Risk-Averse PDE-Constrained Optimization using the Conditional Value-At-Risk

SAND2014-4205C

Drew Kouri

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

May 19, 2014

SIAM Conference on Optimization

May 19-22, 2014. San Diego, California.

Joint work with T. Surowiec (HUBerlin)

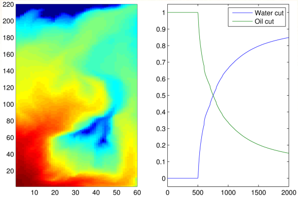
Funding support: ASC.

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

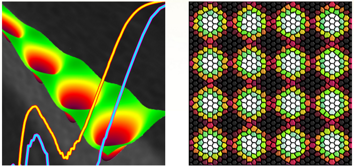
Motivation

Reservoir Optimization



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

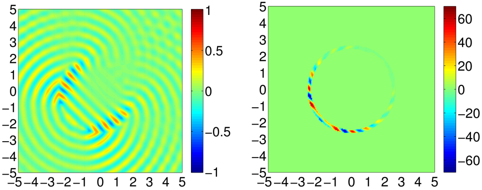
Superconductor Vortex Pinning



Courtesy Argonne National Laboratory

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

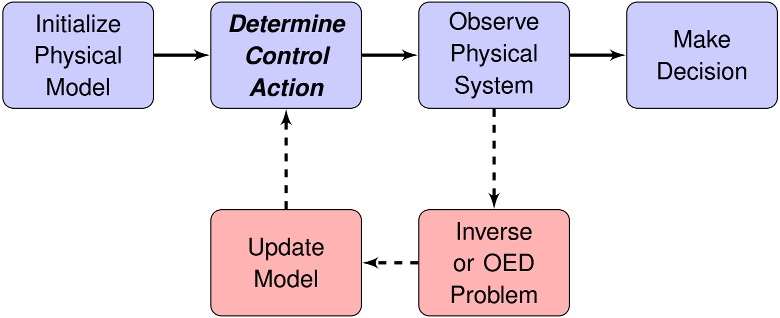
Direct Field Acoustic Testing



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

Optimization Problem Formulation

Goal: Control uncertainty rather than quantify uncertainty.



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

Optimization of PDEs with Random Inputs

Given $\alpha > 0$, $\Omega_o \subseteq \Omega$, $\Omega_c \subseteq \Omega$, and $w \in L^2(\Omega;\mathbb{C})$, we consider

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

where $u(z) = u \in L^{2p}_\rho(\Xi; H^1(\Omega))$ solves the **weak form** of

$$\begin{aligned} -\nabla_x\cdot(\epsilon(\xi,x)\nabla_x u(\xi,x)) + N(u(\xi,x),\xi) &= z(x), && x\in\Omega,\ \xi\in\Xi \\ u(\xi,x) &= 0, && x\in\partial\Omega,\ \xi\in\Xi. \end{aligned}$$

- **Random parameters ξ:**
  - Image space: $\Xi \equiv \Xi_1\times\cdots\times\Xi_M$ with $\Xi_k\subseteq\mathbb{R}$
  - Probability law: $\rho\equiv\rho_1\otimes\cdots\otimes\rho_M$ with $\rho_k:\Xi_k\to[0,\infty)\cup\{+\infty\}$
- **Control space:** $\mathcal{Z}\equiv L^2(\Omega_c)$ Deterministic
- **State space:** $\mathcal{U}\equiv L^{2p}_\rho(\Xi;H^1(\Omega))$ Stochastic
- **Risk Measure:** $\sigma : L^p_\rho(\Xi)\to\mathbb{R}\cup\{+\infty\}$
  see, Rockafellar, Uryasev, Shapiro, Dentcheva, Ruszczynski, ...

Sandia National Laboratories

Motivation - Control Uncertainty

Optimal control should be **"risk averse."** For example:

- Reduce **variance** or **deviation**

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

e.g. *reduce uncertainty and variability in controlled system.*

- Control **rare events**, **tail probabilities**, or **quantiles**

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

e.g. *reduce failure regions and certify reliability.*

- Minimize over **quantiles**

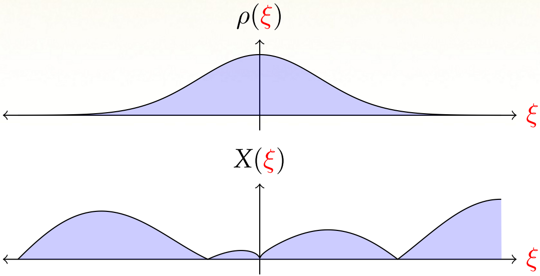
$$\mathrm{CVaR}_\beta[X] = \frac{1}{1-\beta} \int_{X \geq \mathrm{VaR}_\beta[X]} X(\xi)\rho(\xi)\,\mathrm{d}\xi = \mathbb{E}[X \mid X \geq \mathrm{VaR}_\beta[X]]$$

e.g. *minimize over undesirable events.*

Risk Measures

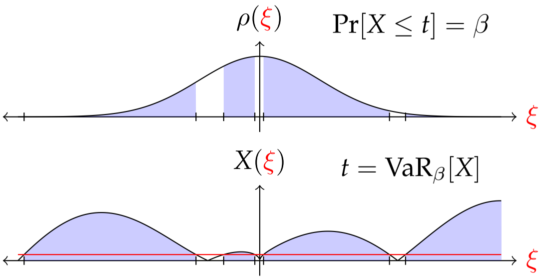
RISK NEUTRAL:

$\sigma[X] = \mathbb{E}[X]$



CONDITIONAL VALUE-AT-RISK:

$\sigma[X] = \mathrm{CVaR}_\beta[X]$



Risk Measures

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

$\sigma : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is a **monetary** risk measure if for $X,\, Y \in L^1_\rho(\Xi)$

- **Monotonicity:** $X \geq Y$ a.e. $\implies$ $\sigma[X] \geq \sigma[Y]$
- **Translation Equivariance:** $\sigma[X + t] = \sigma[X] + t, \quad \forall t \in \mathbb{R}$

σ is a **convex** risk measure if

- σ is a monetary risk measure
- **Convexity:** $\sigma[tX + (1-t)Y] \leq t\sigma[X] + (1-t)\sigma[Y], \quad \forall t \in [0,1]$

σ is a **coherent** risk measure if

- σ is a convex risk measure
- **Positive Homogeneity:** $\sigma[tX] = t\sigma[X], \quad \forall t > 0.$

Examples of coherent risk measures with $X \in \mathcal{X} = L^1_\rho(\Xi)$:

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

Sandia National Laboratories

Conditional Value-at-Risk

$$f(z;\xi) = \frac{1}{2}\int_{\Omega_o}(u(\xi,x\,;z) - w(x))^2\,\mathrm{d}x + \frac{\alpha}{2}\int_{\Omega_c} z^2(x)\,\mathrm{d}x$$

Primal: Rockafellar and Uryasev

- Assume the CDF, i.e., $\Psi(z,t) = \Pr\left[f(z) \le t\right]$, is continuous w.r.t. t.
- Then,

$$\min_{z\in\mathcal{Z}}\inf_{t\in\mathbb{R}}\left\{t + \frac{1}{1-\beta}\mathbb{E}[(f(z)-t)_+]\right\} \iff \min_{(t,z)\in\mathbb{R}\times\mathcal{Z}}\left\{t + \frac{1}{1-\beta}\mathbb{E}[(f(z)-t)_+]\right\}.$$

Dual: Fenchel duality

- CVaR is proper, lower semicontinuous, and convex.
- Moreover, CVaR is coherent and

$$\min_{z\in\mathcal{Z}}\inf_{t\in\mathbb{R}}\left\{t + \frac{1}{1-\beta}\mathbb{E}[(f(z)-t)_+]\right\} \iff \min_{z\in\mathcal{Z}}\sup_{\varrho\in\mathcal{A}}\int_{\Xi}\varrho(\xi)\rho(\xi)\,\mathrm{d}\xi.$$

- $\mathcal{A} = \left\{\varrho \in (L^1_\rho(\Xi))^* \;:\; \varrho(\xi) \in \left[0, \frac{1}{1-\beta}\right]\ \rho\text{-a.e.},\ \int_\Xi \varrho(\xi)\rho(\xi)\,\mathrm{d}\xi = 1\right\}.$

Sandia National Laboratories

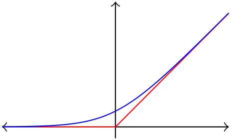
Primal Approach

To discretize and solve optimal control problem, we ...

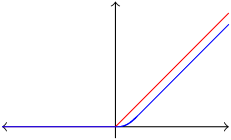
- Use deterministic quadrature to approximation the expected value;
- Smooth $(\cdot)_+$ to ensure convergence of quadrature approximation;
- Minimize reformulation of CVaR_β over $\mathbb{R} \times \mathcal{Z}$.

Examples:

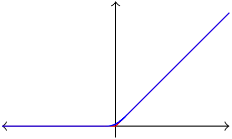
$$(x)^+_{\varepsilon,1} = x + \varepsilon \log\left(1 + \exp\left(\frac{-x}{\varepsilon}\right)\right)$$



$$(x)^+_{\varepsilon,2} = \begin{cases} 0 & \text{if } x \le 0 \\ \left(\frac{x^3}{\varepsilon^2} - \frac{x^4}{2\varepsilon^3}\right) & \text{if } x \in (0,\varepsilon) \\ x - \frac{\varepsilon}{2} & \text{if } x \ge \varepsilon \end{cases}$$



$$(x)^+_{\varepsilon,3} = \left(x + \frac{\varepsilon}{2}\right)^+_{\varepsilon,2}$$



Smoothed CVaR Results

Results for smoothed plus function:

- $(x)^+_{\varepsilon,1} > (x)^+_{\varepsilon,3} \geq (x)_+ \geq (x)^+_{\varepsilon,2}$ for all $x \in \mathbb{R}$.
- $|(x)^+_{\varepsilon,\ell} - (x)_+| \leq c_\ell \varepsilon$ for all $x \in \mathbb{R}$ where $c_1 = \log(2)$, $c_2 = \frac{1}{2}$, and $c_3 = \frac{3}{32}$.

Results for smoothed CVaR: $\sigma^\beta_{\varepsilon,\ell}[X] = \inf\left\{t + \frac{1}{1-\beta}\mathbb{E}[(X-t)^+_{\varepsilon,\ell}] \;:\; t \in \mathbb{R}\right\}$

- $\left|\sigma^\beta_{\varepsilon,\ell}[X] - \mathrm{CVaR}_\beta[X]\right| \leq \frac{c_\ell}{1-\beta}\varepsilon$ for all $X \in L^1_\rho(\Xi)$.
- Smoothed CVaR, $\sigma^\beta_{\varepsilon,\ell}$, is a convex risk measure.
- $X \mapsto \left\{t + \frac{1}{1-\beta}\mathbb{E}[(X-t)^+_{\varepsilon,\ell}]\right\}$ is Hadamard differentiable.
- $t \mapsto \left\{t + \frac{1}{1-\beta}\mathbb{E}[(X-t)^+_{\varepsilon,\ell}]\right\}$ is continuously differentiable.

Sandia National Laboratories

Smoothed-CVaR Optimal Control Results

- **Existence of Optimal Controls:** If the state equation is uniquely solveable, then $\exists$ a minimizer of

$$J_{\varepsilon,\ell}(t,z) \equiv \frac{1}{2}\left(t + \frac{1}{1-\beta}\mathbb{E}\left[\left(\int_{\Omega_o}(u(\cdot,x;z) - w(x))^2\,\mathrm{d}x\right)^{+}_{\varepsilon,\ell}\right]\right) + \frac{\alpha}{2}\int_{\Omega_c} z^2(x)\,\mathrm{d}x$$

- **Differentiability:** $J_{\varepsilon,\ell}(t,z)$ is Hadamard differentiable w.r.t. z and continuously differentiable w.r.t. t.
- **Consistency:** Suppose $\varepsilon_k \searrow 0$ and $(t_{k,\ell}, z_{k,\ell})$ is a minimizer of $J_{\varepsilon_k,\ell}(t,z)$. Then, $\exists$ a subsequence of $(t_{k,\ell}, z_{k,\ell})$ which weakly converges to a minimizer of

$$J(t,z) \equiv \frac{1}{2}\left(t + \frac{1}{1-\beta}\mathbb{E}\left[\left(\int_{\Omega_o}(u(\cdot,x;z) - w(x))^2\,\mathrm{d}x\right)_{+}\right]\right) + \frac{\alpha}{2}\int_{\Omega_c} z^2(x)\,\mathrm{d}x.$$

- **Convergence Rate:** Suppose $(t_{\varepsilon,\ell}, z_{\varepsilon,\ell})$ is a minimizer of $J_{\varepsilon,\ell}(t,z)$ and (t^*, z^*) is a minimizer of $J(t,z)$. Then, $(|t^* - t_{\varepsilon,\ell}|^2 + \|z^* - z_{\varepsilon,\ell}\|^2_{\mathcal{Z}})^{\frac{1}{2}} \le C_\ell \varepsilon^{\frac{1}{2}}$.

Quadrature Discretization

Replace expected values with quadrature approximation

$$\mathbb{E}_Q[X] = \sum_{k=1}^{Q} \omega_k X(\xi_k), \quad \sigma^Q_{\varepsilon,\ell}[X] = \inf\left\{ t + \frac{1}{1-\beta}\mathbb{E}_Q[(X-t)^+_{\varepsilon,\ell}] \;:\; t \in \mathbb{R} \right\}, \quad \text{and}$$

$$J^Q_{\varepsilon,\ell}(t,z) \equiv \frac{1}{2}\left(t + \frac{1}{1-\beta}\sum_{k=1}^{Q}\omega_k \left(\int_{\Omega_o} (u(\xi_k, x; z) - w(x))^2 \,\mathrm{d}x - t \right)^+_{\varepsilon,\ell} \right) + \frac{\alpha}{2}\int_{\Omega_c} z^2(x)\,\mathrm{d}x$$

Require the solutions $u(\xi_k) = u_k \in H^1(\Omega)$, $k = 1, \ldots, Q$, solves the **weak form** of

$$\begin{aligned} -\nabla_x \cdot (\epsilon(\xi_k, x)\nabla_x u_k(x)) + N(u_k(x), \xi_k) &= z(x), & x &\in \Omega \\ u_k(x) &= 0, & x &\in \partial\Omega. \end{aligned}$$

- Decoupled PDE system is equivalent to **stochastic collocation** for a single PDE
- Use favorite numerical PDE technique to solve deterministic PDEs
- Convergence of quad. approx. depends on regularity of state and adjoint w.r.t. ξ.

Existence of Optimal Controls: There exists a minimizer of $J^Q_{\varepsilon,\ell}(t,z)$ in the set $\{\ (t,z) \in \mathbb{R} \times \mathcal{Z} \ :\ \|z\|_{\mathcal{Z}} \le K\ \}$ for any $K > 0$.

Add constraint because quadrature weights may be **negative**!

Dual Approach

Approximation: Add strongly concave regularization

$$\max_{\varrho\in\mathcal{A}} \frac{1}{2}\int_\Xi \varrho(\xi)\rho(\xi)\int_{\Omega_o}(u(\xi,x;z)-w(x))^2\,\mathrm{d}x\mathrm{d}\xi+\frac{\alpha}{2}\int_{\Omega_c} z^2(x)\,\mathrm{d}x-\frac{\gamma}{2}\int_\Xi \varrho^2(\xi)\rho(\xi)\,\mathrm{d}\xi$$

where

$$\mathcal{A}=\left\{\varrho\in(L^1_\rho(\Xi))^*\;:\;\varrho(\xi)\in\left[0,\frac{1}{1-\beta}\right]\ \rho\text{-a.e.},\ \int_\Xi \varrho(\xi)\rho(\xi)\,\mathrm{d}\xi=1\right\}.$$

First-Order Necessary Conditions: Robinson's CQ holds and the maximal density ϱ_γ satisfies

$$\begin{aligned}\varrho_\gamma(\xi)=&\frac{1}{\gamma}\left(\frac{1}{2}\int_{\Omega_o}(u(\xi,x;z)-w(x))^2\,\mathrm{d}x-\mu\right)\\&+\left(\frac{1}{\gamma}\left(\mu-\frac{1}{2}\int_{\Omega_o}(u(\xi,x;z)-w(x))^2\,\mathrm{d}x\right)\right)_+\\&-\left(\frac{1}{\gamma}\left(\frac{1}{2}\int_{\Omega_o}(u(\xi,x;z)-w(x))^2\,\mathrm{d}x-\mu\right)-\frac{1}{1-\beta}\right)_+\end{aligned}$$

for all μ in some nonempty, compact interval $\Lambda\subset\mathbb{R}$.

Consistency: For any $\gamma_k\searrow 0$, there exists a subsequence of maximal densities ϱ_{γ_k} which weak* converges to the true maximal density ϱ^*.

Sandia National Laboratories

Optimal Control of Steady Burger's Equation

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

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

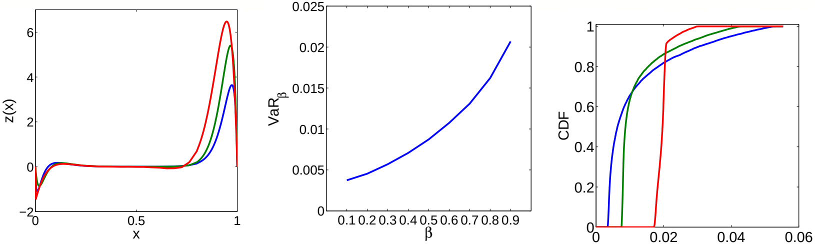
where $u = u(z) \in L^3_\rho(\Xi; H^1(0,1))$ solves the weak form of

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

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

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

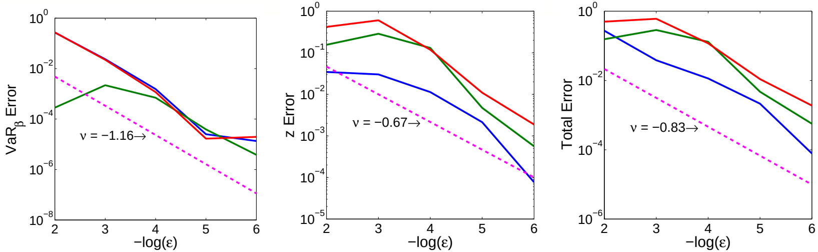
Optimal Controls



$\beta = 0.1$, $\beta = 0.5$, $\beta = 0.9$

Left: Optimal controls. **Center:** VaR varying β. **Right:** CDF of tracking term.

Convergence of Smoothing

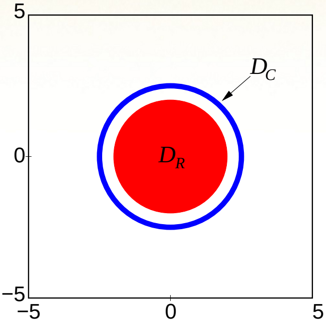


$\beta = 0.1$, $\beta = 0.5$, $\beta = 0.9$

Left: Error in the VaR. **Center:** Error in optimal controls. **Right:** Total error.

Direct Field Acoustic Testing (DFAT)

- **Physical Domain:** $D = (-5,5)^2$
- **Parameter Space:** $\Xi = [-\sqrt{3}, \sqrt{3}]^M$
- **Probability Measure:**
  $\rho(\xi)\mathrm{d}\xi = \left(2\sqrt{3}\right)^{-M} \mathrm{d}\xi$
- **Stochastic Material:** $\epsilon(\xi, x)$
  KL expansion of Matérn covariance
- **Desired State:** $\theta = \frac{\pi}{4}$, $k = 10$
  $\bar{w}(x) = \exp\left(i\left((k\cos\theta)x_1 + (k\sin\theta)x_2\right)\right)$



Let $\alpha > 0$ and $\vartheta = 0.1$. Consider the optimal control problem

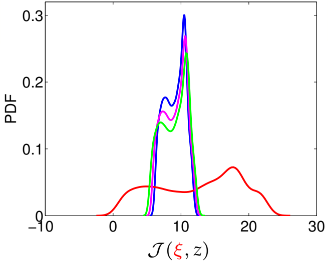
$$\min_{z\in L^2(D;\mathbb{C})} \frac{1}{2}\sigma\left[\int_{D_R} (u(z;\xi,x) - \bar{w}(x))\overline{(u(z;\xi,x) - \bar{w}(x))}\,\mathrm{d}x\right] + \frac{\alpha}{2}\int_{D_c} z(x)\overline{z(x)}\,\mathrm{d}x$$

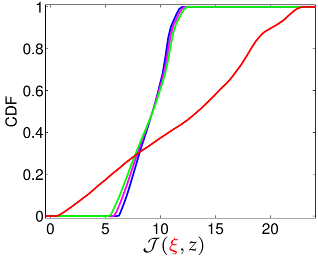
where $u = u(z) \in L^2_\rho(\Xi; H^1(D;\mathbb{C}))$ solves

$$\begin{aligned} -\Delta u(\xi,x) - k^2(1+\vartheta\epsilon(\xi,x))^2 u(\xi,x) &= z(x) \qquad && \forall\ (\xi,x) \in \Xi \times D \\ \frac{\partial u}{\partial n}(\xi,x) &= iku(\xi,x) && \forall\ (\xi,x) \in \Xi \times \partial D. \end{aligned}$$

Results: $M = 6$, $\gamma = 5$, and $\alpha = 10^{-4}$

$$\mathcal{J}(\xi, z) = \int_{D_R} (u(z;\xi,x) - \bar{w}(x))\overline{(u(z;\xi,x) - \bar{w}(x))}\,\mathrm{d}x.$$





- Mean value: $\xi \leftarrow \mathbb{E}[\xi]$
- Mean plus CVaR: $\sigma[X] = \frac{1}{2}\mathbb{E}[X] + \frac{1}{2}\mathrm{CVaR}_{0.1}[X]$
- Risk neutral: $\sigma[X] = \mathbb{E}[X]$
- CVaR: $\sigma[X] = \mathrm{CVaR}_{0.1}[X]$

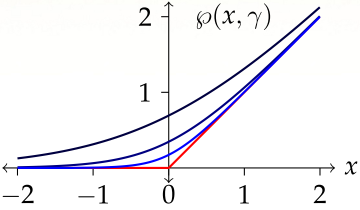
Mean plus CVaR Value-at-Risk: $t = 6.59073$

CVaR Value-at-Risk: $t = 6.90629$

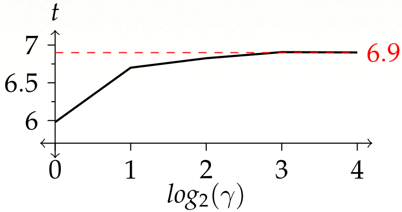
Sandia National Laboratories

Results: The Effect of Smoothing - CVaR

Smoothed Plus Function



VaR Approximation



γ	$\|z\|_{\mathcal{Z}}$	Abs. Err.	Rate	t	Abs. Err.	Rate
1	37.8369	-	-	5.9792	-	-
2	37.0495	1.2177	-	6.7004	0.7212	-
4	36.7416	0.7111	0.7760	6.8258	0.1254	2.5239
8	36.6653	0.4396	0.6939	6.9066	0.0808	0.6341

Theoretical convergence rate is $\frac{1}{2}$ (Kouri and Surowiec).

Conclusions:

- **Primal** and **dual** formulations of CVaR
- **Primal:** Use reformulation, deterministic quadrautre, and smoothing
- Can prove consistency for smoothing and explicit convergence rate
- **Dual:** Add strongly concave regularization and formulate opt. conditions
- Can prove consistency of regularized solution

Future Work:

- **Coherent risk measures** for *risk-averse* optimization under uncertainty (Ruszczynski, Shapiro, Rockafellar, Uryasev, Föllmer, Schied, ...)
- **Probabilistic constraints** to control *tail-probabilities* and *rare events* (Shapiro, Uryasev, Henrion, Kibzun, ...)
- **Ambiguous stochastic programming** to incorporate *data* (Shapiro, Bertsimas, Bayraksan, ...)

$$\min_{z \in \mathcal{Z}} \sup_{P \in \mathcal{M}_0} \int_{\Xi} f(u(z;\xi), z, \xi)\,\mathrm{d}P(\xi)$$

Sandia National Laboratories