Optimization and Control Unde SAND2018-3851C

Drew Kouri

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

April 19, 2018

SIAM Conference on Uncertainty Quantification
Garden Grove, CA.

Funding: DARPA EQUiPS

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Outline

Motivating Applications

General Problem Formulation

Quantifying Risk

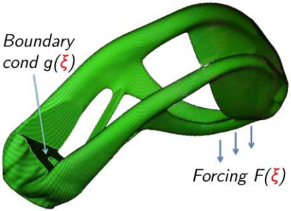
Probabilistic Optimization

What if our uncertainty is uncertain?

Computational Solution Methods

References

Topology Optimization & Additive Manufacturing



Given $V_0 \in (0,1)$ compute a density that solves:

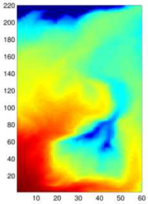
$$\underset{0 \le z \le 1}{\text{Minimize}} \ \mathcal{R}\left(\int_D F \cdot S(z)\,dx + \int_{\Gamma_t} t \cdot S(z)\,dx\right)$$

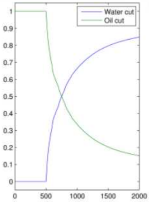
s.t. $\int_D z(x)\,dx \le V_0|D|$, where $S(z) = u$ solves the **linear elasticity equations**

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

- Uncertain external forces (loads) and boundary conditions.
- Uncertain internal forces, e.g., residual stresses due to AM.
- Uncertain material properties (porosity, etc.) due to AM.
- **Reliability formulation:** Compute light-weight designs that minimize the probability of structural failure.

Reservoir Optimization: Secondary Oil Recovery





Given $D \subset \mathbb{R}^3$ and interest rate $r \geq 0$:

$$\underset{z=(q,\widehat{q})}{\text{Minimize}} \ \mathcal{R}\left(\int_0^T e^{rt}\, C([S(z)](t), z(t), t)\, dt\right)$$

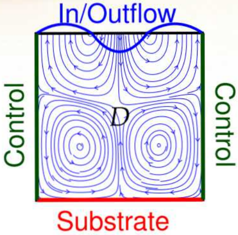
where $S(z) = (s, v, p)$ solves the **reservoir equations**

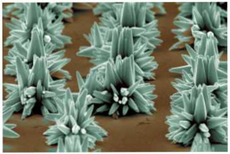
$$\begin{aligned} -\mathbf{K}\lambda(s)\nabla p &= v, && \text{in } D, \text{ a.s.} \\ \nabla \cdot v &= q, && \text{in } D, \text{ a.s.} \\ \phi\, \partial_t s + \nabla \cdot (f(s)v) &= \widehat{q}, && \text{in } D, \text{ a.s.} \end{aligned}$$

(plus initial and boundary conditions).

- Porosity, ϕ, and permeability, $\mathbf{K}$, are estimated from data (e.g., seismic inversion).
- Total mobility, λ, and fractional flow function, f, may be uncertain.
- **Risk-neutral formulation:** Determine injection rates that minimize cost on average.
- **Risk-averse formulation:** Determine injection rates that minimize the average of the 10% worst costs.

Control of Chemical Vapor Deposition Reactors





Consider the optimal control problem

$$\min_{z} \; \frac{1}{2}\mathcal{R}\left(\int_D (\nabla \times U(z))\,\mathrm{d}x\right) + \frac{\gamma}{2}\int_{\Gamma_c} |z|^2\,\mathrm{d}x$$

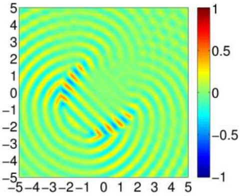
where $S(z) = (U(z), P(z), T(z)) = (u, p, t)$ solves the **Boussinesq flow equations**

$$\begin{aligned} -\nu\nabla^2 u + (u\cdot\nabla)u + \nabla p + \eta t g = 0 &\quad \text{in } D, \text{ a.s.} \\ \nabla\cdot u = 0 &\quad \text{in } D, \text{ a.s.} \\ -\kappa\Delta t + u\cdot\nabla t = 0 &\quad \text{in } D, \text{ a.s.} \\ \kappa\nabla t\cdot n + h(z - t) = 0 &\quad \text{on } \Gamma_c, \text{ a.s.} \end{aligned}$$

(plus additional boundary conditions).

- Uncertain viscosity, thermal conductivity, substrate temperature, etc. imply flow velocity, pressure and temperature are uncertain.
- **Risk-averse formulation:** Determine wall temperature that minimizes the average of *low-probability*, large vorticity scenarios.

Direct Field Acoustic Testing



Consider the optimal control problem

$$\min_{z} \ \frac{1}{2}\mathcal{R}\left(\int_{D_o} (U(z)-w)\overline{(U(z)-w)}\,\mathrm{d}x\right) + \frac{\gamma}{2}\int_{D_c} |z|^2\,\mathrm{d}x$$

where $U(z) = u$ solves the **Helmholtz equation**

$$\begin{aligned} -\Delta u - \kappa^2(1+\sigma\epsilon)^2 u &= \mathbb{1}_{D_c} z \quad && \text{in } D, \text{ a.s.} \\ \nabla u \cdot n &= i\kappa u \quad && \text{on } \partial D, \text{ a.s.} \end{aligned}$$

- The refractive index of the device under investigation is often uncertain.
- **Risk-neutral formulation:** Determine speaker output that produces a material response that matches a desired vibration profile on average.
- **Risk-averse formulation:** Determine speaker output that produces a response that is “good” on average for the 10% worst scenarios.

Outline

Motivating Applications

General Problem Formulation

Quantifying Risk

Probabilistic Optimization

What if our uncertainty is uncertain?

Computational Solution Methods

References

General PDE-Optimization under Uncertainty

Making Deterministic Problems Stochastic

Deterministic PDE-Constrained Optimization:
U and Z are reflexive Banach spaces, Z_{ad} is a closed convex subset of Z, Y is a Banach space, $J : U \times Z \to \mathbb{R}$ and $c : U \times Z \to Y$:

$$\underset{z \in Z_{\text{ad}}}{\text{Minimize}} \ \widehat{J}(z)$$

where $\widehat{J}(z) := J(S(z), z)$ and $S(z) = u \in U$ solves the PDE

$$c(u, z) = 0.$$

Stochastic PDE-Constrained Optimization:
$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. Objective function and PDE are now **parametrized**, i.e., $J : U \times Z \times \Omega \to \mathbb{R}$ and $c : U \times Z \times \Omega \to Y$:

$$\underset{z \in Z_{\text{ad}}}{\text{Minimize}} \ \mathcal{J}(z) = \mathcal{R}(\widehat{J}(z))$$

where $\widehat{J}(z) := J(S(z), z, \cdot)$ and $S(z) = u : \Omega \to U$ solves the PDE

$$c(u, z, \omega) = 0.$$

Notation

$(\Omega, \mathcal{F})$ is a measurable space

$\mathbb{P}, P : \mathcal{F} \to [0,1]$ are probability measures

1. **Expectation:** $\mathbb{E}_P[X] = \int_\Omega X(\omega)\,\mathrm{d}P(\omega)$ and $\mathbb{E}[X] = \mathbb{E}_\mathbb{P}[X]$
2. **Variance:** $\mathbb{V}_P[X] = \mathbb{E}_P[(X - \mathbb{E}_P[X])^2]$ and $\mathbb{V}[X] = \mathbb{V}_\mathbb{P}[X]$
3. **Standard Deviation:** $\sigma_P[X] = \mathbb{V}_P[X]^{1/2}$ and $\sigma[X] = \sigma_\mathbb{P}[X]$
4. **Distribution:** $F_X(x) = \mathbb{P}(X \le x)$
5. **Quantile:** $q_\beta(X) = \inf\{t \in \mathbb{R} \mid F_X(x) > \beta\} = F_X^{-1}(\beta)$

Tensor Product Function Spaces

Lebesgue Spaces: For $1 \leq p < \infty$,

$$L^p(\Omega, \mathcal{F}, \mathbb{P}) := \left\{ v : \Omega \to \mathbb{R} \mid v \; \mathcal{F}\text{-measurable}, \int_\Omega |v(\omega)|^p \, \mathrm{d}\mathbb{P}(\omega) < \infty \right\},$$

$$L^\infty(\Omega, \mathcal{F}, \mathbb{P}) := \{ v : \Omega \to \mathbb{R} \mid v \; \mathcal{F}\text{-measurable}, \operatorname{ess\,sup} |v(\omega)| < \infty \}.$$

If $f, g \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ then $f = g \iff f(\omega) = g(\omega)$ for $\mathbb{P}$ almost all $\omega \in \Omega$.

Tensor Spaces: Given a real Banach space W then

$$L^p(\Omega, \mathcal{F}, \mathbb{P}) \otimes W := \operatorname{span} \{ vx \mid v \in L^p(\Omega, \mathcal{F}, \mathbb{P}), \; x \in W \}.$$

Many norms exist for the vector space $L^p(\Omega, \mathcal{F}, \mathbb{P}) \otimes W$ and given a norm $L^p(\Omega, \mathcal{F}, \mathbb{P}) \otimes W$ is not necessarily complete.

Bochner Spaces: For $1 \leq p < \infty$ and W a real Banach space

$$L^p(\Omega, \mathcal{F}, \mathbb{P}; W) := \left\{ v : \Omega \to W \mid v \text{ strongly } \mathcal{F}\text{-measurable}, \int_\Omega \|v(\omega)\|_W^p \, \mathrm{d}\mathbb{P}(\omega) < \infty \right\}$$

and similarly for $p = \infty$. $L^p(\Omega, \mathcal{F}, \mathbb{P}; W)$ is the completion of $L^p(\Omega, \mathcal{F}, \mathbb{P}) \otimes W$ with respect to the Bochner norm

$$\|u\|_{L^p(\Omega, \mathcal{F}, \mathbb{P}; W)} := \left(\int_\Omega \|u(\omega)\|_W^p \, \mathrm{d}\mathbb{P}(\omega) \right)^{\frac{1}{p}} \quad \text{and} \quad \|u\|_{L^\infty(\Omega, \mathcal{F}, \mathbb{P}; W)} := \operatorname{ess\,sup} \|u(\omega)\|_W.$$

Again, if $f, g \in L^p(\Omega, \mathcal{F}, \mathbb{P}; W)$ then $f = g \iff f(\omega) = g(\omega)$ for $\mathbb{P}$ almost all $\omega \in \Omega$.

Assumptions on PDE Solution Map $S(z)$

1. For each $z \in Z$, $c(u, z, \omega) = 0$ is well posed, i.e.,
   (i) $\exists!\, S(z) : \Omega \to U$ such that $c(S(z), z, \cdot) = 0$ a.s. for all z;
   (ii) $\exists\, 0 < c(\cdot) \in L^q(\Omega, \mathcal{F}, \mathbb{P})$, $1 \le q \le \infty$ and an increasing function $\rho : [0, \infty) \to [0, \infty)$ both independent of z such that

$$\|S(z)\|_U \le c\rho(\|z\|_Z) \quad \text{a.s.} \quad \forall z \in Z_{\text{ad}}.$$

2. $S(z)$ is strongly measurable $\forall z \in Z_{\text{ad}} \implies S(z) \in L^q(\Omega, \mathcal{F}, \mathbb{P}; U)$.
3. $z \mapsto S(z)$ satisfies the continuity property

$$z_n \rightharpoonup z \text{ in } Z \quad \implies \quad S(z_n) \rightharpoonup S(z) \text{ in } U, \text{ a.s.}$$

4. $\exists V \supseteq Z_{\text{ad}}$, V open, such that $S : V \to L^q(\Omega, \mathcal{F}, \mathbb{P}; U)$ is continuously Fréchet differentiable.

Senstivity Equation: To compute the sensitivity of $S(z)$ in the direction $h \in Z$ solve:

$$c_u(S(z), z, \cdot)S'(z)h + c_z(S(z), z, \cdot)h = 0 \quad \text{a.s.}$$

Example: Linear Elliptic PDE

Let $D \subset \mathbb{R}^n$ be a bounded Lipschitz domain, $U = H_0^1(D)$, $Y = Z = H^{-1}(D)$ and $A : \Omega \to \mathbb{R}^{n \times n}$:

$$\langle c(u, z, \omega), v \rangle_{U^*,U} := \int_D (A(\omega) \nabla u(x)) \cdot \nabla v(x) \, \mathrm{d}x - \langle z, v \rangle_{U^*,U} \quad \text{for } v \in H_0^1(D).$$

If $\exists \, 0 < \underline{c} \leq \overline{c} < \infty$ such that

$$\underline{c} \leq \frac{\zeta^\top A(\omega) \zeta}{\zeta^\top \zeta} \leq \overline{c} \quad \text{a.s.}$$

then Lax-Milgram $\implies$ existence of a unique solution $u \in H_0^1(D)$ to $c(u, z, \cdot) = 0$ for fixed z a.s. Moreover,

$$\underline{c} \| \nabla S(z) \|_{L^2(D)}^2 \leq \| z \|_{H^{-1}(D)} \| S(z) \|_{H_0^1(D)} \quad \text{a.s.}$$

Hence, Poincaré's inequality guarantees that

$$\| \nabla S(z) \|_{L^2(D)} \leq C_{d,D} \| z \|_{H^{-1}(D)} \quad \text{a.s.}$$

and $S : H^{-1}(D) \to L^\infty(\Omega, \Sigma, \mathbb{P}; H_0^1(D))$.

Note: S with domain restricted to $L^2(D)$ is compact since $L^2(D) \subset\subset H^{-1}(D)$.

Uncertain Objective Functions

General Assumptions:

1. **Integrability:** $\widehat{J}(z) \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ for all $z \in Z$;
2. **Weak Lower Semicontinuity:** If $z_n \rightharpoonup z$ then

$$\liminf_{n\to\infty} \mathbb{E}[\vartheta \widehat{J}(z_n)] \geq \mathbb{E}[\vartheta \widehat{J}(z)]$$

for all $\vartheta \in (L^p(\Omega, \mathcal{F}, \mathbb{P}))^*$ satisfying $\vartheta \geq 0$ a.s.

Compare to *normal integrands*, i.e., the epigraph of $\widehat{J}$ is measurable and closed valued.

Uncertain Objective Functions

Separable Objective Functions: $J(u,z,\omega) = g(u,\omega) + \wp(z)$

1. **Carathéodory:** $g(\cdot,\omega)$ is continuous a.s. and $g(u,\cdot)$ is measurable $\forall\, u \in U$.
2. **Growth Condition:**
   If $q < \infty$, then $\exists\ 0 \le a \in L^p(\Omega,\mathcal{F},\mathbb{P})$ and $c > 0$ such that

   $$|g(u,\omega)| \le a(\omega) + c\|u\|_U^{q/p} \quad \forall\, u \in U \text{ a.s.}$$

   If $q = \infty$, then $\forall\, c > 0\ \ \exists\ \gamma_c \in L^p(\Omega,\mathcal{F},\mathbb{P})$ such that

   $$|J(u,\omega)| \le \gamma_c(\omega) \quad \text{a.s.} \quad \forall\, u \in U, \quad \|u\|_U \le c.$$

3. **Convexity:** $g(\cdot,\omega)$ is convex a.s. (optional)

Uncertain Objective Functions

The Separable Case

Superposition (Nemytskii) Operator:
$\mathcal{G} : L^q(\Omega, \mathcal{F}, \mathbb{P}; U) \to L^p(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathcal{G}(u) = g(u(\cdot), \cdot)$.

1. If g is Carathéodory and satisfies the growth condition, then $\mathcal{G} : L^q(\Omega, \mathcal{F}, \mathbb{P}; U) \to L^p(\Omega, \mathcal{F}, \mathbb{P})$ is continuous.
2. If, in addition, g is convex, then $\mathcal{G}$ is Gâteaux directionally differentiable.
3. If, in addition, g is locally Lipschitz, then $\mathcal{G}$ is Hadamard directionally differentiable.
4. If $g(\cdot, \omega)$ is continuously Fréchet differentiable for a.s. and there exists $\alpha > 0$ and $K \in L^s(\Omega, \mathcal{F}, \mathbb{P})$ with

$$s = \begin{cases} pq/(q-(1+\alpha)p) & \text{if } q > (1+\alpha)p \\ \infty & \text{if } q = (1+\alpha)p \end{cases}$$

such that

$$\|g_u(u,\omega) - g_u(v,\omega)\|_{U^*} \le K(\omega)\|u-v\|_U^\alpha \quad \text{a.s.}$$

Then $\mathcal{G}$ is Fréchet differentiable.

Example: Quadratic Objective Function

Let W be a real Hilbert space, $w \in W$ and $\mathbf{C} \in \mathcal{L}(U, W)$. Consider

$$J(u, z, \omega) = \frac{1}{2}\|\mathbf{C}u - w\|_W^2 + \frac{\gamma}{2}\|z\|_Z^2, \quad \gamma > 0.$$

J is separable with $g(u, \omega) = \frac{1}{2}\|\mathbf{C}u - w\|_W^2$.

1. **Carathéodory:** Satisfied since g has no dependence on ω.
2. **Growth Condition:** Satisfied (using Young's inequality) with

$$a = \|w\|_W^2 \quad \text{and} \quad c = \|\mathbf{C}\|_{\mathcal{L}(U,W)}^2.$$

3. **Convexity:** Clearly satisifed.
4. **Differentiability:** Satisfied with $K = \|\mathbf{C}\|_{\mathcal{L}(U,W)}^2$ and $\alpha = 1$.

Result: $\mathcal{G} : L^q(\Omega, \mathcal{F}, \mathbb{P}; U) \to L^p(\Omega, \mathcal{F}, \mathbb{P})$ is continuous and Fréchet differentiable as long as $q \geq 2p$.

The Functional $\mathcal{R}$

Assumptions & Existence of Minimizers

$\mathcal{R} : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R} \cup \{+\infty\}$

- $\mathcal{R}$ is convex and lower semicontinuous
- $\mathcal{R}$ satisfies $\mathcal{R}(C) = C$ for all constants C;
- $\mathcal{R}$ is monotonic, i.e., if $X \geq X'$ a.s., then $\mathcal{R}(X) \geq \mathcal{R}(X')$.

Existence: If Z_{ad} is convex, closed and bounded, then there exists a minimizer of $\mathcal{J}(z) = \mathcal{R}(\widehat{J}(z))$ in Z_{ad}.

Proof: Apply the direct method of the calculus of variations.

Note: The same result holds if $Z = Z_{\text{ad}}$ and $\widehat{J}(z)$ is a.s. coercive, i.e., $Z_{\text{ad}} = Z$ and $\widehat{J}(z)$ has the coercivity propery that $\exists\, r > 0$ and coercive $\varphi : Z \to \mathbb{R} \cup \{+\infty\}$, such that

$$\|z\|_Z \geq r \quad \Longrightarrow \quad \widehat{J}(z) \geq \varphi(z) \text{ a.s.}$$

Outline

Modeling Risk Preference

What is risk? *Possibility of loss or injury* (Merriam Webster)

…In our optimization problem, $J(S(z;\cdot),\cdot)$ is a **risk**!

We **cannot** directly minimize $J(S(z;\cdot),\cdot) + \wp(z) \in \mathcal{X} := L^p(\Omega, \mathcal{F}, \mathbb{P})$

…How should we quantify our risk?

- **Traditional Stochastic Programming:** Minimize *on average*

$$\mathcal{R}(F(z)) = \mathbb{E}[\mathcal{F}(z)].$$

- **Risk-Averse Stochastic Programming:** Model *risk preferences*

$$\mathcal{R}(F(z)) = \mathbb{E}[\mathcal{F}(z)] + c\mathbb{E}[(\mathcal{F}(z) - \mathbb{E}[\mathcal{F}(z)])_+^p]^{1/p}.$$

- **Probabilistic Optimization:** Minimize the *probability of loss*

$$\mathcal{R}(\mathcal{F}(z)) = \mathbb{P}(\mathcal{F}(z) > \tau).$$

- **Stochastic Orders:** Model risk preference with a *benchmark* Y

$$\mathbb{P}(X \leq x) \leq \mathbb{P}(Y \leq x) \quad \forall x \in \mathbb{R}.$$

Quantifying Risk & Controlling Uncertainty

- Reduce **variability** of optimized system:

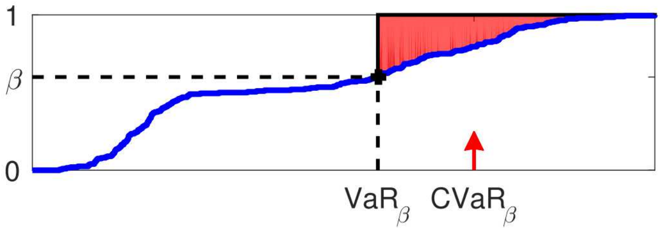
$$\mathbb{E}[(X - \mathbb{E}[X])^2] \qquad \text{or} \qquad \mathbb{E}[(X - \mathbb{E}[X])_+^p]^{1/p}$$

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

$$\mathbb{P}(X > t) \qquad \text{or} \qquad q_\beta(X) = \inf\{\, t \in \mathbb{R} \,:\, \mathbb{P}(X \le t) \ge \beta \,\}$$

- Minimize over **undesirable events**:

$$\mathrm{CVaR}_\beta(X) = \frac{1}{1-\beta}\int_\beta^1 F_X^{-1}(\alpha)\,\mathrm{d}\alpha \approx \mathbb{E}[X \,|\, X \ge q_\beta(X)]$$



Mitigating Uncertainty by Shaping Distributions

Law Invariance & Stochastic Dominance

Law Invariance:

- $\mathcal{R}$ is **law invariant** if

$$F_X(t) = F_{X'}(t) \ \ \forall t \in \mathbb{R} \quad \Longrightarrow \quad \mathcal{R}(X) = \mathcal{R}(X').$$

If $\mathcal{R}$ is law invariant, then it is a function of distributions.

Stochastic Dominance:

- X **dominates** X' with respect to the 1^{st} **stochastic order**, denoted $X \succeq_{(1)} X'$, if

$$F_X(t) \le F_{X'}(t) \ \ \forall t \in \mathbb{R}.$$

- X **dominates** X' with repsect to the 2^{nd} **stochastic order**, denoted $X \succeq_{(2)} X'$, if

$$\int_{-\infty}^{t} F_X(\eta)\,\mathrm{d}\eta \le \int_{-\infty}^{t} F_{X'}(\eta)\,\mathrm{d}\eta \quad \forall t \in \mathbb{R}$$
$$\iff \quad \mathbb{E}[(t-X)_+] \le \mathbb{E}[(t-X')_+] \ \ \forall t \in \mathbb{R}.$$

Here, $(x)_+ = \max\{0, x\}$.

Consequences: Suppose $\mathcal{R}$ is law invariant:

- If $X \ge X'$ a.s. implies $\mathcal{R}(X) \ge \mathcal{R}(X')$, then $X \succeq_{(1)} X'$ implies $\mathcal{R}(X) \ge \mathcal{R}(X')$;
- If $\mathcal{R}$ is lsc and convex, then $-X' \succeq_{(2)} -X$ implies $\mathcal{R}(X) \ge \mathcal{R}(X')$.
- Law invariant $\mathcal{R}$ prefer dominated random variables!

Mean-Plus-Variance Risk

Markowitz, Portfolio Selection, 1952

A common risk functional in engineering application is

$$\mathcal{R}(X) = \mathbb{E}[X] + c\mathbb{V}[X] \quad \text{for } c > 0.$$

Downsides:

- $\mathcal{R}$ penalizes variation **below** the mean.
- $\mathcal{R}$ is **not** monotonic.

Example: Shapiro, Dentcheva, Ruszczynski (2014)

Suppose $\Omega = \{\omega_1, \omega_2\}$ with associated probabilities $\mathfrak{p} \in (0,1)$ and $(1-\mathfrak{p})$. Consider the stochastic program

$$\underset{z_1,\,z_2}{\text{Minimize}}\ \mathcal{R}(-\zeta_1 z_1 - \zeta_2 z_2) \quad \text{subject to} \quad z_1 + z_2 = 1 \ \text{ and } \ z_1,\, z_2 \geq 0$$

where $\zeta_1,\ \zeta_2 : \Omega \to \mathbb{R}$ are

$$\zeta_1(\omega_1) = a > 0, \quad \zeta_1(\omega_2) = 0, \quad \text{and} \quad \zeta_2(\omega_1) = \zeta_2(\omega_2) = 0.$$

If $\mathfrak{p} \leq 1 - (ca)^{-1}$, then $\mathcal{R}(-\zeta_1) = -\mathfrak{p}a + ca^2\mathfrak{p}(1-\mathfrak{p}) > \mathcal{R}(-\zeta_2) = 0$ eventhough $-\zeta_1 \leq -\zeta_2$ for all $\omega \in \Omega$.

Coherent Risk Measures

$\mathcal{R} : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R} \cup \{\infty\}$ is **coherent** if

(R1) **Convexity:** For all $X, X' \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ and for all $0 \le t \le 1$,

$$\mathcal{R}(tX + (1-t)X') \le t\mathcal{R}(X) + (1-t)\mathcal{R}(X')$$

(R2) **Monotonicity:** For any $X, X' \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ satifying

$$X \ge X' \text{ a.s.} \qquad \Longrightarrow \qquad \mathcal{R}(X) \ge \mathcal{R}(X')$$

(R3) **Translation Equivariance:** For all $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $t \in \mathbb{R}$,

$$\mathcal{R}(X + t) = \mathcal{R}(X) + t$$

(R4) **Positive Homogeneity:** For all $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $t \ge 0$,

$$\mathcal{R}(tX) = t\mathcal{R}(X)$$

Ph. Artzner, F. Delbaen, J.-M. Eber & D. Heath, *Coherent measures of risk.* Math. Finance, 1999.

Biconjugate Representation:

- $\mathcal{R}$ is proper, **convex** and lsc $\iff$

$$\mathcal{R}(X) = \sup\left\{\mathbb{E}[\vartheta X] - \mathcal{R}^*(\vartheta) \mid \vartheta \in \mathrm{dom}(\mathcal{R}^*)\right\}.$$

- $\mathcal{R}$ is **translation equivariant** and **monotonic** $\iff$

$$\mathrm{dom}(\mathcal{R}^*) \subseteq \{\vartheta \in (L^p(\Omega,\mathcal{F},\mathbb{P}))^* \mid \mathbb{E}[\vartheta] = 1,\ \vartheta \geq 0 \text{ a.s.}\}$$

- $\mathcal{R}$ is **positive homogeneous** $\iff$

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

Example (Conditional Value-at-Risk (CVaR)): $\mathcal{R}(X) = \frac{1}{1-\beta}\int_\beta^1 q_X(\beta)\,\mathrm{d}\beta$

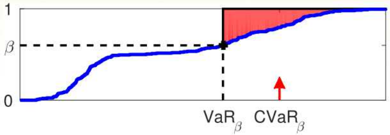
$$\mathrm{dom}(\mathcal{R}^*) = \left\{\vartheta \in (L^p(\Omega,\mathcal{F},\mathbb{P}))^* \mid \mathbb{E}[\vartheta] = 1,\ 0 \leq \vartheta \leq \frac{1}{1-\beta} \text{ a.s.}\right\}.$$

Differentiability: If $\mathcal{R} : L^p(\Omega,\mathcal{F},\mathbb{P}) \to \mathbb{R}$ is **coherent**, then $\mathcal{R}$ is **Fréchet differentiable** $\iff$ $\exists\,\vartheta \in (L^p(\Omega,\mathcal{F},\mathbb{P}))^*$ with $\vartheta \geq 0$ a.s., $\mathbb{E}[\vartheta] = 1$, and $\mathcal{R}(X) = \mathbb{E}[\vartheta X]$ for all $X \in L^p(\Omega,\mathcal{F},\mathbb{P})$.

CVaR and Kusuoka Representation

Let $F_X(x) = \mathbb{P}(X \le x)$, then CVaR is

$$\mathrm{CVaR}_\beta(X) := \frac{1}{1-\beta}\int_\beta^1 F_X^{-1}(\alpha)\,\mathrm{d}\alpha$$



In fact, all **law-invariant coherent** risk measures have the representation

$$\mathcal{R}(X) = \sup_{\mu\in\mathfrak{M}} \int_0^1 \mathrm{CVaR}_\beta(X)\,\mathrm{d}\mu(\beta)$$

where $\mathfrak{M}$ is a set of **probability measures** on $[0,1]$.

Spectral Risk Measures: Given a probability measure ν on $[0,1]$,

$$\begin{aligned}\mathcal{R}(X) &= \int_0^1 \mathrm{CVaR}_\beta(X)\,\mathrm{d}\nu(\beta)\\ &= \int_0^1 h(\beta)F_X^{-1}(\beta)\,\mathrm{d}\beta \quad \text{where} \quad h(\beta) := \int_0^\beta \frac{1}{1-\alpha}\,\mathrm{d}\nu(\alpha)\end{aligned}$$

S. Kusuoka, *On law-invariant coherent risk measures*, Advances in Math. Econ., 2001.

Risk Measure Examples

Risk Neutral:

$$\mathcal{R}(X) = \mathbb{E}[X]$$

is **law invariant** and **coherent**.

Mean-Plus-Deviation:

$$\mathcal{R}(X) = \mathbb{E}[X] + c\mathbb{E}[|X - \mathbb{E}[X]|^p]^{1/p}, \quad c > 0$$

is **law invariant** and satisfies (R1), (R3) and (R4), but **not** (R2).

Mean-Plus-Upper-Semideviation:

$$\mathcal{R}(X) = \mathbb{E}[X] + c\mathbb{E}[(X - \mathbb{E}[X])_+^p]^{1/p}, \quad c \in [0, 1]$$

is **law invariant** and **coherent**.

Conditional Value-at-Risk:

$$\mathcal{R}(X) = \frac{1}{1-\beta}\int_\beta^1 F_X^{-1}(\eta)\,\mathrm{d}\eta = \inf_{t\in\mathbb{R}}\left\{t + \frac{1}{1-\beta}\mathbb{E}[(X-t)_+]\right\}, \quad 0 \le \beta < 1$$

is **law invariant** and **coherent**.

Entropic Risk:

$$\mathcal{R}(X) = \lambda^{-1}\ln\mathbb{E}[\exp(\lambda X)], \ \lambda > 0$$

is **law invariant** and satisfies (R1), (R2) and (R3), but **not** (R4).

More Measures of Risk

One can quantify risk using the optimized certainty equivalent risk measure

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

where $v : \mathbb{R} \to \mathbb{R}$ is a convex *regret* function that satisfies

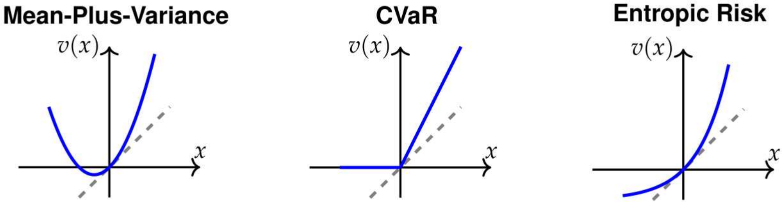
$$v(0) = 0, \quad v(x) > x \quad \forall\, x \neq 0$$

Relation to Utility: $u(x) = -v(-x)$ is a *utility* function

Properties: $\mathcal{R}$ is convex and translation equivariant

$\mathcal{R}$ is **positive homogeneous** $\iff$ v is piecewise linear with kink at 0

$\mathcal{R}$ is **monotonic** $\iff$ v is nondecreasing

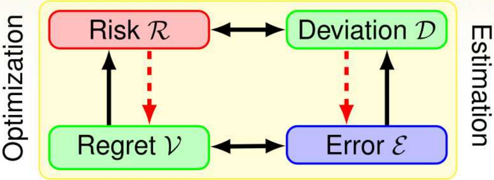


A. Ben Tal & M. Teboulle, *An old-new concept of convex risk measures: The optimized certainty equivalents*, Math. Finance, 2007.

The Risk Quadrangle

$$\mathcal{R}(X) = \mathbb{E}[X] + \mathcal{D}(X) = \min_t \{t + \mathcal{V}(X - t)\}$$

$$\mathcal{V}(X) = \mathbb{E}[X] + \mathcal{E}(X)$$



$$\mathcal{D}(X) = \mathcal{R}(X) - \mathbb{E}[X] = \min_t \mathcal{E}(X - t)$$

$$\mathcal{E}(X) = \mathcal{V}(X) - \mathbb{E}[X]$$

- $\mathcal{R}$ quantifies **hazard** — Used in optimization as objective function or constraint
- $\mathcal{E}$ quantifies **nonzeroness** — Used in regression analysis, e.g., polynomial chaos
- $\mathcal{V}$ quantifies **displeasure** for postive values — Used to define risk via *disutility*
- $\mathcal{D}$ quantifies **nonconstancy** — Used to define risk via *variability*

Quantile Quadrangle: $0 < \alpha < 1$

$$\mathcal{R}(X) = \mathrm{CVaR}_\alpha(X) \qquad \mathcal{D}(X) = \mathrm{CVaR}_\alpha(X - \mathbb{E}[X])$$

$$\mathcal{V}(X) = \tfrac{1}{1-\alpha}\mathbb{E}[X_+] \qquad \mathcal{E}(X) = \mathbb{E}[\tfrac{\alpha}{1-\alpha}X_+ + X_-]$$

$$\mathcal{S}(X) = q_\alpha(X)$$

Safety Margins Quadrangle: $c > 0$

$$\mathcal{R}(X) = \mathbb{E}[X] + c\sigma(X) \qquad \mathcal{D}(X) = c\sigma(X)$$

$$\mathcal{V}(X) = \mathbb{E}[X] + c\|X\|_2 \qquad \mathcal{E}(X) = c\|X\|_2$$

$$\mathcal{S}(X) = \mathbb{E}[X]$$

R. T. Rockafellar & S. Uryasev, *The fundamental risk quadrangle in risk management, optimization, and statistical estimation*, Surveys in OR & Managment Science, 2013.

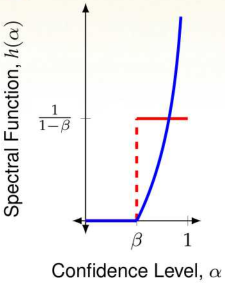
Superquantile Quadrangle

Choosing the **uniform** probability measure on $[\beta, 1]$,

$$\nu(S) = \frac{1}{1-\beta} \int_S \mathbb{1}_{[\beta,1]}(\alpha)\,\mathrm{d}\alpha,$$

produces the **second-order** CVaR

$$\mathcal{R}(X) = \frac{1}{1-\beta} \int_\beta^1 \mathrm{CVaR}_\alpha(X)\,\mathrm{d}\alpha$$



Second-order CVaR is a product of the **risk quadrangle**:

$$\mathcal{R}(X) = \frac{1}{1-\beta} \int_\beta^1 \mathrm{CVaR}_\alpha(X)\,\mathrm{d}\alpha \qquad \mathcal{D}(X) = \frac{1}{1-\beta} \int_\beta^1 \mathrm{CVaR}_\alpha(X - \mathbb{E}[X])\,\mathrm{d}\alpha$$

$$\mathcal{V}(X) = \frac{1}{1-\beta} \int_0^1 (\mathrm{CVaR}_\alpha(X))_+\,\mathrm{d}\alpha \qquad \mathcal{E}(X) = \frac{1}{1-\beta} \int_0^1 (\mathrm{CVaR}_\alpha(X))_+\,\mathrm{d}\alpha - \mathbb{E}[X]$$

$$\mathcal{S}(X) = \mathrm{CVaR}_\beta(X)$$

R. T. Rockafellar & J. O. Royset, *Random variables, monotone relations, and convex analysis*, Math. Programming, 2014.

Example — CVaR

Optimal Control of 1D Elliptic Equation

Let $\gamma = 10$, $D = (-1, 1)$, and $w \equiv 1$ and consider

$$\underset{z \in L^2(-1,1)}{\text{minimize}} \quad J(z) = \frac{1}{2}\text{CVaR}_\beta \left[\int_{-1}^{1} (S(z)(\cdot, x) - 1)^2 \, \mathrm{d}x \right] + \frac{\gamma}{2} \int_{-1}^{1} z(x)^2 \, \mathrm{d}x$$

where $S(z) = u \in L^2(\Omega, \mathcal{F}, \mathbb{P}; H_0^1(-1,1))$ solves the weak form of

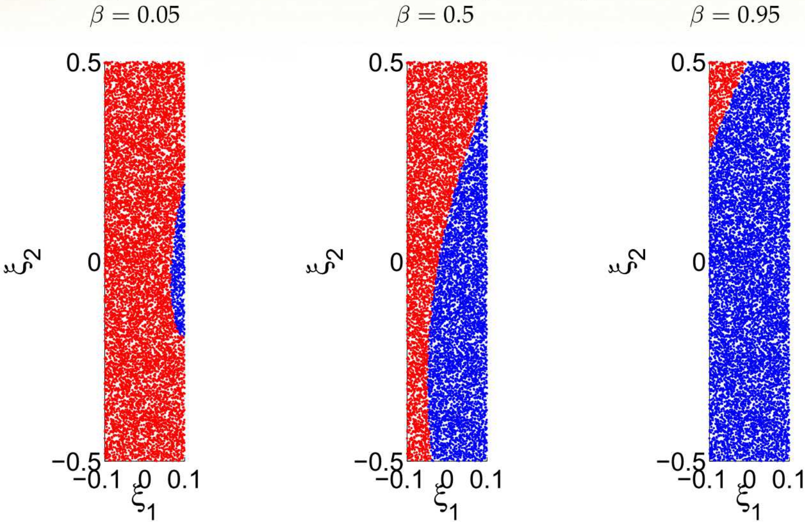
$$\begin{aligned} -\partial_x \left(\epsilon(\omega, x) \partial_x u(\omega, x) \right) &= f(\omega, x) + z(x) && x \in D, \text{ a.s.}, \\ u(\omega, -1) = 0, \quad u(\omega, 1) &= 0 && \text{a.s.} \end{aligned}$$

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

$$\epsilon(\omega, x) = 0.1 \cdot \mathbb{1}_{(-1,\omega_1)} + 10 \cdot \mathbb{1}_{(\omega_1,1)}, \quad \text{and} \quad f(\omega, x) = \exp(-(x - \omega_2)^2).$$

Example — CVaR

Sample Approximation: Monte Carlo with 10,000 samples.



$\vartheta^* = 0$ and $\vartheta^* = (1-\beta)^{-1}$

Outline

Motivating Applications

General Problem Formulation

Quantifying Risk

Probabilistic Optimization

What if our uncertainty is uncertain?

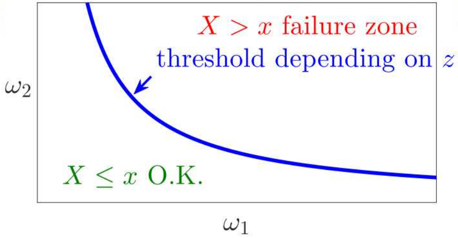
Computational Solution Methods

References

Probabilistic Hazard

Standard Engineering Prospective

$X = \widehat{J}(z) =$ "cost" signaling "danger"



Probability of failure: $\mathcal{R}(X) = p_x(X) = \mathbb{P}(X > x)$

- How to compute or at least estimate?
- How to cope with control variables z in optimization? Both $p_x(X)$ and the threshold change with z!

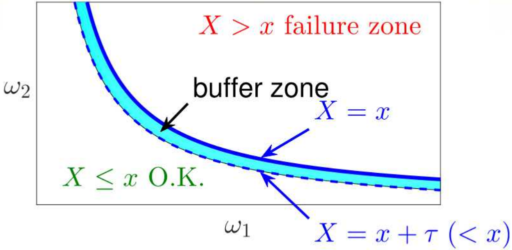
Troubles with this concept:

- Poor mathematical behavior is a serious handicap.
- Failure probability ignores the **degree** of failure.

Buffered Probabilities

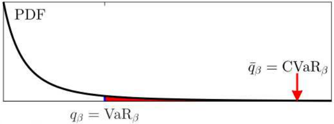
Rockafellar & Royset (2013), Mafusalov & Uryasev (2014), Norton & Uryasev (2014)

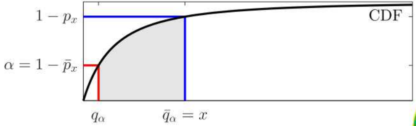
Utilizing **CVaR** in place of **quantile** in reliability



Buffered probability of failure: $\mathcal{R}(X) = \bar{p}_x(X) = \mathbb{P}(X > \tau(x))$
where $\tau(x)$ is determined by $\mathrm{CVaR}_{(1-\bar{p}_x(X))}(X) = \mathbb{E}[\,X \,|\, X > \tau(x)\,] = x$.

$$\mathrm{bPOE}_x[X] = 1 - \alpha \quad \text{where } \alpha \text{ solves} \quad \mathrm{CVaR}_\alpha[X] = x.$$





Buffered Probability Properties

- Optimization representation:

$$\mathrm{bPOE}_x[X] = \min_{t \geq 0} \mathbb{E}[(t(X - x) + 1)_+]$$

- Takes into account **values of outcomes in the distribution tail**
- Closed, quasi-convex and monotonic in random variable X
- Lowest quasi-convex (in X) upper bound of POE
- Continuous with respect to threshold $x \in [\mathbb{E}[X], \operatorname{ess\,sup} X)$
- Easy to manage (optimize with convex and linear programming)
- $$\mathrm{CVaR}_\alpha[X] \leq x \quad \Longleftrightarrow \quad \mathrm{bPOE}_x[X] \leq 1 - \alpha$$

Objective function in optimization representation is **nonsmooth**!

Question: Is it possible to account for higher-order tail moments?

Higher-Moment Coherent Risk Measures

Higher-Moment Coherent Risk (HMCR) measures with $p \geq 1$ and $\beta \in [0,1)$

$$\mathrm{HMCR}_{p,\beta}[X] = \inf_{t \in \mathbb{R}} \left\{ t + \frac{1}{1-\beta} \mathbb{E}[(X-t)_+^p]^{1/p} \right\}$$

1. As the name suggests, $\mathrm{HMCR}_{p,\beta}$ is coherent and law invariant
2. When $p = 1$, we have that $\mathrm{HMCR}_{1,\beta}[X] = \mathrm{CVaR}_\beta[X]$
3. $\mathrm{HMCR}_{p,\beta}$ is generated from the risk quadrangle with regret measure

$$\mathcal{V}(X) = \frac{1}{1-\beta} \mathbb{E}[(X)_+^p]^{1/p}$$

Properties of HMCR: Suppose X is not degenerate (constant)

1. $p \mapsto \mathrm{HMCR}_{p,\beta}[X]$ is nondecreasing
2. $\beta \mapsto \mathrm{HMCR}_{p,\beta}[X]$ is nondecreasing and continuous
3. In fact, $\beta \mapsto \mathrm{HMCR}_{p,\beta}[X]$ is strictly increasing on $[0, 1-\pi_X)$ with

$$\pi_X = \mathrm{prob}(X = \operatorname{ess\,sup} X)$$

4. $\mathrm{HMCR}_{p,0}[X] = \mathbb{E}[X]$ and $\mathrm{HMCR}_{p,1}[X] = \operatorname{ess\,sup} X$

$\beta \mapsto \mathrm{HMCR}_{p,\beta}[X]$ has a nondecreasing and continuous inverse!

P. A. Krokhmal, *Higher moment coherent risk measures*, Quantitative Finance, 2007.

Higher-Moment bPOE Properties

Kouri (2018)

- Optimization representation:

$$\mathrm{bPOE}_{p,x}[X] = \min_{t\geq 0} \mathbb{E}[(t(X-x)+1)_+^p]^{1/p}$$

- Takes into account **moments** of outcomes in the distribution tail
- Closed, quasi-convex and monotonic in random variable X
- Continuous with respect to threshold $x \in [\mathbb{E}[X], \operatorname{ess\,sup} X)$
- Objective function in optimization representation is **smooth** in X
- $$\mathrm{HMCR}_{p,\alpha}[X] \leq x \quad \Longleftrightarrow \quad \mathrm{bPOE}_{p,x}[X] \leq 1-\alpha$$
- $$\mathrm{bPOE}_{x}[X] \leq (\mathrm{bPOE}_{2,x}[X])^2 \leq \ldots \leq (\mathrm{bPOE}_{p,x}[X])^p$$

Example: Second-Moment Buffered Probability

Suppose $X \sim N(0,1)$ with cdf Φ and pdf ϕ. Let $x \geq 0$ then

$$Z := (t(X - x) + 1) \sim N(1 - tx, t) \quad \forall t > 0$$

Therefore, the buffered probability of X exceeding x is

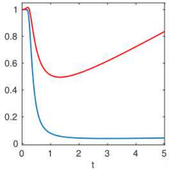
$$\text{bPOE}_x[X] = \min_{t \geq 0}\{1 - \Phi(x - 1/t) + t\phi(x - 1/t)\}$$

and the second order buffered probability of X exceeding x is

$$(\text{bPOE}_{2,x}[X])^2 = \min_{t \geq 0}\{(1 + t^2)(1 - \Phi(x - 1/t)) + (t^2 x + t)\phi(x - 1/t)\}$$

x	$\text{POE}_x[X]$	$\text{bPOE}_x[X]$	$(\text{bPOE}_{2,x}[X])^2$	$\text{bPOE}_{2,x}[X]$
0	0.5	1	1	1
1	0.15866	0.89894	1	1
2	0.02275	0.32584	0.99608	0.99804
3	0.00135	0.03802	0.49553	0.70394
4	0.00003	0.00150	0.12966	0.36008
5	2.87e-7	0.00002	0.01890	0.13746
6	9.87e-10	3.84e-7	0.00158	0.03973

$x = 3$



Order 1 and Order 2

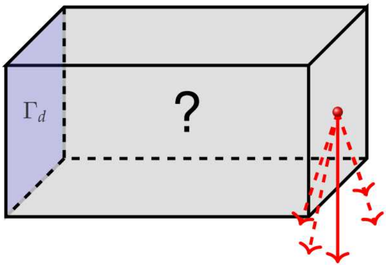
3D Topology Optimization with Buffered Probability

Given compliance tolerance c_0, probability $p_0 \in (0,1)$, order $q \geq 1$,

$$\min_{0 \leq z \leq 1} \int_D z \, \mathrm{d}x =: \mathrm{vol}(z) \quad \text{subject to} \quad \mathrm{bPOE}_{q,c_0}\left(\int_D F \cdot S(z)\mathrm{d}x\right) \leq 1 - p_0$$

where $S(z) = u$ solves the **linear elasticity equations**

$$\begin{aligned} -\nabla \cdot (\mathbf{E}(z) : \varepsilon u) &= F, && \text{in } D \\ \varepsilon u &= \tfrac{1}{2}(\nabla u + \nabla u^\top), && \text{in } D \\ u &= 0, && \text{on } \Gamma_D \\ \varepsilon u : n &= 0, && \text{on } \partial D \setminus \Gamma_D \end{aligned}$$



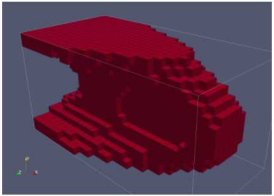
Numerical Results

Spatial Discretization: Q1 FEM on a uniform $32 \times 16 \times 16$ mesh

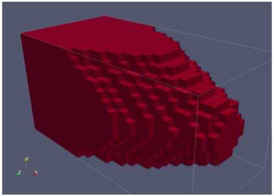
Stochastic Discretization: $Q = 120$ Monte Carlo samples

Problem Data: $p_0 = 0.75$ and $c_0 = 2\mathbb{E}\left[\int_D F \cdot S(\mathbf{1})\mathrm{d}x\right]$

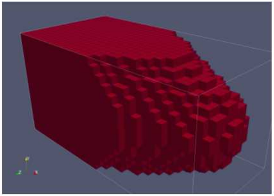
Mean Value



Risk Neutral



bPOE



	MV	RN	bPOE
Volume Fraction	49.061%	47.634%	67.204%

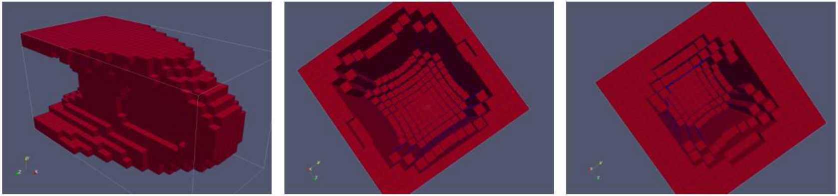
Numerical Results

Spatial Discretization: Q1 FEM on a uniform $32 \times 16 \times 16$ mesh

Stochastic Discretization: $Q = 120$ Monte Carlo samples

Problem Data: $p_0 = 0.75$ and $c_0 = 2\mathbb{E}\left[\int_D F \cdot S(\mathbf{1})\mathrm{d}x\right]$

Mean Value **Risk Neutral** **bPOE**



Topology changes from beam to shell!

	MV	RN	bPOE
Volume Fraction	49.061%	47.634%	67.204%

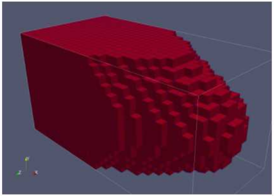
Numerical Results

Spatial Discretization: Q1 FEM on a uniform $32 \times 16 \times 16$ mesh

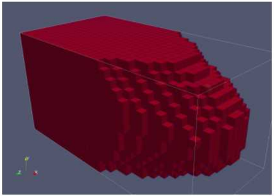
Stochastic Discretization: $Q = 120$ Monte Carlo samples

Problem Data: $p_0 = 0.75$ and $c_0 = 2\mathbb{E}\left[\int_D F \cdot S(\mathbf{1})\mathrm{d}x\right]$

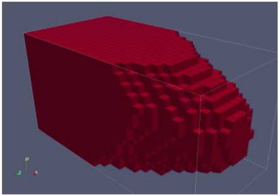
bPOE_{c_0}



bPOE_{2,c_0}



bPOE_{3,c_0}



Order	1	2	3
Volume Fraction	67.204%	77.369%	80.075%

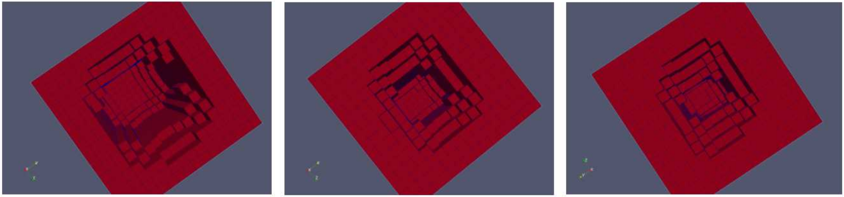
Numerical Results

Spatial Discretization: Q1 FEM on a uniform $32 \times 16 \times 16$ mesh

Stochastic Discretization: $Q = 120$ Monte Carlo samples

Problem Data: $p_0 = 0.75$ and $c_0 = 2\mathbb{E}\left[\int_D F \cdot S(\mathbf{1})\mathrm{d}x\right]$

bPOE_{c_0} $\qquad$ bPOE_{2,c_0} $\qquad$ bPOE_{3,c_0}



Topology changes from beam to shell!

Order	1	2	3
Volume Fraction	67.204%	77.369%	80.075%

Outline

Motivating Applications

General Problem Formulation

Quantifying Risk

Probabilistic Optimization

What if our uncertainty is uncertain?

Computational Solution Methods

References

What if our uncertainty is uncertain?

Distributionally Robust Stochastic Programming

$(\Omega, \mathcal{F})$ is a measurable space and prob. measure is *unknown*.
Consider

$$\min_{z \in Z_{\mathrm{ad}}} \mathcal{R}(\widehat{J}(z)) = \sup_{P \in \mathfrak{A}} \mathbb{E}_P[\widehat{J}(z)].$$

Ambiguity Set: $\mathfrak{A} \subset \{P : \mathcal{F} \to [0,1] \mid P(\Omega) = 1\}$ defined by data.
For example:

- **Moment Matching:** Given generalized moment data $m_1, \ldots, m_N$,

$$\mathfrak{A} = \{P : \mathcal{F} \to [0,1] \mid P(\Omega) = 1,\, \mathbb{E}_P[\psi_i] = m_i,\, i = 1, \ldots, N\}.$$

- **Φ-Divergence (e.g., Kullback-Leibler, χ^2, TV, Hellinger, . . .):** Given a nominal P_0 and $\epsilon > 0$,

$$\mathfrak{A} = \{P : \mathcal{F} \to [0,1] \mid P(\Omega) = 1,\, D_\Phi(P, P_0) \le \epsilon\}.$$

- **Wasserstein Distance:** Given a nominal P_0 and $\epsilon > 0$,

$$\mathfrak{A} = \left\{P : \mathcal{F} \to [0,1] \mid P(\Omega) = 1,\, \sup_{f \in \mathcal{L}} \int_\Omega f(\omega)\, \mathrm{d}(P - P_0)(\omega) \le \epsilon\right\}.$$

Example: Moment Matching

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

$$\mathfrak{A} = \left\{ P : \mathcal{F} \to [0,1] \mid P(\Omega) = 1, \quad \begin{matrix} \mathbb{E}_P[\psi_i] = m_i, \; i = 1, \ldots, N_e \\ \mathbb{E}_P[\psi_i] \le m_i, \; i = N_e + 1, \ldots, N \end{matrix} \right\}.$$

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

$$\mathcal{R}(\widehat{J}(z)) = \sup_{P \in \mathfrak{A}} \mathbb{E}_P[\widehat{J}(z)] = \sum_{i=1}^{N+1} p_i J([S(z)](\omega_i), z, \omega_i)$$

W. W. Rogosinski, *Moments of non-negative mass*, Proceedings of the Royal Society of London: Series A, Math. and Phys. Sciences, 1958.

Example: Φ-Divergence

Supppose

(i) A **nominal** probability measure P_0 is given,

(ii) The random variable $X \in L^p(\Omega, \mathcal{F}, P_0)$, and

(iii) $\Phi : \mathbb{R} \to [0, \infty]$ is convex lower semicontinuous satisfying

$$\Phi(1) = 0 \quad \text{and} \quad \Phi(x) = \infty \quad \forall\, x < 0.$$

Define, for fixed $\epsilon > 0$,

$$\mathfrak{A} = \{\vartheta \in (L^p(\Omega, \mathcal{F}, P_0))^* \;|\; \mathbb{E}_{P_0}[\vartheta] = 1,\; \vartheta \geq 0,\; \mathbb{E}_{P_0}[\Phi(\vartheta)] \leq \epsilon\}\,.$$

Then $\quad \mathcal{R}(X) = \sup_{\vartheta \in \mathfrak{A}} \mathbb{E}_{P_0}[\vartheta X] = \inf_{\lambda \geq 0,\, \mu} \{\lambda\epsilon + \mu + \mathbb{E}_{P_0}[(\lambda\Phi)^*(X - \mu)]\}$

is a **law-invariant coherent** risk measure!

Example (Kullback-Leibler Divergence): $\Phi(x) = x\ln(x) - x + 1,\, x \geq 0$

$$\mathcal{R}(X) = \inf_{\lambda > 0} \left\{\lambda c + \lambda \ln \mathbb{E}_{P_0}\left[e^{X/\lambda}\right]\right\}.$$

A. Ben Tal & M. Teboulle, *Penalty functions and duality in stochastic programming via phi-divergence functionals*, Mathematics of Operations Research, 1987.

Robust Probabilitistic Optimization

Shapiro, Mafusalov, Uryasev, Kouri (2018)

When $\mathbb{P}$ is unknown, we can similarly *robustify* the POE and bPOE.

Probability: In general, we have that

$$\mathrm{POE}_x^\star(X) = \sup_{P\in\mathfrak{A}} P(X > x) = \sup_{P\in\mathfrak{A}} \mathbb{E}_P[\mathbb{1}_A] = \mathcal{R}(\mathbb{1}_A)$$

where $A = \{\omega \in \Omega \,|\, X(\omega) > x\}$ and $\mathcal{R}$ is a **coherent** risk measure!

Buffered Probability: Under mild regularity conditions, we have

$$\begin{aligned}\mathrm{bPOE}_x^\star(X) = \sup_{P\in\mathfrak{A}} \min_{t\geq 0} \mathbb{E}_P[(t(X-x)+1)_+] &= \min_{t\geq 0} \sup_{P\in\mathfrak{A}} \mathbb{E}_P[(t(X-x)+1)_+] \\ &= \min_{t\geq 0} \mathcal{R}((t(X-x)+1)_+)\end{aligned}$$

where $\mathcal{R}$ is a **coherent** risk measure! For Φ-divergence ambiguity,

$$\mathrm{bPOE}_x^\star(X) = \min_{t\geq 0,\lambda\geq 0,\mu} \{\lambda\epsilon + \mu + \mathbb{E}_{P_0}[(\lambda\Phi)^*((t(X-x)+1)_+ - \mu)]\}\,.$$

Distributionally Robust Contaminant Mitigation

Problem Description

Model contaminant spread by advection-diffusion on $D = (0,1)^2$.
Determine controls that mitigate the contaminant

$$\min_z \mathcal{R}\left(\frac{\kappa_s}{2}\int_D S(z)^2\,\mathrm{d}x\right) + \wp(z) \quad \text{subject to} \quad 0 \le z \le 1$$

where $S(z) = u : \Omega \to H^1(D)$ solves

$$\begin{aligned} -\nabla\cdot(\epsilon(\omega)\nabla u(\omega)) + V(\omega)\cdot\nabla u(\omega) &= f(\omega) - Bz, && \text{in } D, \text{ a.s.} \\ u(\omega) &= 0, && \text{on } \Gamma_d, \text{ a.s.} \\ \epsilon(\omega)\nabla u(\omega)\cdot n &= 0, && \text{on } \partial D \setminus \Gamma_d, \text{ a.s.,} \end{aligned}$$

$$Bz = \sum_{k=1}^{9} z_k \exp\left(\frac{-\|x-p_k\|_2^2}{2\sigma^2}\right) \quad \text{and} \quad \wp(z) = \kappa_c\|z\|_1 = \kappa_c\sum_{k=1}^{9} z_k.$$

Control	1	2	3	4	5	6	7	8	9
x_1	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75
x_2	0.25	0.25	0.25	0.50	0.50	0.50	0.75	0.75	0.75

Total of 37 random variables.

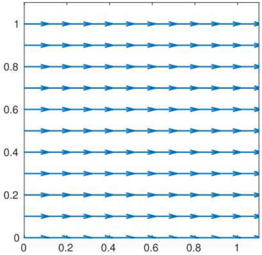
Risk-Averse Contaminant Mitigation

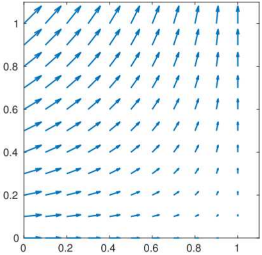
Nominal Distribution: $\xi_k(\omega) \sim U(-1,1)$ with $k = 1, \ldots, 37$

Diffusivity:

$$\log(c\,\epsilon(\omega,x) - 0.5) = 1 + \xi_1(\omega)\left(\frac{\sqrt{\pi}L_c}{2}\right)^{1/2} + \sum_{n=2}^{10} \zeta_k \phi_k(x) \xi_k(\omega)$$

Advection:





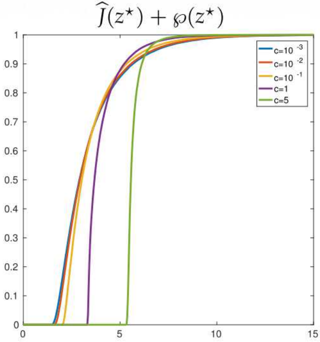
Source:

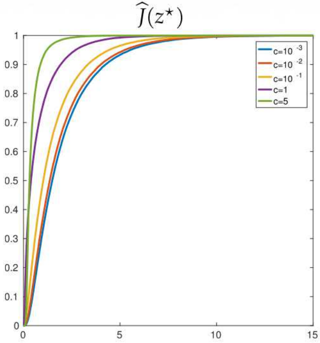
$$f(\omega,x) = \sum_{k=1}^{5} \xi_{8+5k}(\omega) \exp\left(\frac{-(x_1 - \xi_{9+5k}(\omega))^2}{2\xi_{10+5k}(\omega)^2}\right) \exp\left(\frac{-(x_2 - \xi_{11+5k}(\omega))^2}{2\xi_{12+5k}(\omega)^2}\right).$$

Numerical Results

DRO with KL-Divergence Ambiguity

$$\mathcal{R}(X) = \inf_{\lambda>0} \left\{ \lambda c + \lambda \ln \mathbb{E}\left[e^{X/\lambda} \right] \right\}$$





$\log_{10}(c)$	1	2	3	4	5	6	7	8	9	obj
10^{-3}	—	0.410	—	—	1.000	—	—	—	—	3.465
10^{-2}	—	0.560	—	—	1.000	—	—	—	—	3.637
10^{-1}	—	1.000	—	—	1.000	—	—	—	—	4.186
1	—	1.000	—	0.580	1.000	0.709	—	—	—	5.939
5	1.000	1.000	0.249	1.000	1.000	1.000	—	—	—	8.124

Outline

Motivating Applications

General Problem Formulation

Quantifying Risk

Probabilistic Optimization

What if our uncertainty is uncertain?

Computational Solution Methods

References

Methods for Stochastic Optimization

1. **Stochastic Approximation (SA):** Stochastic subgradient descent only requires a single sample at every iteration.
2. **Progressive Hedging:** Decoupled deterministic optimization via alternating directions method of multipliers (ADMM).
3. **Sample Average Approximation (SAA):** (Quasi) Monte Carlo approximation of expected value.
4. **Adaptive Stochastic Collocation:** Deterministic quadrature approximation of expected value. Adaptivity using trust regions.

Note: The convergence of SA and SAA is **probabilistic**!

Note: Risk measures and probabilistic functionals are often nonsmooth $\Longrightarrow$ polynomial approximation and derivative-based optimization may not apply.

The Finite Noise Assumption

Suppose there exists a random vector $\xi : \Omega \to \Xi \subseteq \mathbb{R}^M$ and functions $\bar{J} : U \times Z \times \Xi \to \mathbb{R}$ and $\bar{c} : U \times Z \times \Xi \to Y$ such that

$$J(u, z, \omega) = \bar{J}(u, z, \xi(\omega)) \quad \text{and} \quad c(u, z, \omega) = \bar{c}(u, z, \xi(\omega)).$$

Moreover, assume the probability law $\mathbb{P} \circ \xi^{-1}$ has Lebesgue density $\rho : \Xi \to \mathbb{R}$, i.e., $\mathrm{d}\mathbb{P} \circ \xi^{-1} = \rho \, \mathrm{d}\xi$.

This permits the change of variables from $\omega \in \Omega$ to $\xi \in \Xi$.
Analysis now performed in weighted Lebesgue space

$$L^p_\rho(\Xi) = \left\{ v : \Xi \to \mathbb{R} \;\middle|\; \int_\Xi |v(\xi)|^p \rho(\xi) \, \mathrm{d}\xi < \infty \right\}.$$

$L^\infty_\rho(\Xi)$ and $L^p_\rho(\Xi; W)$ are similarly defined.

Independence: For adaptive stochastic collocation, we will assume that the components of ξ are independent and

$$\Xi = [a_1, b_1] \times \cdots \times [a_M, b_M] \quad \text{and} \quad \rho = \rho_1 \otimes \cdots \otimes \rho_M.$$

Stochastic Approximation

Set $\widehat{J}(z) = \bar{J}(u(z), z, \cdot)$. Let Z be Hilbert and Z_{ad} be closed convex. Given $z_k \in Z_{\text{ad}}$ and $G(z_k, \xi) = G_k(\xi)$ such that $\mathbb{E}[G_k(\xi)] \in \partial \mathbb{E}[\widehat{J}(z_k)]$, the SA iteration is

$$z_{k+1} = \Pi_{Z_{\text{ad}}}(z_k - \mu_k G_k(\xi_k)), \quad \mu_k > 0,$$

where ξ_k for $k = 1, \ldots$ is an iid sequence of realizations and

$$\Pi_{Z_{\text{ad}}}(z) = \operatorname*{arg\,min}_{\zeta \in Z_{\text{ad}}} \|z - \zeta\|_Z.$$

Note: $\Pi_{Z_{\text{ad}}}$ is (firmly) nonexpansive.

Note: For PDE-constrained optimization, SA requires a single **deterministic** state and adjoint solve per iteration!

Must solve:

$$\bar{c}(u, z_k, \xi_k) = 0 \quad \text{and} \quad \bar{c}_u(u_k, z_k, \xi_k)^* \lambda = -\bar{J}_u(u_k, z_k, \xi_k).$$

H. Robbins & S. Monro, *A stochastic approximation method*, An. Math. Statist., 1951.

Analysis for Linear-Elliptic Quadratic Control

Recall: **(Spaces)** $Z = L^2(D)$ and $Z_{\mathrm{ad}} = \{z \in Z \mid z_a \le z \le z_b\}$, $U = H_0^1(D)$, $Y = H^{-1}(D)$, W is a Hilbert space such that $U \hookrightarrow W$,

$$\bar{J}(u,z,\xi) = \frac{1}{2}\|\mathbf{C}u - w\|_W^2 + \frac{\gamma}{2}\|z\|_Z^2$$

where $\mathbf{C} \in \mathcal{L}(U,W)$, $w \in W$ and $\gamma > 0$, and for $v \in U$

$$\langle \bar{c}(u,z,\xi), v\rangle_{-1,1} = \int_D (\overline{A}(\xi)\nabla u(x)) \cdot \nabla v(x)\,\mathrm{d}x - \int_D z(x)v(x)\,\mathrm{d}x.$$

Note: $\mathcal{J}(z) = \mathbb{E}[\widehat{J}(z)]$ is strongly convex with constant γ.

Stochastic Approximation: Given $z_k \in Z$ and $u_k \in U$ that solves $\bar{c}(u_k, z_k, \xi_k) = 0$

$$G_k(\xi_k) = \gamma z_k + \lambda_k$$

where λ_k solves the adjoint equation

$$\int_D (\overline{A}(\xi_k)\nabla\lambda_k(x)) \cdot \nabla v(x)\,\mathrm{d}x = -\langle \mathbf{C}u_k - w, \mathbf{C}v\rangle_W \quad \forall\, v \in U.$$

Analysis for Linear-Elliptic Quadratic Control

Let $z^* \in Z_{\text{ad}}$ minimize $\mathcal{J}(z)$ over Z_{ad} then (since $\Pi_{Z_{\text{ad}}}$ is nonexpansive)

$$\begin{aligned}\mathbb{E}[\|z_{k+1} - z^*\|_Z^2] &= \mathbb{E}[\|\Pi_{Z_{\text{ad}}}(z_k - \mu_k G_k(\xi_k)) - \Pi_{Z_{\text{ad}}}(z^*)\|_Z^2] \\ &\leq \mathbb{E}[\|z_k - z^*\|_Z^2] + \mu_k^2 \mathbb{E}[\|G_k(\xi_k)\|_Z^2] - 2\mu_k \mathbb{E}[\langle G_k(\xi_k), z_k - z^*\rangle_Z]\end{aligned}$$

z_k only depends on $\xi_1, \ldots, \xi_{k-1}$ (which are iid), thus

$$\begin{aligned}\mathbb{E}[\langle z_k - z^*, G_k(\xi_k)\rangle_Z] &= \mathbb{E}[\mathbb{E}[\langle z_k - z^*, G_k(\xi_k)\rangle_Z | \xi_1, \ldots, \xi_{k-1}]] && \text{Law of Total Exp.} \\ &= \mathbb{E}[\langle z_k - z^*, \mathbb{E}[G_k(\xi_k) | \xi_1, \ldots, \xi_{k-1}]\rangle_Z] && \text{Fubini's Theorem} \\ &= \mathbb{E}[\langle z_k - z^*, \mathbb{E}[\nabla \overline{F}(z_k)]\rangle_Z] \\ &\geq \mathbb{E}[\langle z_k - z^*, \mathbb{E}[\nabla \overline{F}(z_k) - \nabla \overline{F}(z^*)]\rangle_Z && \text{Optimality of } z^* \\ &\geq \gamma \mathbb{E}[\|z_k - z^*\|_Z^2]. && \text{Strong Convexity of } \mathcal{J}\end{aligned}$$

Analysis for Linear-Elliptic Quadratic Control

Since Z_{ad} is bounded and $u,\ \lambda$ depend continuously on z, we have

$$\mathbb{E}[\|G(z,\xi)\|_Z^2] \le M^2 \quad \forall z \in Z_{\mathrm{ad}}$$
$$\implies \quad \mathbb{E}[\|z_{k+1} - z^*\|_Z^2] \le \mathbb{E}[\|z_k - z^*\|_Z^2] + \mu_k^2 M^2 - 2\mu_k\gamma\mathbb{E}[\|z_k - z^*\|_Z^2].$$

Now, set $\mu_k = \theta/k$, then

$$\mathbb{E}[\|z_{k+1} - z^*\|_Z^2] \le \left(1 - \frac{2\gamma\theta}{k}\right)\mathbb{E}[\|z_k - z^*\|_Z^2] + \frac{\theta^2 M^2}{k^2}$$ Previous Results

$$\le \frac{\max\{\theta^2 M^2 (2\gamma\theta - 1)^{-1}, \|z_1 - z^*\|_Z^2\}}{k}.$$ Use Induction

Minimizing the right hand side with $\theta > 0$ gives $\theta^* = 1/\gamma$.

Note: The expected decay at each iteration is $\mathcal{O}(k^{-1})$

$\implies$ to reach tolerance ε requires $\mathcal{O}(\varepsilon^{-1})$ iterations (on average)!

Progressive Hedging

Problem Assumptions: Suppose $\widehat{J}(\cdot,\xi)$ is a **convex random loss** and ξ is discretely distributed. Consider the convex program

$$\underset{z\in Z_{\mathrm{ad}}}{\text{Minimize}} \left\{ \mathbb{E}[\widehat{J}(z,\xi)] = \sum_{k=1}^{N} p_k \widehat{J}(z,\xi_k) \right\}.$$

Progressive Hedging Algorithm:

Given $\hat{z} \in Z$ and a Z-valued r.v. $W(\xi)$ with $\mathbb{E}[W(\xi)] = 0$.

1. Compute $\zeta(\xi) \in Z_{\mathrm{ad}}$ a.s. that approximately solves

$$\underset{z\in Z_{\mathrm{ad}}}{\text{Minimize}} \left\{ \widehat{J}(z,\xi) + \langle W(\xi), z\rangle_Z + \tfrac{r}{2}\|z-\hat{z}\|_Z^2 \right\} \quad \text{a.s.}$$

2. Update $\hat{z} = \mathbb{E}[\zeta(\xi)] = p_1\zeta_1 + \ldots + p_N\zeta_N$.
3. Update $W(\xi) = W(\xi) + r(\zeta(\xi) - \hat{z})$.

Step 1 requires solving decoupled deterministic convex opt. problems!

However, objective function $\widehat{J}(\cdot,\xi)$ must be convex . . .

R. T. Rockafellar & R. J.-B. Wets, *Scenarios and policy aggregation in optimization under uncertainty*, Math. Oper. Res., 1991.

Sample Average Approximation

Idea: Approximate expected value in $\mathcal{J}$ using Monte Carlo.

Let $\xi_1, \ldots, \xi_N$ be iid random samples of ξ, then solve

$$\underset{z \in Z_{\text{ad}}}{\text{Minimize}} \left\{ \widehat{\mathcal{J}}_N(z) = \frac{1}{N} \sum_{k=1}^{N} \widehat{J}(z, \xi_k) \right\}.$$

Apply nonlinear programming algorithms to solve numerically.

Linear-Elliptic Quadratic Control:

(i) Let $z_N^* \in Z_{\text{ad}}$ minimize $\widehat{\mathcal{J}}_N$ over Z_{ad}

(ii) Let $z^* \in Z_{\text{ad}}$ minimize $\mathcal{J}$ over Z_{ad}.

Strong convexity of $\mathcal{J}$ and optimality of z_N^*, z^* imply

$$\begin{aligned} \gamma \|z_N^* - z^*\|_Z^2 &\le \langle z_N^* - z^*, \nabla \mathcal{J}(z_N^*) - \nabla \mathcal{J}(z^*) \rangle_Z \\ &\le \langle z_N^* - z^*, \nabla \mathcal{J}(z_N^*) - \nabla \widehat{\mathcal{J}}_N(z_N^*) \rangle_Z \end{aligned}$$

Therefore, $\gamma \|z^* - z_N^*\|_Z \le \left\| \mathbb{E}[\lambda] - \frac{1}{N} \sum_{k=1}^{N} \lambda_k \right\|_Z = \mathcal{O}(N^{-\frac{1}{2}})$ **Probabilistic!**

Stochastic Collocation

Idea: Approximate expected value in $\mathcal{J}$ using quadrature.

Let $\xi_1, \ldots, \xi_N$ be quad. points with weights $w_1, \ldots, w_N$, then solve

$$\operatorname*{Minimize}_{z \in Z_{\mathrm{ad}}} \left\{ \widehat{\mathcal{J}}_N(z) = \sum_{k=1}^{N} w_k \widehat{J}(z, \xi_k) \right\}.$$

Apply nonlinear programming algorithms to solve numerically.

Linear-Elliptic Quadratic Control:

(i) Let $z_N^* \in Z_{\mathrm{ad}}$ minimize $\widehat{\mathcal{J}}_N$ over Z_{ad}

(ii) Let $z^* \in Z_{\mathrm{ad}}$ minimize $\mathcal{J}$ over Z_{ad}.

Strong convexity of $\mathcal{J}$ and optimality of z_N^*, z^* imply

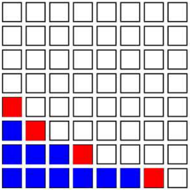
$$\begin{aligned}\gamma \|z_N^* - z^*\|_Z^2 &\le \langle z_N^* - z^*, \nabla \mathcal{J}(z_N^*) - \nabla \mathcal{J}(z^*) \rangle_Z \\ &\le \langle z_N^* - z^*, \nabla \mathcal{J}(z_N^*) - \nabla \widehat{\mathcal{J}}_N(z_N^*) \rangle_Z\end{aligned}$$

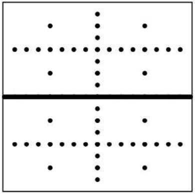
Therefore, $\gamma \|z^* - z_N^*\|_Z \le \left\| \mathbb{E}[\lambda] - \sum_{k=1}^{N} w_k \lambda_k \right\|_Z =$ **Quad. Error**

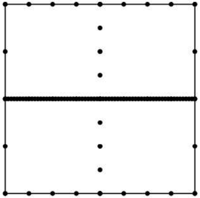
Sparse Grids and Adaptivity

Gerstner and Griebel 2003

- **1D Operators:** For $k = 1, \dots, M$, $\mathbb{E}_k^0 \equiv 0$ and
$$\Delta_k^i \equiv \mathbb{E}_k^i - \mathbb{E}_k^{i-1} \quad \text{where} \quad \mathbb{E}_k^i(g) \xrightarrow{i \to \infty} \int_{\Xi_k} \rho_k(\xi) g(\xi) \mathrm{d}\xi$$
- **Sparse-Grid Operator:** For an index set $\mathcal{I} \subset \mathbb{N}^M$,
$$\mathbb{E}_{\mathcal{I}} \equiv \sum_{\mathbf{i} \in \mathcal{I}} (\Delta_1^{i_1} \otimes \cdots \otimes \Delta_M^{i_M})$$
- **Admissibility:** $\mathbf{i} \in \mathcal{I}$ and $\mathbf{i} \geq \mathbf{j} \implies \mathbf{j} \in \mathcal{I}$
- **Error:** Given the index set $\mathcal{I} \subset \mathbb{N}^M$, the error is
$$\mathbb{E} - \mathbb{E}_{\mathcal{I}} = \sum_{\mathbf{i} \notin \mathcal{I}} (\Delta_1^{i_1} \otimes \cdots \otimes \Delta_M^{i_M})$$
- **Adaptivity:** Pick $\mathbf{i} \notin \mathcal{I}$ s.t. $\mathcal{I} \cup \{\mathbf{i}\}$ admissible and $\Delta_1^{i_1} \otimes \cdots \otimes \Delta_M^{i_M}$ "large"







Trust-Region Algorithm

Given: z_0, $m_0(s) \approx \mathcal{J}(z_0 + s)$, $\mathcal{J}_0 \approx \mathcal{J}$, $\Delta_0 \geq 0$, and $\mathrm{gtol} > 0$.
While $\|\nabla m_k(s)\|_{\mathcal{Z}} > \mathrm{gtol}$

1. **Model Update:** Choose a new $m_k(s) \approx \mathcal{J}(z_k + s)$. $\leftarrow$ **ADAPTIVITY**
2. **Step Computation:** Approximate a solution, s_k, to the subproblem

$$\min_{s \in \mathcal{Z}} m_k(s) \quad \text{subject to} \quad \|s\|_{\mathcal{Z}} \leq \Delta_k.$$

3. **Objective Update:** Choose a new $\mathcal{J}_k(z) \approx \mathcal{J}(z)$. $\leftarrow$ **ADAPTIVITY**
4. **Step Acceptance:** Compute

$$\rho_k = \frac{\mathcal{J}_k(z_k) - \mathcal{J}_k(z_k + s_k)}{m_k(0) - m_k(s_k)}.$$

   If $\rho_k \geq \eta \in (0,1)$, then $z_{k+1} = z_k + s_k$ else $z_{k+1} = z_k$.
5. **Trust Region Update:** Choose a new trust region radius, Δ_{k+1}.

EndWhile

Inexact Gradients and Objective Functions

Kouri, Heinkenschloss, Ridzal, and van Bloemen Waanders (2013, 2014)

Inexact Gradients

There exists $c > 0$ independent of k such that

$$\|\nabla m_k(0) - \nabla \mathcal{J}(z_k)\|_{\mathcal{Z}} \leq c \min\{\|\nabla m_k(0)\|_{\mathcal{Z}}, \Delta_k\}$$

(Carter 1989, Heinkenschloss and Vicente 2001).

Inexact Objective Functions

There exists $K > 0$, $\omega \in (0,1)$, and $\theta(z,s) \to 0$ as $r \to 0$ such that

$$|(\mathcal{J}(z_k) - \mathcal{J}(z_k + s_k)) - (\mathcal{J}_k(z_k) - \mathcal{J}_k(z_k + s_k))| \leq K\theta(z_k, s_k)$$
$$\theta(z_k, s_k)^{\omega} \leq \eta \min\{(m_k(0) - m_k(s_k)), r_k\}.$$

Here, $\eta > 0$ is tied to algorithmic parameters and $\lim_{k\to\infty} r_k = 0$.
(Carter 1989, Ziems and Ulbrich 2013).

- **Cannot** compute $\mathcal{J}(z_k)$ and $\nabla \mathcal{J}(z_k)$;
- Control *a posteriori* errors using **adaptive sparse grids**.

Optimal Control of Steady Burger's Equation

Let $\gamma = 10^{-3}$, $\Omega_o = \Omega_c = \Omega = (0,1)$, and $w \equiv 1$ and consider

$$\min_{z \in L^2(0,1)} \quad \mathcal{J}(z) = \frac{1}{2}\mathbb{E}\left[\int_0^1 (u(\cdot, x; z) - 1)^2 \,\mathrm{d}x\right] + \frac{\gamma}{2}\int_0^1 z(x)^2 \,\mathrm{d}x$$

where $u = S(z) \in L^3_\rho(\Xi; H^1(0,1))$ solves the weak form of

$$\begin{aligned} -\nu(\xi)\partial_{xx}u(\xi,x) + u(\xi,x)\partial_x u(\xi,x) &= f(\xi,x) + z(x) \qquad && (\xi,x) \in \Xi \times \Omega, \\ u(\xi,0) = d_0(\xi), \quad u(\xi,1) &= d_1(\xi) && \xi \in \Xi. \end{aligned}$$

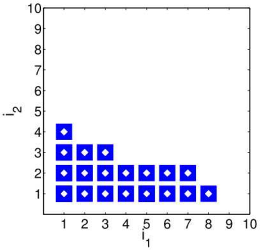
$\Xi = [-1,1]^4$ is endowed with the uniform density $\rho(\xi) \equiv 2^{-4}$, and the random field coefficients are

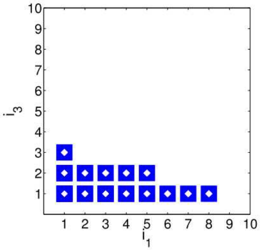
$$\nu(\xi) = 10^{\xi_1 - 2}, \quad f(\xi,x) = \frac{\xi_2}{100}, \quad d_0(\xi) = 1 + \frac{\xi_3}{1000}, \quad \text{and} \quad d_1(\xi) = \frac{\xi_4}{1000}.$$

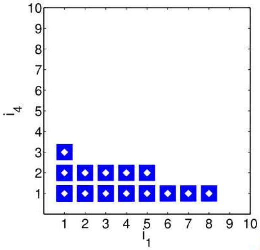
Adaptive Sparse Grid Results

Spatial: Piecewise Linear Finite Elements
Stochastic: Maximum Level 8 Clenshaw-Curtis Sparse Grids

Algorithm	NonlinPDE	CP_{obj}	LinearPDE	CP_{grad}	Rel. Err.
Newton-CG	45,224 (1.0)	7,537	489,906 (1.0)	7,537	–
Grad. Adapt.	45,531 (1.0)	7,537	3,405 (143.9)	249	2.89×10^{-6}
Full Adapt.	603 (75.0)	23	3,405 (143.9)	249	2.89×10^{-6}







Outline

References

1. **Ph. Artzner, F. Delbaen, J.-M. Eber & D. Heath,** *Coherent measures of risk.* Math. Finance, 1999.
2. **A. Ben Tal & M. Teboulle**, *Penalty functions and duality in stochastic programming via phi-divergence functionals*, Mathematics of Operations Research, 1987.
3. **A. Ben Tal & M. Teboulle**, *An old-new concept of convex risk measures: The optimized certainty equivalents*, Mathematical Finance, 2007.
4. **D. P. Kouri, M. Heinkenschloss, D. Ridzal, B. G. van Bloemen Waanders**, *A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty*, SISC, 2013.
5. **D. P. Kouri, M. Heinkenschloss, D. Ridzal, B. G. van Bloemen Waanders**, *Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty*, SISC, 2014.
6. **D. P. Kouri & A. Shapiro**, *Optimization of PDEs with Uncertain Inputs*, IMA Volumes in Mathematics and its Applications, 2018.
7. **D. P. Kouri & T. M. Surowiec**, *Existence and optimality conditions for risk-averse PDE-constrained optimization*, SIAM/ASA JUQ (accepted), 2018.
8. **P. A. Krokhmal**, *Higher moment coherent risk measures*, Quantitative Finance, 2007.
9. **S. Kusuoka**, *On law-invariant coherent risk measures*, Advances in Mathematical Economics, 2001.
10. **A. Mafusalov & S. Uryasev**, *Buffered probability of Exceedance: Mathematical Properties and Optimization Algorithms*, UF Technical Report, 2016.
11. **H. Markowitz**, *Portfolio selection*, The Journal of Finance, 1952.
12. **H. Robbins & S. Monro**, *A stochastic approximation method*, Annals of Mathematical Statistics, 1951.
13. **R. T. Rockafellar & S. Uryasev**, *The fundamental risk quadrangle in risk management, optimization, and statistical estimation*, Surveys in Operations Research & Managment Science, 2013.
14. **R. T. Rockafellar & J. O. Royset**, *On buffered failure probability in design and optimization of structures*, Reliability Engineering & System Safety, 2010.
15. **R. T. Rockafellar & J. O. Royset**, *Random variables, monotone relations, and convex analysis*, Mathematical Programming, 2014.
16. **R. T. Rockafellar & R. J.-B. Wets**, *Scenarios and policy aggregation in optimization under uncertainty*, Mathematics of Operations Research, 1991.
17. **W. W. Rogosinski**, *Moments of non-negative mass*, Proceedings of the Royal Society of London: Series A, Mathematical and Physical Sciences, 1958.
18. **A. Shapiro**, *Distributionally robust stochastic programming*, SIAM Journal on Optimization, 2017.
19. **A. Shapiro**, *Monte Carlo sampling methods*, Stochastic Programming, Handbooks in Operations Research and Management Science, 2003.
20. **A. Shapiro, D. Dentcheva & A. Ruszczynski**. *Lectures on Stochastic Programming: Modeling and Theory, Second Edition*, MOS-SIAM Series on Optimization, 2014.