aligned} -\nabla \cdot (\epsilon(\xi)\nabla u(\xi)) + N(u(\xi), \xi) &= \chi_{\Omega_c} z, && \text{in } \Omega, \text{ a.s.} \\ u(\xi) &= g(\xi), && \text{on } \partial\Omega, \text{ a.s.} \end{aligned}$$

Topology Optimization: Given $0 < V_0 < 1$ and $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$.

$$\min_{z \in \mathcal{Z}} \; J(z) \equiv \mathcal{R}\left[\int_{\Omega} F(\xi, x) \cdot (U(z))(\xi, x) \, \mathrm{d}x\right] \quad \text{s.t.} \quad 0 \le z \le 1, \quad \int_{\Omega} z(x) \, \mathrm{d}x \le V_0|\Omega|$$

where $U(z) = u : \Xi \to H^1(\Omega)^d$ solves the **weak form** of

$$\begin{aligned} -\nabla \cdot (\mathbf{E}(z) : \epsilon(u(\xi)) &= F(\xi), && \text{in } \Omega, \text{ a.s.} \\ \epsilon(u(\xi)) &= \frac{1}{2}(\nabla u(\xi) + \nabla u(\xi)^\top), && \text{in } \Omega, \text{ a.s.} \\ u(\xi) &= g(\xi), && \text{on } \partial\Omega, \text{ a.s.} \end{aligned}$$

General PDE-Optimization under Uncertainty

$(\Omega, \mathcal{F}, P)$ is a complete probability space and $\boldsymbol{\xi} : \Omega \to \Xi$ is a random variable. Consider

$$\min_{z \in \mathcal{Z}_{\text{ad}}} J(z) = \mathcal{R}(f((U(z))(\boldsymbol{\xi}), z, \boldsymbol{\xi}))$$

where $U(z) = u \in L^p_P(\Omega; \mathcal{U})$ solves the **weak form** PDE

$$e(u, z, \boldsymbol{\xi}) = 0 \quad \text{and} \quad \mathcal{Z}_{\text{ad}} \subseteq \mathcal{Z}.$$

Assumptions:

- $\mathcal{U}$ is a reflexive Banach space and $\mathcal{Z}$ is a Hilbert spaces.
- For each $z \in \mathcal{Z}_{\text{ad}}$ and $\xi \in \Xi$, $e(u, z, \xi) = 0$ is well posed, i.e.,
  - $\exists!\, U(z) = u \in L^p_P(\Xi; \mathcal{U})$ such that $e(U(z), z, \xi) = 0$;
  - $\exists\, c > 0$ indpendent of z and $\xi \in \Xi$ such that $\|U(z)\|_{\mathcal{U}} \leq c(\|z\|_{\mathcal{Z}} + 1)$.
- e is a.s. sequentially weakly continuous.
- f is a.s. sequentially weakly lsc and $\xi \mapsto f((U(z))(\xi), z, \xi) \in L^q_{P\circ\boldsymbol{\xi}^{-1}}(\Xi)$.
- $\mathcal{Z}_{\text{ad}}$ is convex, closed and bounded – or –
  $\mathcal{Z}_{\text{ad}} = \mathcal{Z}$ and $z \mapsto f((U(z))(\xi), z, \xi)$ is a.s. coercive i.e.,
  $\exists\, r > 0$ and coercive $\varphi : \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$, both independent of ξ, s.t.

$$\|z\|_{\mathcal{Z}} \geq r \quad \Longrightarrow \quad f((U(z))(\xi), z, \xi) \geq \varphi(z) \ \text{ a.s.}$$

Sandia National Laboratories

Known v.s. Unknown Probability Distribution

Known Probability Distribution:

- $\Xi \subseteq \mathbb{R}^M$ is known and $P \circ \boldsymbol{\xi}^{-1}$ has Lebesgue density $\rho : \Xi \to [0, \infty)$.
- Enables UQ techniques including gPC, collocation, and sampling.
- All analysis performed in $L^p_\rho(\Xi)$ instead of $L^p_{P\circ\boldsymbol{\xi}^{-1}}(\Xi)$.

Unknown Probability Distribution:

- Must determine optimal solutions that are robust to unknown pdf.
- Use data to estimate pdf (i.e., experimental data or inverted coefficients).
- Formulate optimization problem as a minimax problem

$$\min_{z \in \mathcal{Z}_{\mathrm{ad}}} \sup_{P \in \mathcal{A}} \mathbb{E}_P[f((U(z), z, \cdot)]$$

- $\mathcal{A}$ is the **ambiguity** set and is defined with data, i.e., moment matching.
- Must discretize the probability measures $P \in \mathcal{A}$.
- Require specialized optimization algorithms to efficiently solve.

Sandia National Laboratories

Outline

Known Probability Distribution

Unknown Probability Distribution

Sandia National Laboratories

Risk Measures

Assumptions:

- $\mathcal{R} : L^q_{P\circ\boldsymbol{\xi}^{-1}}(\Xi) \to \mathbb{R} \cup \{+\infty\}$
  see Rockafellar, Uryasev, Shapiro, Dentcheva, Ruszczynski, . . .
- $\mathcal{R}$ is convex, lsc and satisfies $\mathcal{R}(C) = C$ for all constants C;
- $\mathcal{R}$ is monotonic, i.e., if $X_1 \geq X_2$ a.s., then $\mathcal{R}(X_1) \geq \mathcal{R}(X_2)$.

Result: There exists a minimizer of J in $\mathcal{Z}_{\text{ad}}$.

Risk Neutral v.s. Risk Averse

- **Risk Neutral:** $\mathcal{R} \equiv \mathbb{E}$.
  - Optimal solution minimizes on average.
  - Efficiently solved with adaptive sparse grid trust-region algorithm Kouri, Heinkenschloss, Ridzal, van Bloemen Waanders.
- **Risk Averse:** $\mathcal{R}(X) > \mathbb{E}[X] \quad \forall$ nonconstant $X \in L^q_{P\circ\boldsymbol{\xi}^{-1}}(\Xi)$.
  - More conservative than $\mathcal{R} \equiv \mathbb{E}$.
  - Can minimize measures of deviation and/or tail events.

Choosing a Risk Measure

Controlling Uncertainty

- Reduce **variability** of optimized system:

$$\mathbb{E}[(X - \mathbb{E}[X])^2] \qquad \text{or} \qquad \mathbb{E}[(X - \mathbb{E}[X])_+^q]^{\frac{1}{q}}$$

- Control **rare events**, reduce **failure regions**, and certify **reliability**:

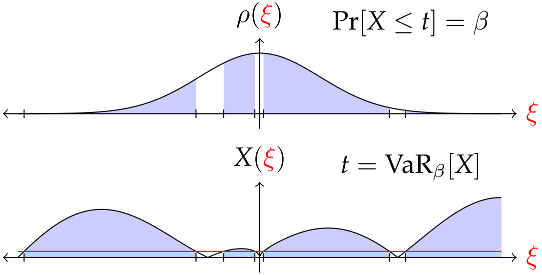
$$\Pr[X \le t] \qquad \text{or} \qquad \mathrm{VaR}_\beta[X] = \inf\{\, t \in \mathbb{R} \,:\, \Pr[X \le t] \ge \beta \,\}$$

- Minimize over **undesirable events**:

$$\mathrm{CVaR}_\beta[X] = \frac{1}{1-\beta}\int_{X \ge \mathrm{VaR}_\beta[X]} X(\omega)\, \mathrm{d}P(\omega) = \mathbb{E}[X \,|\, X \ge \mathrm{VaR}_\beta[X]]$$

CONDITIONAL VALUE-AT-RISK:

$\mathcal{R}(X) = \mathrm{CVaR}_\beta[X]$



Classification of Risk Measures

Shapiro, Dentcheva, Ruszczynski, Rockafellar, Uryasev, . . .

$\mathcal{R} : L^q_\rho(\Xi) \to \mathbb{R} \cup \{\infty\}$ is a **monetary** risk measure if for $X,\ X_1,\ X_2 \in L^q_\rho(\Xi)$

- **Monotonicity:** $X_1 \geq X_2$ a.e. $\implies \mathcal{R}(X_1) \geq \mathcal{R}(X_2)$
- **Translation Equivariance:** $\mathcal{R}(X + t) = \mathcal{R}(X) + t, \quad \forall t \in \mathbb{R}$

$\mathcal{R}$ is a **convex** risk measure if

- $\mathcal{R}$ is a monetary risk measure
- **Convexity:** $\mathcal{R}(tX_1 + (1-t)X_2) \leq t\mathcal{R}(X_1) + (1-t)\mathcal{R}(X_2), \quad \forall t \in [0, 1]$

$\mathcal{R}$ is a **coherent** risk measure if

- $\mathcal{R}$ is a convex risk measure
- **Positive Homogeneity:** $\mathcal{R}(tX) = t\mathcal{R}(X), \quad \forall t > 0.$

Examples of coherent risk measures with $X \in L^q_\rho(\Xi)$:

- Risk Neutral: $\mathcal{R}(X) = \mathbb{E}[X]$
- Mean Plus Semideviation: $\mathcal{R}(X) = \mathbb{E}[X] + c\mathbb{E}[(X - \mathbb{E}[X])_+],\ c \in (0, 1)$
- Conditional Value-at-Risk: $\mathcal{R}(X) = \inf\{\, t + c\mathbb{E}[(X - t)_+] \ :\ t \in \mathbb{R}\,\},\ c > 1$

Duality Theory of Risk Measures

The Fenchel-Moreau Theorem $\implies$ if $\mathcal{R}$ is a **convex** risk measure, then

$$\mathcal{R}(X) = \sup_{\vartheta \in \mathrm{dom}(\mathcal{R}^*)} \{\mathbb{E}[\vartheta X] - \mathcal{R}^*(\vartheta)\}$$

where $\mathcal{R}^*$ is the conjugate of $\mathcal{R}$, i.e., $\mathcal{R}^*(\vartheta) = \sup_{X \in \mathrm{dom}(\mathcal{R})} \{\mathbb{E}[\vartheta X] - \mathcal{R}(X)\}$.

Moreover, if $\mathcal{R}$ is a **coherent** risk measure, then

$$\mathcal{R}(X) = \sup_{\vartheta \in \mathrm{dom}(\mathcal{R}^*)} \mathbb{E}[\vartheta X].$$

$\mathrm{dom}(\mathcal{R}^*)$ is the **risk envelope** $\implies$ related to **ambiguity set**.

Example (Conditional Value-at-Risk):

$$\mathcal{R}(X) = \mathrm{CVaR}_\beta[X] = \inf_t \left\{ t + (1-\beta)^{-1}\mathbb{E}[(X-t)_+] \right\} = \sup_{\vartheta \in \mathrm{dom}(\mathcal{R}^*)} \mathbb{E}[\vartheta X]$$

$$\mathrm{dom}(\mathcal{R}^*) = \left\{ \vartheta \in (L^q_\rho(\Xi))^* \;:\; \mathbb{E}[\vartheta] = 1,\; 0 \le \vartheta \le \frac{1}{1-\beta}\;\rho\text{-a.e.} \right\}.$$

Outline

Known Probability Distribution

Unknown Probability Distribution

Sandia National Laboratories

Distributionally Robust PDE-Optimization

Recall: $(\Xi, \mathcal{F})$ is a measurable space and prob. measure is *unknown*.

- $\mathcal{M}$ is the Banach space of regular Borel measures on $\mathcal{F}$, i.e.,

$$C(\Xi)^* \cong \mathcal{M}.$$

- $\mathcal{M}^+ \subset \mathcal{M}$ is the set of positive measures, i.e.,

$$\mu \in \mathcal{M}^+ \quad \Longrightarrow \quad \mu(V) \geq 0 \quad \forall V \in \mathcal{F}.$$

- **Ambiguity Set:** $\mathcal{A} \subset \mathcal{M}$ defined by data. For example:
  - **Moment Matching:** Given generalized moment data $m_1, \ldots, m_N$,

$$\mathcal{A} = \left\{ P \in \mathcal{M}^+ \,:\, P(\Xi) = 1,\, \int_\Xi \psi_i(\xi)\, \mathrm{d}P(\xi) = m_i,\, i = 1, \ldots, N \right\}.$$

  - **Φ-Divergence (e.g., Kullback-Leibler):** Given an estimated prob. measure P_0 and $\epsilon > 0$,

$$\mathcal{A} = \left\{ P \in \mathcal{M}^+ \,:\, P(\Xi) = 1,\, D_\Phi(P, P_0) \leq \epsilon \right\}.$$

- **Distributionally-robust (a.k.a. data-driven) optimization problem:**

$$\min_{z \in \mathcal{Z}_{\mathrm{ad}}} \sup_{P \in \mathcal{A}} \int_\Xi f((U(z))(\xi), z, \xi)\, \mathrm{d}P(\xi).$$

Measure Approximation

General Approach:

1. Let $\{\varphi_i\}_{i=1}^n$ be a partition of unity on Ξ and $\mu \in \mathcal{M}$ be any measure.
2. Define the "localized" measures

$$\mu_i(V) = \int_V \varphi_i(\xi)\,\mathrm{d}\mu(\xi).$$

3. Note $\mu(\Xi) = \mu_1(\Xi) + \ldots + \mu_n(\Xi)$.
4. Define the projection operators $\Pi_n^\mu : C(\Xi) \to \text{span}\{\varphi_1, \ldots, \varphi_n\}$ as

$$\Pi_n^\mu y = \sum_{i=1}^n \mu_i(\Xi)^{-1} \int_\Xi y(\xi)\,\mathrm{d}\mu_i(\xi)\ \varphi_i \quad \forall y \in C(\Xi)$$

and $\Lambda_n^\mu : \mathcal{M} \to \text{span}\{\mu_1, \ldots, \mu_n\}$ as

$$\Lambda_n^\mu \nu = \sum_{i=1}^n \mu_i(\Xi)^{-1} \int_\Xi \varphi_i(\xi)\,\mathrm{d}\nu(\xi)\ \mu_i \quad \forall \nu \in \mathcal{N},$$

Approximation Properties

- **Lemma:** If $\mu \in \mathcal{M}^+$ is σ-finite, then $\Lambda_n^\mu \nu$ is absolutely continuous with respect to μ with density

$$f_n^\mu[\nu] = \sum_{i=1}^{n} \mu_i(\Xi)^{-1} \int_\Xi \varphi_i(\xi)\, \mathrm{d}\nu(\xi) \varphi_i \quad \text{for any } \nu \in \mathcal{M}.$$

- **Lemma:** Λ_n^μ is invariant on the space of probability measures.
- **Lemma:** Π_n^μ is the adjoint of Λ_n^μ.
- **Theorem:** Let $V_i = \mathrm{supp}(\varphi_i)$ and $\|\cdot\|_{u,V_i}$ denote the uniform norm on V_i. Then, there exists $c_i > 0$ such that

$$|\langle \nu - \Lambda_n^\mu \nu, y\rangle_{\mathcal{M},C(\Xi)}| \le \sum_{i=1}^{n} c_i \left\{ \int_\Xi \sqrt{\varphi_i(\xi)}\, \mathrm{d}|\nu|(\xi) \right\} \inf_{\bar{y}_i \in \mathbb{R}} \|\sqrt{\varphi_i}(y - \bar{y}_i)\|_{u,V_i}.$$

Sandia National Laboratories

Measure Approximation

Piecewise Constants:

1. Let $\{V_i\}_{i=1}^n$ be a tesselation of Ξ and define $\varphi_i = \chi_{V_i}$.
2. The "localized" measures are

$$\mu_i(V) = \mu(V \cap V_i).$$

3. The projection operator $\Pi_n^\mu : C(\Xi) \to \text{span}\{\varphi_1, \ldots, \varphi_n\}$ is

$$\Pi_n^\mu y = \sum_{i=1}^n \mu(V_i)^{-1} \int_{V_i} y(\xi)\, \mathrm{d}\mu(\xi)\ \chi_{V_i} \quad \forall\, y \in C(\Xi)$$

and $\Lambda_n^\mu : \mathcal{M} \to \text{span}\{\mu_1, \ldots, \mu_n\}$ is

$$\Lambda_n^\mu \nu = \sum_{i=1}^n \mu(V_i)^{-1} \nu(V_i) \mu_i \quad \forall\, \nu \in \mathcal{N},$$

- **Theorem:** Suppose V_i are convex, bounded, and Lipschitz, and $\mu \in \mathcal{M}$. Then $\exists\, c > 0$ only depending on M such that

$$\|\nu - \Lambda_n^\mu \nu\|_{W^{1,\infty}(\Xi)^*} \leq c \sum_{i=1}^n \left(1 + \frac{|\mu|(V_i)}{|\mu(V_i)|}\right) |\nu|(V_i)\ \mathrm{diam}(V_i).$$

Sandia National Laboratories

Example — Voronoi Tesselation

Suppose $\Xi = [0,1]$ and P has pdf

$$\text{pdf}(\xi) = \frac{\beta}{1-e^{-\beta}} e^{-\beta\xi} \quad \text{for} \quad \beta > 0.$$

Approx. P using piecewise constant projection and μ set to the uniform prob. measure:

$$\text{approx-pdf}(\xi) = \sum_{i=1}^{n} \frac{(e^{-\beta a_{i-1}} - e^{-\beta a_i})}{(1-e^{-\beta})(a_i - a_{i-1})} \chi_{[a_{i-1},a_i]}(\xi).$$

β	n	Error	Sum W. Diam.	Max. Diam.	Max. W. Diam.
1	10	3.592×10^{-2}	1.438×10^{-1}	2.518×10^{-1}	5.899×10^{-2}
	100	3.740×10^{-3}	1.496×10^{-2}	4.269×10^{-2}	1.471×10^{-3}
	1000	3.751×10^{-4}	1.501×10^{-3}	6.089×10^{-3}	2.733×10^{-5}
	10000	3.750×10^{-5}	1.500×10^{-4}	7.955×10^{-4}	4.404×10^{-7}
10	10	2.282×10^{-1}	1.304×10^{-1}	7.572×10^{-1}	1.010×10^{-1}
	100	3.053×10^{-2}	1.451×10^{-2}	5.328×10^{-1}	8.191×10^{-3}
	1000	3.551×10^{-3}	1.502×10^{-3}	3.133×10^{-1}	5.424×10^{-4}
	10000	3.763×10^{-4}	1.517×10^{-4}	1.300×10^{-1}	2.710×10^{-5}
100	10	3.076×10^{-1}	1.226×10^{-1}	9.758×10^{-1}	1.194×10^{-1}
	100	4.128×10^{-2}	1.327×10^{-2}	9.531×10^{-1}	1.261×10^{-2}
	1000	5.022×10^{-3}	1.348×10^{-3}	9.301×10^{-1}	1.247×10^{-3}
	10000	5.899×10^{-4}	1.360×10^{-4}	9.072×10^{-1}	1.224×10^{-4}

Approximation and Optimization Algorithms

Given an arbitrary $\mu \in \mathcal{M}^+$ with $\mu(\Xi) = 1$, we approximate

$$J(z) = \sup_{P \in \mathcal{A}} \int_\Xi f((U(z))(\xi), z, \xi)\, \mathrm{d}P(\xi)$$

using our measure discretization, i.e.,

$$J_n(z) = \sup_{p \in \mathcal{A}_n} \sum_{i=1}^n \frac{p_i}{\mu_i(\Xi)} \int_\Xi f(U(z)(\xi), z, \xi)\, \mathrm{d}\mu_i(\xi), \quad \mathcal{A}_n = \left\{ p \in \mathbb{R}^n \;:\; \sum_{i=1}^n \frac{p_i}{\mu_i(\Xi)} \mu_i \in \mathcal{A} \right\}.$$

- **Theorem (Piecewise Constants):** If $\xi \mapsto f(U(z)(\xi), z, \xi) \in W^{1,\infty}(\Xi)$ and z_n minimizes J_n defined on a family of tesselations $\{V_{ni}\}_{i=1}^n$ satisfying

  $$\lim_{n \to \infty} \sup_{P \in \mathcal{A}} \sum_{i=1}^n P(V_{ni}) \mathrm{diam}(V_{ni}) = 0.$$

  Then, z_n has a w-converging subsequence and the w-limit minimizes J.
- J and J_n may **not** be differentiable!
  - J and J_n are *Fréchet subdifferentiable.*
  - Compute value and subgradient using linear/convex optimization.
  - **Cannot** use derivative-based optimization algorithms.
  - Subgradient descent and bundle methods converge *sublinearly*.
- **Expensive PDEs $\implies$ Need rapid optimization algorithms.**

Sandia National Laboratories

Example — Moment Matching

Let $\psi_i : \Xi \to \mathbb{R}$ be $\mathcal{F}$-measurable functions and $m_i \in \mathbb{R}$ for $i = 1, \ldots, N$

$$\mathcal{A} = \left\{ P \in \mathcal{M}^+ \;:\; P(\Xi) = 1, \quad \begin{array}{l} \int_\Xi \psi_i(\xi)\,\mathrm{d}P(\xi) = m_i,\ i = 1, \ldots, N_e \\ \int_\Xi \psi_i(\xi)\,\mathrm{d}P(\xi) \le m_i,\ i = N_e + 1, \ldots, N \end{array} \right\}.$$

Theorem (Shapiro): If $\mathcal{A} \neq \emptyset$, then for each $z \in \mathcal{Z}$ there exists ξ_i and $p_i \ge 0$ with $p_1 + \cdots + p_{N+1} = 1$ such that

$$\sup_{P \in \mathcal{A}} \int_\Xi f((U(z))(\xi), z, \xi)\,\mathrm{d}P(\xi) = \sum_{i=1}^{N+1} p_i f((U(z))(\xi_i), z, \xi_i)$$

Approximation: Localized measures μ_j

$$\mathcal{A}_n = \left\{ p \in \mathbb{R}^n \;:\; \sum_{j=1}^n p_j = 1, \quad \begin{array}{l} \sum_{j=1}^n \frac{p_j}{\mu_j(\Xi)} \int_\Xi \psi_i(\xi)\,\mathrm{d}\mu_j(\xi) = m_i,\ i = 1, \ldots, N_e \\ \sum_{j=1}^n \frac{p_j}{\mu_j(\Xi)} \int_\Xi \psi_i(\xi)\,\mathrm{d}\mu_j(\xi) \le m_i,\ i = N_e + 1, \ldots, N \end{array} \right\}.$$

Theorem (Kouri): If $\mathcal{A}_n \neq \emptyset$, then for each $z \in \mathcal{Z}$ there exists $p_i \ge 0$ with at most $\min\{n, N+1\}$ nonzero such that $p_1 + \cdots + p_{N+1} = 1$ and

$$\sup_{q \in \mathcal{A}_n} \sum_{j=1}^n \frac{q_j}{\mu_j(\Xi)} \int_\Xi f((U(z))(\xi), z, \xi)\,\mathrm{d}\mu_j(\xi) = \sum_{j=1}^{N+1} \frac{p_j}{\mu_j(\Xi)} \int_\Xi f((U(z))(\xi), z, \xi)\,\mathrm{d}\mu_j(\xi).$$

Sandia National Laboratories

Example — Moment Matching

Optimal Control of 1D Elliptic Equation

Let $\alpha = 10^{-4}$, $\Omega_o = \Omega_c = \Omega = (-1,1)$, and $w \equiv 1$ and consider

$$\underset{z \in L^2(-1,1)}{\text{minimize}} \quad J(z) = \frac{1}{2}\mathcal{R}\left[\int_{-1}^{1} (U(z)(\cdot,x) - 1)^2 \, \mathrm{d}x\right] + \frac{\alpha}{2}\int_{-1}^{1} z(x)^2 \, \mathrm{d}x$$

where $U(z) = u \in L^2_\rho(\Xi; H^1_0(0,1))$ solves the weak form of

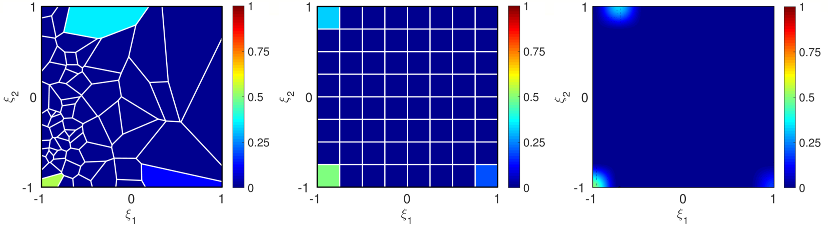
$$\begin{aligned} -\partial_x \left(\epsilon(\xi,x)\partial_x u(\xi,x)\right) &= f(\xi,x) + z(x) & (\xi,x) \in \Xi \times \Omega, \\ u(\xi,-1) = 0, \quad u(\xi,1) &= 0 & \xi \in \Xi. \end{aligned}$$

$\Xi = [-0.1, 0.1] \times [-0.5, 0.5]$, the true distribution is a tensor product of truncated exponentials, and the random field coefficients are

$$\epsilon(\xi,x) = 0.1\chi_{(-1,\xi_1)} + 10\chi_{(\xi_1,1)}, \quad \text{and} \quad f(\xi,x) = \exp(-(x-\xi_2)^2).$$

Example — Moment Matching

$$P(\Xi) = 1, \quad \int_\Xi \xi_1 \, dP(\xi) \approx -0.537, \quad \text{and} \quad \int_\Xi \xi_2 \, dP(\xi) \approx -0.313$$



- **Left:** Voronoi ($n = 64$) with 1000 MC samples per cell.
- **Center:** Uniform ($n = 64$) with level 4 sparse grids.
- **Right:** C^2 parition of unity ($n = 64$) with level 4 sparse grids, i.e., shifted/scaled tensor products of

$$\theta(x) = \begin{cases} 4x^2(3-4x) & \text{if } 0 < x \le \frac{1}{2} \\ 4(x-1)^2(4x-1) & \text{if } \frac{1}{2} < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Example — Moment Matching

	n	Obj. Val.	Center	Prob.	Center	Prob.	Center	Prob.
Voronoi 1000	16	0.13457	$(-0.864, -0.893)$	0.435	$(-0.634, 0.841)$	0.328	$(0.195, -0.848)$	0.237
	64	0.13777	$(-0.882, -0.933)$	0.540	$(-0.331, 0.849)$	0.346	$(0.467, -0.909)$	0.114
	256	0.14056	$(-0.981, -0.983)$	0.605	$(0.116, 0.922)$	0.351	$(0.330, -0.960)$	0.044
	1024	0.14133	$(-0.126, -0.987)$	0.484	$(-0.916, 0.988)$	0.342	$(-0.939, -0.994)$	0.174
	4096	0.14207	$(-0.978, -0.997)$	0.368	$(-0.813, 0.988)$	0.343	$(0.350, -0.991)$	0.289
Square $\ell = 4$	16	0.13221	$(-0.750, -0.750)$	0.709	$(-0.750, 0.750)$	0.150	$(0.750, 0.750)$	0.142
	64	0.13779	$(-0.857, -0.875)$	0.496	$(-0.875, 0.875)$	0.321	$(0.875, -0.875)$	0.193
	256	0.14058	$(-0.063, -0.938)$	0.457	$(-0.938, 0.938)$	0.333	$(-0.938, -0.938)$	0.210
	1024	0.14194	$(-0.969, -0.969)$	0.438	$(-0.969, 0.969)$	0.338	$(0.906, -0.969)$	0.223
	4096	0.14286	$(-1.000, -1.000)$	0.433	$(-0.968, 1.000)$	0.342	$(1.000, -1.000)$	0.225
C^2 $\ell = 4$	16	0.13444	$(-1.000, -1.000)$	0.696	$(1.000, 1.000)$	0.164	$(-1.000, 1.000)$	0.140
	64	0.13953	$(-1.000, -1.000)$	0.501	$(-0.714, 1.000)$	0.329	$(1.000, -1.000)$	0.170
	256	0.14154	$(-1.000, -1.000)$	0.663	$(0.867, 1.000)$	0.231	$(-1.000, 1.000)$	0.106
	1024	0.14244	$(-1.000, -1.000)$	0.441	$(-0.935, 1.000)$	0.340	$(1.000, -1.000)$	0.218
	4096	0.14286	$(-1.000, -1.000)$	0.433	$(-0.968, 1.000)$	0.342	$(1.000, -1.000)$	0.225
	$\star$	0.15640	$(-0.995, -0.996)$	0.657	$(0.432, 1.000)$	0.323	$(-0.993, 0.999)$	0.019

$\star$ Computed using Gaivoronski's stochastic descent algorithm for moment matching.

Example — CVaR

Optimal Control of 1D Elliptic Equation

Let $\alpha = 10$, $\Omega_o = \Omega_c = \Omega = (-1, 1)$, and $w \equiv 1$ and consider

$$\underset{z \in L^2(-1,1)}{\text{minimize}} \quad J(z) = \frac{1}{2}\mathcal{R}\left[\int_{-1}^{1} (U(z)(\cdot, x) - 1)^2 \, \mathrm{d}x\right] + \frac{\alpha}{2}\int_{-1}^{1} z(x)^2 \, \mathrm{d}x$$

where $U(z) = u \in L^2_\rho(\Xi; H^1_0(0,1))$ solves the weak form of

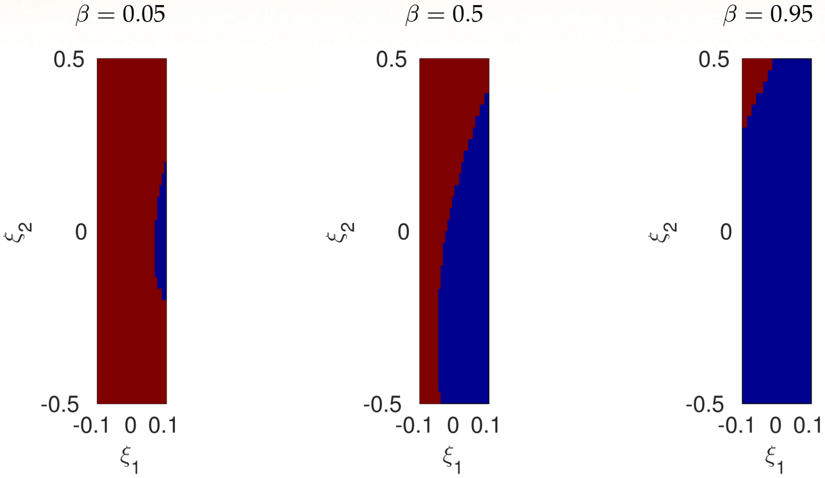
$$\begin{aligned} -\partial_x \left(\epsilon(\xi, x)\partial_x u(\xi, x)\right) &= f(\xi, x) + z(x) && (\xi, x) \in \Xi \times \Omega, \\ u(\xi, -1) = 0, \quad u(\xi, 1) &= 0 && \xi \in \Xi. \end{aligned}$$

$\Xi = [-0.1, 0.1] \times [-0.5, 0.5]$ is endowed with the uniform density $\rho \equiv 5$ and the random field coefficients are

$$\epsilon(\xi, x) = 0.1\chi_{(-1,\xi_1)} + 10\chi_{(\xi_1,1)}, \quad \text{and} \quad f(\xi, x) = \exp(-(x - \xi_2)^2).$$

Example — CVaR

Discretization: Uniform ($n = 900$) with level 4 sparse grids.



$$\mathcal{A}_n = \left\{ p \in \mathbb{R}^n \; : \; \sum_{i=1}^{n} p_i = 1, \; 0 \le p_i \le \frac{\mu(V_i)}{1-\beta}, \; i = 1, \ldots, n \right\}$$

Conclusions:

- **Risk Neutral:**
  - Can efficiently solve using adaptive sparse grids and trust regions.
- **Risk Averse:**
  - Risk measures often not differentiable;
  - Define smooth risk measures using the risk quadrangle;
  - Can use Newton's method/quad. and can prove error bounds.
- **Unknown Distribution:**
  - Incorporate data into distributionally-robust opt. formulation;
  - Objective function not differentiable;
  - Nonsmooth optimization algorithms converge slowly.

Future Work:

- **Risk measures:** Develop error indicators and use locally adaptive sparse grids with trust-region algorithm.
- **Unknown distribution:** Develop opt. algorithm with adaptive tessellation and sampling that exploits PDE constraint.
- Incorporate **(buffered) probabilistic objectives and constraints** to control *tail-probabilities* and *rare events* (Rockafellar, Uryasev, Royset, Shapiro, Henrion, Kibzun, ...)