SAND2015-10185PE

Risk-Averse Optimization of Large-Scale Multiphysics Systems

D. P. Kouri (**PI**, Sandia), W. Aquino (Duke), D. Ridzal (Sandia),
R. T. Rockafellar (UFlorida), A. Shapiro (Georgia Tech),
S. Uryasev (UFlorida)

November 10, 2015

DARPA EQUiPS Kick-Off Meeting, DARPA Conference Center, Arlington, VA.

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.

Outline

Introduction
- Team Members & Expertise
- Computational & Software Infrastructure

Motivating Applications

Technical Challenges
- Risk Mitigation Strategy

Our Approach
- Risk Quadrangle
- Buffered Probability
- Time Consistency
- Optimization-Based Sampling

Outline

Introduction
- Team Members & Expertise
- Computational & Software Infrastructure

Motivating Applications

Technical Challenges
- Risk Mitigation Strategy

Our Approach
- Risk Quadrangle
- Buffered Probability
- Time Consistency
- Optimization-Based Sampling

Team Members

Drew Kouri
(Sandia, **PI**)

Denis Ridzal
(Sandia)

Wilkins Aquino
(Duke)

Stan Uryasev
(Florida)

R. Tyrrell Rockafellar
(Florida)

Alexander Shapiro
(Georgia Tech)

Our Expertise

- **Risk-Averse Optimization:**
  - **A. Shapiro**, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia, 2009.
  - **A. Shapiro**. "Time consistency of dynamic risk measures", Operations Research Letters, vol. 40, pp. 436-439, 2012.
  - **A. Shapiro**. "On a time consistency concept in risk averse multi-stage stochastic programming", Operations Research Letters, vol. 37, pp. 143-147, 2009.
  - **R. T. Rockafellar** and **S. Uryasev**, "The Fundamental Risk Quadrangle in Risk Management, Optimization, and Statistical Estimation", Surveys in Operations Research and Management Science, 18, 2013.
  - **R. T. Rockafellar**, **S. Uryasev**, and M. Zabarankin. "Generalized Deviations in Risk Analysis", Finance and Stochastics, 10, 2006.
  - **R. T. Rockafellar** and **S. Uryasev**, "Conditional Value-at-Risk for General Loss Distributions", Journal of Banking and Finance, vol. 26(7), 2002.
  - **R. T. Rockafellar** and J. Royset, "Engineering Decisions under Risk Averseness", ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A: Civ. Eng., 1(2), 2015.
  - **R. T. Rockafellar**, "Coherent Approaches to Risk in Optimization Under Uncertainty", OR Tools and Applications: Glimpses of Future Technologies, 2007.
  - **D. P. Kouri** and T. M. Surowiec, "Risk-Averse PDE-Constrained Optimization using the Conditional Value-at-Risk", SIAM Journal on Optimization, to appear.

Our Expertise

- **Probabilistic Computation & Statistical Estimation:**
  - S. Y. Chun, **A. Shapiro**, and **S. Uryasev**, "Conditional Value-at-Risk and Average Value-at-Risk: Estimation and Asymptotics", Operations Research, vol. 60, pp. 739-756, 2012.
  - B. Pagnoncelli, S. Ahmed, and **A. Shapiro**, "Sample Average Approximation Method for Chance Constrained Programming: Theory and Applications", Journal Optimization Theory and Applications, vol. 142, pp. 399-416, 2009.
  - A. Veremyev, P. Tsyurmasto, **S. Uryasev**, and **R. T. Rockafellar**, "Calibrating Probability Distributions with Convex-Concave-Convex Functions: Application to CDO Pricing", Computational Management Science, 2013.
  - **R. T. Rockafellar** and J. Royset, "On Buffered Failure Probability in Design and Optimization of Structures", J. Reliability Engineering and System Safety, vol. 95, 2010.
- **Discrete Density Estimation & Stochastic Reduced Order Models**
  - B. Ergashev, K. Pavlikov, **S. Uryasev**, and E. Sekeris, "Estimation of Truncated Data Samples in Operational Risk Modeling". Journal of Risk and Insurance, 2015.
  - S. Sarkar, J. E. Warner, **W. Aquino**, and M. D. Grigoriu, "Stochastic Reduced Order Models for Uncertainty Quantification of Intergranular Corrosion Rates", Corrosion Science, 2013.
  - J. E. Warner, M. D. Grigoriu, and **W. Aquino**, "Stochastic Reduced Order Models for Random Vectors: Application to Random Eigenvalue Problems", Probabilistic Engineering Mechanics, vol. 31, 2013.

Our Expertise

- **PDE-Constrained Optimization, Inexactness, & Adaptivity:**
  - **D. P. Kouri**, M. Heinkenschloss, **D. Ridzal**, and B. G. van Bloemen Waanders, “A Trust-Region Algorithm with Adaptive Stochastic Collocation for PDE Optimization under Uncertainty”, SIAM Journal on Scientific Computing, vol. 35(4), 2013.
  - **D. P. Kouri**, M. Heinkenschloss, **D. Ridzal**, and B. G. van Bloemen Waanders, “Inexact Objective Function Evaluations in a Trust-Region Algorithm for PDE-Constrained Optimization under Uncertainty”, SIAM Journal on Scientific Computing, vol. 36(6), 2014.
  - **D. P. Kouri**, “A Multilevel Stochastic Collocation Algorithm for Optimization of PDEs with Uncertain Coefficients”, SIAM/ASA Journal on Uncertainty Quantification, vol. 2(1), 2014.
  - M. Heinkenschloss and **D. Ridzal**, “A Matrix-Free Trust-Region SQP Method for Equality Constrained Optimization”, SIAM Journal on Optimization, vol. 24(3), 2014.
  - M. Heinkenschloss and **D. Ridzal**, “An Inexact Trust-Region SQP Method with Applications to PDE-Constrained Optimization”, Numerical Mathematics And Advanced Applications, EUMATH 2007, Springer-Verlag, Heidelberg, 2008.
  - M. Heinkenschloss and **D. Ridzal**, “Integration of Sequential Quadratic Programming and Domain Decomposition Methods for Nonlinear Optimal Control Problems”, In Domain Decomposition Methods in Science and Engineering XVII, Lecture Notes in Computational Science and Engineering Vol. 60, Springer-Verlag, Heidelberg, 2008.

Our Expertise

- **Numerical Methods and Software for PDEs:**
  - P. B. Bochev, **D. Ridzal**, "Rehabilitation of the Lowest-Order Raviart-Thomas Element on Quadrilateral Grids", SIAM Journal on Numerical Analysis, vol. 47(1), 2007.
  - P. B. Bochev, H. C. Edwards, K. J. Peterson, **D. Ridzal**, "Solving PDEs with Intrepid", Scientific Programming, vol. 20(2), 2011.
  - N. D. Roberts, P. B. Bochev, L. D. Demkowicz, **D. Ridzal**, "A toolbox for a class of Discontinuous Petrov-Galerkin methods using Trilinos", Sandia Report, SAND2011-6678, 2011.
  - **W. Aquino**, J. C. Brigham, N. Sukumar, C. J. Earls, "Generalized Finite Element Method using Proper Orthogonal Decomposition", International Journal for Numerical Methods in Engineering, vol. 79, 2009.
  - *Intrepid*, www.trilinos.org/packages/intrepid, Lead developers: **D. Ridzal**, P. B. Bochev.
  - *Sierra Structural Dynamics*, www.sandia.gov/asc/integrated_codes.html.

- **Software for Large-Scale Optimization:**
  - *Rapid Optimization Library (ROL)*, www.trilinos.org/packages/rol, Lead developers: **D. P. Kouri** and **D. Ridzal**.
  - *Portfolio Safeguard (PSG)*, www.aorda.com/aod/psg.action, Lead developer: **S. Uryasev**.

Outline

Introduction
- Team Members & Expertise
- Computational & Software Infrastructure

Motivating Applications

Technical Challenges
- Risk Mitigation Strategy

Our Approach
- Risk Quadrangle
- Buffered Probability
- Time Consistency
- Optimization-Based Sampling

Computational & Software Infrastructure

- **Optimization:** Rapid Optimization Library (ROL)
  Lead Developers: **Kouri** and **Ridzal**
  - Matrix-free simulation-based optimization interface
  - State-of-the-art algorithms for large-scale unconstrained and constrained optimization
  - Interface for optimization under uncertainty including risk, buffered probability, and sample generation
- **Application:** Intrepid (**Ridzal**), Sacado, MueLu, etc.
  - Intrepid implements local finite-element computations
  - Trilinos packages for automatic differentiation, algebraic multigrid, (non)linear solvers, mesh data structures, etc.
  - Rapid development of large-scale multiphysics applications with scalable solvers and interfaces for optimization

---

- **Optimization:** Portfolio Safeguard
  Chief Research Consultant and Owner: **Uryasev**
- Nonlinear and mixed-integer nonlinear optimization
- Numerous risk and probabilistic functions implemented
- Large scale: 1,000,000 scenarios and 200,000 variables

Outline

Introduction
- Team Members & Expertise
- Computational & Software Infrastructure

Motivating Applications

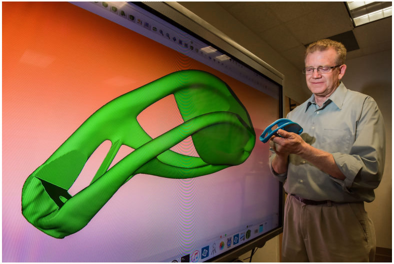
Technical Challenges
- Risk Mitigation Strategy

Our Approach
- Risk Quadrangle
- Buffered Probability
- Time Consistency
- Optimization-Based Sampling

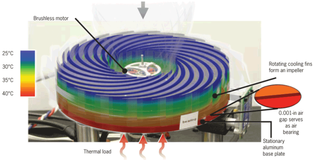
Motivating Applications

Direct-field acoustic testing

Topology optimization for structural design



Optimal control of thermal fluids

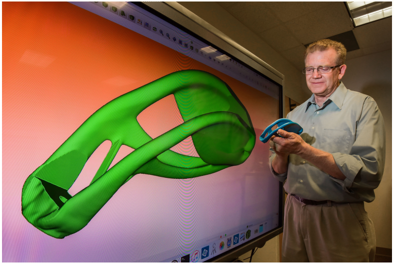


Optimal control of plasma instabilities

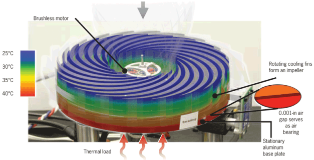
Motivating Applications

Direct-field acoustic testing

Topology optimization for structural design



Optimal control of thermal fluids

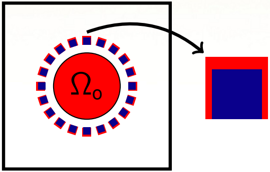


Optimal control of plasma instabilities

*Overarching challenge: Incorporate the treatment of **risk** in the mathematical formulation, theoretical analysis and numerical solution.*

Direct-Field Acoustic Testing

"RISK-NEUTRAL" OPTIMAL CONTROL OF THE HELMHOLTZ EQUATION



Setup: Red denotes uncertainty in the refraction index inside Ω_o and uncertainty in the loudspeaker enclosures (including thickness and sound speed). Blue stands for the loudspeaker control regions, whose union is denoted by Ω_c.

$$\underset{z \in L^2(\Omega_c)}{\text{Minimize}} \int_{\Xi} \rho(\xi) \int_{\Omega_o} |(U(z))(\xi, x) - w(x)|^2 dx\, d\xi + \alpha \int_{\Omega_c} |z(x)|^2 dx,$$

where $U(z) = u : \Xi \to H^1(\Omega)$ solves

$$-\Delta u(\xi) - K(\xi)u(\xi) = \mathbb{1}_{\Omega_c} z, \qquad \text{in } \Omega, \text{ a.s.}$$

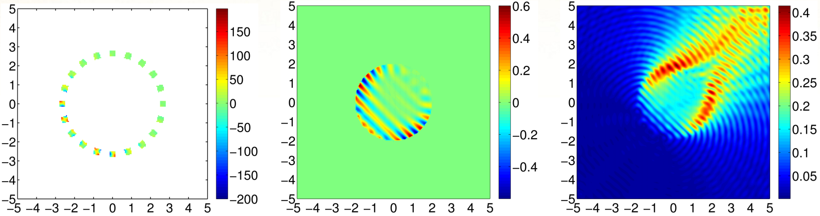
with Robin boundary conditions

$$\frac{\partial u}{\partial n}(\xi) = iku(\xi), \qquad \text{on } \partial\Omega, \text{ a.s.}$$

- Large-scale optimization, incl. solution of Helmholtz equation in 3D.
- Uncertainty in material properties, loudspeaker specifications, etc.
- Risk-neutral (expected value) formulation of the misfit objective: The state $u(\xi)$ should match w on average; *cf.* experimental science.
- Challenge 1: Large-dimensional space of uncertain parameters, Ξ.
- Partial solution: Adaptive objective function and gradient evaluations.

Direct-Field Acoustic Testing

"RISK-NEUTRAL" OPTIMAL CONTROL OF THE HELMHOLTZ EQUATION



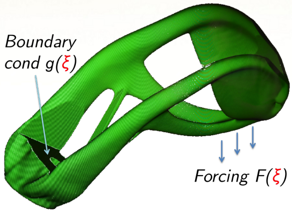
Results: *Left pane:* The real parts of the computed optimal controls. *Center pane:* The real part of the expected value of the optimal state, restricted to the region of interest. *Right pane:* The real part of the standard deviation of the optimal state.

dim	PDE Solves	CP_{obj}	CP_{grad}	Obj. Value
42	11,543	145	145	5.2542
44	32,739	233	481	5.2637
46	60,617	243	1,453	5.2641
48	79,221	247	2,961	5.2641
50	90,157	251	4,569	5.2641
60	100,911	271	7,621	5.2641
70	103,979	291	8,233	5.2641
80	105,607	311	8,253	5.2641

Cost: Roughly <u>200x</u> the cost of deterministic optimization, for <u>80-dim</u> uncertainty.

Topology Optimization for Structural Design

"RISK-AVERSE" OPTIMAL DESIGN FOR LINEAR ELASTICITY EQUATIONS



Setup: The forcing or load $F(\xi)$ on the right part of the bracket is uncertain. Additionally, there is an uncertain Dirichlet condition on the displacement at the bolt location, see $g(\xi)$.

Given volume fraction $V_0 \in (0,1)$, max compliance η, $\Omega \subset \mathbb{R}^3$,

$$\underset{0 \le z \le 1}{\text{Minimize}} \ \text{Prob}\left[\int_\Omega F(\xi,x)\cdot(U(z))(\xi,x)\,dx \ge \eta\right]$$

s.t. $\int_\Omega z(x)\,dx \le V_0|\Omega|$, where $U(z) = u : \Xi \to (H^1(\Omega))^3$ solves

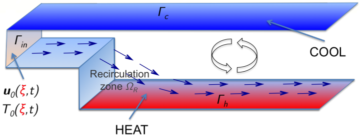
$$\begin{aligned} -\nabla\cdot(\mathbf{E}(z):\varepsilon u(\xi)) &= F(\xi), && \text{in } \Omega, \ \text{a.s.} \\ \varepsilon u(\xi) &= \frac{1}{2}(\nabla u(\xi) + \nabla u(\xi)^\top), && \text{in } \Omega, \ \text{a.s.} \\ u(\xi) &= g(\xi), && \text{on } \partial\Omega, \ \text{a.s.} \end{aligned}$$

- Uncertainty in external forces (loads) and boundary conditions.
- Reliability formulation of the compliance objective: Compute light-weight designs that reduce the probability of structural failure.
- Challenge 2: Nonsmooth objective functions and constraints.
- Challenge 3: Rare-event detection and computation.

Optimal Control of Thermal Fluids

“TIME-CONSISTENT” RISK-AVERSE CONTROL OF MULTIPHYSICS SYSTEMS

$$\underset{a(t) \,\le\, z(t) \,\le\, b(t)}{\text{Minimize}} \quad \text{Risk}\left(\int_{\Omega_R} |\nabla \times (\mathbf{U}(z))(\xi, x, \tau^{\text{final}})|^2 dx\right)$$



Setup: The velocity field and temperature at the inlet Γ_{in} are time dependent and uncertain, see $\mathbf{u}_0(\xi, t)$ and $T_0(\xi, t)$. A time dependent temperature control $z(t)$ is applied on the top and bottom boundaries, Γ_c and Γ_h, to create thermal flow counteracting the vortex in the recirculation zone Ω_R.

where $\mathbf{U}(z) = \mathbf{u} : \Xi \to (H^1(\Omega))^3 \times L^2([0, \tau^{\text{final}}])$ solves

$$\frac{\partial \mathbf{u}(\xi)}{\partial t} - \nu_1 \Delta \mathbf{u}(\xi) + (\mathbf{u}(\xi) \cdot \nabla)\mathbf{u}(\xi) + \nabla p(\xi) + \nu_2 T(\xi)\mathbf{g} = 0$$

$$\frac{\partial T(\xi)}{\partial t} - \nu_3 \Delta T(\xi) + \mathbf{u}(\xi) \cdot \nabla T(\xi) = 0$$

and $\nabla \cdot \mathbf{u}(\xi) = 0$, in $\Omega \times [0, \tau^{\text{final}}]$, with BCs

$\mathbf{u} = \mathbf{u}_0(\xi, t)$ and $T = T_0(\xi, t)$ on $\Gamma_{\text{in}} \times [0, \tau^{\text{final}}]$, etc.,

and heat-flux control $z(t) = z$ satisfying

$$\frac{\partial T(\xi)}{\partial \mathbf{n}} = h(z - T(\xi)) \quad \text{on} \quad \Gamma_h \quad \text{and} \quad \Gamma_c.$$

- Uncertainty in the velocity field and temperature at the inlet.
- A thermal fluid system with time-dependent temperature control.
- Challenge 4: Time consistency for continuous-time systems.

Outline

Introduction
- Team Members & Expertise
- Computational & Software Infrastructure

Motivating Applications

Technical Challenges
- Risk Mitigation Strategy

Our Approach
- Risk Quadrangle
- Buffered Probability
- Time Consistency
- Optimization-Based Sampling

Target Optimization Formulations

Goal: Develop *efficient* methods to determine *resilient* optimal controls & designs that *mitigate high-consequence rare events*.

Minimize **probability** *subject to* **risk-adjusted** *constraints:*

$$\min_{z \in \mathcal{Z}} \; p_\tau(U(z)) \quad \text{subject to} \quad \mathcal{R}(J(U(z), z)) \le c_0.$$

Minimize **risk** *subject to* **probabilistic** *constraints:*

$$\min_{z \in \mathcal{Z}} \; \mathcal{R}(J(U(z), z)) \quad \text{subject to} \quad p_\tau(U(z)) \le p_0.$$

Notation: z is the control or design and $U(z)$ is the PDE solution.

Multi-Physics Hierarchy

Unsteady — **Phase 2**

$$\frac{\partial \mathbf{u}}{\partial t} - \nu_1 \Delta \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} + \nabla p - \nu_2 T\mathbf{g} = 0$$
$$\nabla\cdot\mathbf{u} = 0$$
$$\frac{\partial T}{\partial t} - \nu_3 \Delta T + \mathbf{u}\cdot\nabla T = 0$$

$$\frac{\partial T}{\partial t} - \nu_3 \Delta T + \mathbf{u}\cdot\nabla T = 0$$

$$\frac{\partial \mathbf{u}}{\partial t} - \nu_1 \Delta \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} + \nabla p - \nu_2 \mathbf{g} = 0$$
$$\nabla\cdot\mathbf{u} = 0$$

Steady — **Phase 1**

$$-\nu_1 \Delta \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} + \nabla p - \nu_2 T\mathbf{g} = 0$$
$$\nabla\cdot\mathbf{u} = 0$$
$$-\nu_3 \Delta T + \mathbf{u}\cdot\nabla T = 0$$

$$-\nu_3 \Delta T + \mathbf{u}\cdot\nabla T = 0$$

$$-\nu_1 \Delta \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} + \nabla p - \nu_2 \mathbf{g} = 0$$
$$\nabla\cdot\mathbf{u} = 0$$

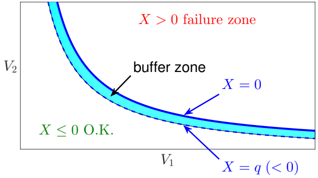
Technical Overview

Risk Quadrangle

Risk	$\mathcal{R} \longleftrightarrow \mathcal{D}$	Deviation
	$\uparrow\downarrow \quad \mathcal{S} \quad \downarrow\uparrow$	
Regret	$\mathcal{V} \longleftrightarrow \mathcal{E}$	Error

- Application specific risk through defintion of regret.
- Rigorously connect optimization and statistical esitmation.
- Develop time-consistent risk measures.

Buffered Probability



- Conservative surrogate for probabilistic computations.
- Nice mathematical properties: quasi-convexity, monotonicity, continuity, etc.
- Efficient optimization formulation.

Outline

Introduction
    Team Members & Expertise
    Computational & Software Infrastructure

Motivating Applications

Technical Challenges
    Risk Mitigation Strategy

Our Approach
    Risk Quadrangle
    Buffered Probability
    Time Consistency
    Optimization-Based Sampling

Technical Challenges

Risk Mitigation Strategy

1. *Large-dimensional space of uncertain parameters*
   - **Optimization-based sampling** generates dimension-independent discrete densities that match desired properties, e.g., distribution tails.
   - **Adaptive discretizations** — in numerical optimization, accuracy is not required far away from a solution.
2. *Rare-event detection and computation*
   - **Buffered probability** provides a numerically tractable and conservative surrogate for probabilistic objectives and constraint.
   - **Optimization-based samples** must account for rare/tail events.
3. *Nonsmooth objective functions and constraints*
   - **Epi-regularized risk quadrangles** are related to trade-off models of risk, regret, deviation, and error.
   - **Higher-moment buffered probability** as a smooth, conservative surrogate for buffered probability computation.
4. *Time consistency for continuous-time systems*
   - **Time-consistent risk measures**: "At every state of the system, optimality of our decisions should not depend on scenarios which we already know cannot happen in the future" (**Shapiro** et al.)
   - Develop time-consistent risk measures using the **risk quadrangle**.

Outline

Introduction
- Team Members & Expertise
- Computational & Software Infrastructure

Motivating Applications

Technical Challenges
- Risk Mitigation Strategy

Our Approach
- Risk Quadrangle
- Buffered Probability
- Time Consistency
- Optimization-Based Sampling

Outline

Introduction
- Team Members & Expertise
- Computational & Software Infrastructure

Motivating Applications

Technical Challenges
- Risk Mitigation Strategy

Our Approach
- Risk Quadrangle
- Buffered Probability
- Time Consistency
- Optimization-Based Sampling

Risk Quadrangle — A New Paradigm

Functionals applied to random variable "costs" X

$$\begin{array}{ccccc} & \text{Risk } \mathcal{R} & \longleftrightarrow & \mathcal{D} \text{ Deviation} & \\ \text{Optimization} & \uparrow\downarrow & \mathcal{S} & \downarrow\uparrow & \text{Statistics} \\ & \text{Regret } \mathcal{V} & \longleftrightarrow & \mathcal{E} \text{ Error} & \end{array}$$

$\mathcal{R}(X)$ provides a numerical surrogate for the cost in X

$\mathcal{D}(X)$ measures the nonconstancy in X

$\mathcal{E}(X)$ measures the nonzeroness in X

$\mathcal{V}(X)$ quantifies the (net) regret in outcomes $X > 0$ versus $X \leq 0$

$\mathcal{S}(X)$ is a "statistic" associated with X through $\mathcal{E}$ as well as $\mathcal{V}$

---

see: *Rockafellar and Uryasev.*
Surveys in Management Science and O.R. 18 (2013)

downloads: www.ise.ufl.edu/uryasev/files/2013/03/quadrangle.pdf

Basic Quadrangle Relationships

- **Regret** and **Error** differ by Expectation

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

- **Risk** and **Deviation** differ by Expectation

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

- **Risk** and **Deviation** obtained from **Regret** and **Error**

$$\mathcal{R}(X) = \min_C \{C + \mathcal{V}(X - C)\}, \qquad \mathcal{D}(X) = \min_C \mathcal{E}(X - C)$$

- **Statistic** is a “byproduct” of **Regret** and **Error** minimization

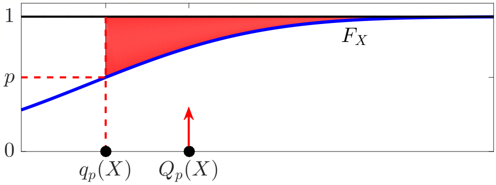
$$\mathcal{S}(X) = \operatorname*{argmin}_C \mathcal{E}(X - C) = \operatorname*{argmin}_C \{C + \mathcal{V}(X - C)\}$$

Example: Mean-Based Quadrangle – Scaling parameter $\lambda > 0$

$\mathcal{E}(X) = \lambda\|X\|_2 = \lambda(\mathbb{E}[X^2])^{1/2}$	$\Longrightarrow$	L^2-error scaled
$\mathcal{S}(X) = \mathbb{E}[X]$	$\Longrightarrow$	mean
$\mathcal{D}(X) = \lambda\sigma(X)$	$\Longrightarrow$	standard deviation, scaled
$\mathcal{R}(X) = \mathbb{E}[X] + \lambda\sigma(X)$	$\Longrightarrow$	safety margin risk
$\mathcal{V}(X) = \mathbb{E}[X] + \lambda\|X\|_2$	$\Longrightarrow$	L^2-regret, scaled

Quantiles and "Superquantiles": VaR and CVaR

F_X = cumulative distribution function for random variable X



Quantile: "value-at-risk" in finance

$$q_p(X) = \mathrm{VaR}_p(X) = F_X^{-1}(p)$$

Superquantile: "conditional value-at-risk" in finance

$$Q_p(X) = \mathrm{CVaR}_p(X) = \mathbb{E}[\,X\,|X \geq q_p(X)\,] = \frac{1}{1-p}\int_p^1 q_t(X)dt$$

Corresponding concepts of "failure":

$q_p(\underline{c}(x)) \leq 0$ probability of failure is $\leq 1 - p$

$Q_p(\underline{c}(x)) \leq 0$ **buffered** probability of failure is $\leq 1 - p$

Example: Quantile-Based Quadrangle

Probability level $p \in (0,1)$

$$\mathcal{E}(X) = \mathbb{E}\left[\frac{p}{1-p}\max\{0,X\} + \max\{0,-X\}\right]$$
$$= \text{Koenker-Basset error, normalized}$$

$$\mathcal{S}(X) = q_p(X) = \mathrm{VaR}_p(X) = \text{quantile}$$

$$\mathcal{D}(X) = Q_p(X - \mathbb{E}[X]) = \mathrm{CVaR}_p(X - \mathbb{E}[X])$$
$$= \text{superquantile deviation}$$

$$\mathcal{R}(X) = Q_p(X) = \mathrm{CVaR}_p(X) = \text{superquantile}$$

$$\mathcal{V}(X) = \frac{1}{1-p}\mathbb{E}[\max\{0,X\}]$$
$$= \text{expected absolute loss, scaled}$$

Applications in Optimization

- Minimizing **Risk** through **Regret**

$$\mathcal{R}(X) = \min_{C\in\mathbb{R}} \left\{C + \mathcal{V}(X - C)\right\}$$

$$\min_{X\in\mathcal{X}} \mathcal{R}(X) = \min_{X\in\mathcal{X}} \min_{C\in\mathbb{R}} \left\{C + \mathcal{V}(X - C)\right\}$$

$$= \min_{X\in\mathcal{X}, C\in\mathbb{R}} \left\{C + \mathcal{V}(X - C)\right\}$$

- Minimizing **Deviation** through **Error**

$$\mathcal{D}(X) = \min_{C\in\mathbb{R}} \mathcal{E}(X - C)$$

$$\min_{X\in\mathcal{X}} \mathcal{D}(X) = \min_{X\in\mathcal{X}} \min_{C\in\mathbb{R}} \mathcal{E}(X - C) = \min_{X\in\mathcal{X}, C\in\mathbb{R}} \mathcal{E}(X - C)$$

- **Statistic** is a "byproduct" of **Regret** and **Error** minimization

$$\mathcal{S}(X) = \operatorname*{argmin}_{C\in\mathbb{R}} \mathcal{E}(X - C) = \operatorname*{argmin}_{C\in\mathbb{R}} \left\{C + \mathcal{V}(X - C)\right\}$$

Example: Minimizing CVaR

$$\mathrm{CVaR}_p(X) = \min_{C \in \mathbb{R}} \left\{ C + \frac{1}{1-p} \mathbb{E}[\max\{0, X - C\}] \right\} \qquad \text{for } p \in (0,1)$$

$$\mathrm{VaR}_p(X) = \operatorname{argmin} \left\{ C + \frac{1}{1-p} \mathbb{E}[\max\{0, X - C\}] \right\}$$

Minimization in $x \in \mathbb{R}^n$

$$\begin{aligned} &\text{minimize} \quad \mathrm{CVaR}_{p_0}(\underline{c}_0(x)) \\ &\text{subject to} \quad \mathrm{CVaR}_{p_i}(\underline{c}_i(x)) \le b_i, \quad i = 1, \dots, m \end{aligned}$$

Minimization in x and auxiliary variables $C_0, C_1, \dots, C_m$

$$\begin{aligned} &\text{minimize} \quad C_0 + \frac{1}{1-p_0} \mathbb{E}[\max\{0, \underline{c}_0(x) - C_0\}] \\ &\text{subject to} \quad C_i + \frac{1}{1-p_i} \mathbb{E}[\max\{0, \underline{c}_i(x) - C_i\}] \le b_i, \quad i = 1, \dots, m \end{aligned}$$

Convex/Linear programming when $c_i(x)$ a linear combination of random variables, or a mixture of distributions (e.g., normal).

Applications in Statistics

Y = dependent random variable
$X_1, \ldots, X_n$ = independent random variables ("factors")

Approximation scheme: $Y \approx f(X_1, \ldots, X_n)$

$\mathcal{C}$ = some specified class of functions $f : \mathbb{R}^n \to \mathbb{R}$

e.g. linear, $f(X_1, \ldots, X_n) = c_0 + c_1 X_1 + \cdots + c_n X_n$

residual: $Z_f = Y - f(X_1, \ldots, X_n)$ for $f \in \mathcal{C}$

Regression

minimize $\mathcal{E}(Z_f)$ over all $f \in \mathcal{C}$ for some error measure $\mathcal{E}$

Standard regression: $\mathcal{E}(Z_f) = (\mathbb{E}[Z_f^2])^{1/2}$ "least squares"

Quantile regression: using $Z^+ = \max\{0, Z\}$, $Z^- = \max\{0, -Z\}$

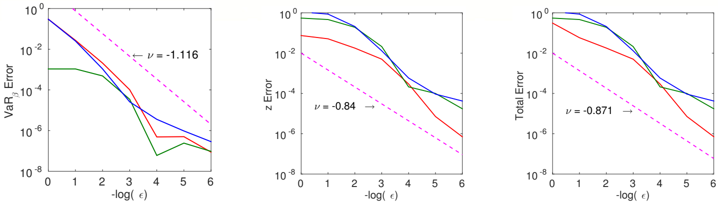
$\mathcal{E}(Z_f) = \mathbb{E}[\frac{p}{1-p} Z_f^+ + Z_f^-]$ at probability level $p \in (0, 1)$

Koenker-Basset error, normalized

Risk Quadrangle

Preliminary Results

- Optimal control of steady, viscid Burgers equation using smoothed CVaR



- **D. P. Kouri** and T. M. Surowiec, “Risk-Averse PDE-Constrained Optimization using the Conditional Value-at-Risk”, SIAM Journal on Optimization, to appear.

Research Tasks

- Develop smooth risk quadrangles to enable risk-averse PDE-optimization.
- Path-following continuation scheme coupled with Newton-type methods.
- Time-consistent, application-specific risk quadrangles for control of thermal fluids.

Outline

Introduction
Team Members & Expertise
Computational & Software Infrastructure

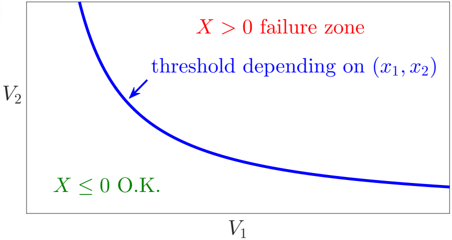
Motivating Applications

Technical Challenges
Risk Mitigation Strategy

Our Approach
Risk Quadrangle
Buffered Probability
Time Consistency
Optimization-Based Sampling

Failure: Standard Engineering Perspective

$X = c(x_1, \ldots, x_n; V_1, \cdots, V_r) =$ "cost" signaling "danger"



Probability of failure: $p_f = \text{prob}\{X > 0\}$

- How to compute or at least estimate?
- How to cope with control variables $x_1, \ldots, x_n$ in optimization?

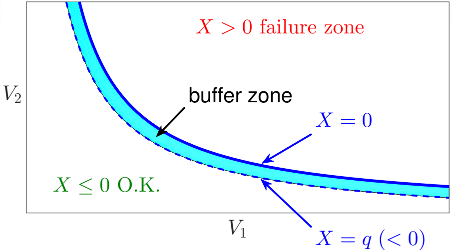
  both p_f and the threshold **shift** with changes in $x_1, \ldots, x_n$!

Troubles with this concept:

- poor mathematical behavior is a serious handicap
- failure probability ignores the **degree** of failure

Buffered Failure: Better Approach to Safety

Utilizing **CVaR (superquantile)** in place of **VaR (quantile)** in reliability



Buffered probability of failure: $P_f = \mathrm{prob}\{X > q\}$

with q determined by $\mathrm{CVaR}_{1-P_f}(X) = \mathbb{E}[X \mid X > q] = 0$

Proposal: use P_f in place of p_f

safer by integrating tail information, and

easier also to work with in computation!

See: *Rockafellar and Royset (2010). On Buffered Failure Probability in Design and Optimization of Structures.*

Definitions: POE and bPOE

X = random variable
$x \in \mathbb{R}$ = threshold
$\alpha \in [0, 1]$ = probability level

Probability of Exeedance (POE) = 1 - Inverse of VaR (Quantile)

$$1 - \{\alpha : q_\alpha(X) = x\} = \mathrm{prob}\{X > x\} = p_x(X)$$

Buffered Probability of Exceedance (bPOE) = 1 - Inverse of CVaR

$$1 - \{\alpha : \mathrm{CVaR}_\alpha(X) = x\} = \bar{p}_x(X)$$

bPOE is an extension of Buffered Probability of Failure for $x \neq 0$.

[1] *Mafusalov and Uryasev (2014). Buffered Probability of Exceedance: Mathematical Properties and Optimization Algorithms.*

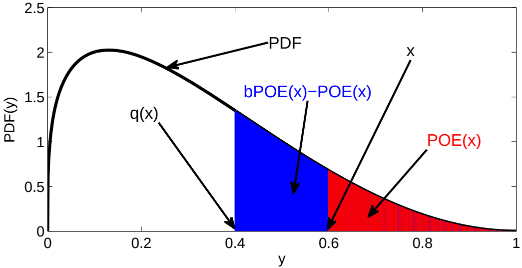
[2] *Norton and Uryasev (2014). Maximization of AUC and Buffered AUC in Classification.*

bPOE Explanation

POE = $\text{prob}\{X > x\}$,

bPOE = $\text{prob}\{X > q(x)\}$, where $q(x)$ satisfies $\mathbb{E}[X|X > q(x))] = x$

"Buffer"= $x - q(x)$



bPOE vs POE

POE characteristics

- POE is concerned with the proportion of events exceeding some threshold $x \in \mathbb{R}$
- DOES NOT consider the magnitude of these events
- Considers only a count of events exceeding the threshold
- Poor mathematical properties: discontinuous!

bPOE characteristics

- bPOE is concerned with the proportion of events, that when considered together, have average magnitude equal to some threshold $x \in \mathbb{R}$
- bPOE is a probability measurement, which takes into account the magnitude of tail events
- Excellent mathematical properties

Why is bPOE important?

POE can "hide" critical information

When the distribution of X is heavy tailed, the magnitude of tail events is important.

Example: Hurricane Damage Data

Threshold (Damage in \$ billions)	POE (%)	bPOE (%)
100	1	3
50	4	10
10	15	69
1	48	100
0.1	79	100

- Hurricane damage data forms a heavy-tailed distribution
- Notice how bPOE reflects heavy-tail (e.g. at threshold \$10 billion)

Preliminary Results

bPOE Mathematical Properties

- bPOE equals a minimum of simple "partial moment" functions! Surprisingly, the minimum value is always between 0 and 1.

$$\bar{p}_x(X) = \min_{\lambda \geq 0} \mathbb{E}[\lambda(X - x) + 1]^+$$

- bPOE is quasi-convex in X w.r.t. addition and concave w.r.t. the mixture operation; it is monotonic w.r.t. X.

- bPOE is a strictly decreasing function of the parameter x.

- $1/\bar{p}_x(X)$ is a convex function of x, and a piecewise-linear function for discrete distributions.

- CVaR and bPOE constraints are equivalent

$$\mathrm{CVaR}_\alpha(X) \leq z \quad \Longleftrightarrow \quad \bar{p}_z(X) \leq 1 - \alpha.$$

Buffered Probability of Exceedance

Preliminary Results

- If $X = \vec{a}^\top \vec{x} + b$ then bPOE opt. reduces to Convex/Linear Programming!
- $\vec{x}$ are opt. variables, $\vec{a}$ is a random vector and b is a random variable

$$\begin{aligned}
\min_{\vec{x}} \; \bar{p}_z(\vec{a}^\top \vec{x} + b) &= \min_{\vec{x}} \min_{\lambda \geq 0} \mathbb{E}\left[\lambda(\vec{a}^\top \vec{x} + b - z) + 1\right]^+ \\
&= \min_{\vec{x}, \lambda \geq 0} \mathbb{E}\left[\vec{a}^\top \lambda \vec{x} + (b - z)\lambda + 1\right]^+ \\
&= \min_{\vec{y}, \lambda \geq 0} \mathbb{E}\left[\vec{a}^\top \vec{y} + (b - z)\lambda + 1\right]^+ .
\end{aligned}$$

Research Tasks

- Prove functional and MC approximation properties of bPOE.
- Smooth bPOE objective to enable use of Newton-type algorithms.
- Developed algorithms to minimize bPOE for nonlinear X.

Outline

Introduction
- Team Members & Expertise
- Computational & Software Infrastructure

Motivating Applications

Technical Challenges
- Risk Mitigation Strategy

Our Approach
- Risk Quadrangle
- Buffered Probability
- Time Consistency
- Optimization-Based Sampling

Dynamic Decision Process

Motivation

- Consider the multistage decision process

$$\text{decision } (x_1) \rightsquigarrow \text{observation } (\xi_2) \rightsquigarrow \text{decision } (x_2) \rightsquigarrow \ldots \rightsquigarrow \text{observation } (\xi_T) \rightsquigarrow \text{decision } (x_T)$$

- **Example:** Time-dependent optimal control of thermal fluids
- $\xi_{[t]} = (\xi_1, \ldots, \xi_t)$ is the history up to time $t \in \{1, \ldots, T\}$
- At time t, the past $\xi_{[t]}$ is observed, but the future is uncertain
- Decision x_t only depends on the history $\xi_{[t]}$, not future observations

Guiding Principle

"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" (**Bellman**)

Time Consistent Risk Measures

Our Approach

- Define *conditional* risk measures $\mathcal{R}_{t,T}$ for every period $t \in \{1, \ldots, T\}$.
- **Time Consistency:** For all $1 \leq \tau < \theta \leq T$, if Z_t and Z_t' such that
  - $Z_t = Z_t'$ for $t = \tau, \ldots, \theta - 1$
  - $\mathcal{R}_{\theta,T}(Z_\theta, \ldots, Z_T) < \mathcal{R}_{\theta,T}(Z_\theta', \ldots, Z_T')$

  then $\mathcal{R}_{\tau,T}(Z_\tau, \ldots, Z_T) < \mathcal{R}_{\tau,T}(Z_\tau', \ldots, Z_T')$.
- Sufficient condition for Bellman's principle of optimality.
- **A. Shapiro**. "Time consistency of dynamic risk measures", Operations Research Letters, vol. 40, pp. 436-439, 2012.
- **A. Shapiro**. "On a time consistency concept in risk averse multi-stage stochastic programming", Operations Research Letters, vol. 37, pp. 143-147, 2009.

Research Tasks

- Extend time consistency theory to continuous time systems.
- Develop application relevant and computationally tractable time-consistent risk measures.

Outline

Introduction

Team Members & Expertise

Computational & Software Infrastructure

Motivating Applications

Technical Challenges

Risk Mitigation Strategy

Our Approach

Risk Quadrangle

Buffered Probability

Time Consistency

Optimization-Based Sampling

Optimization-Based Sampling

Motivation

- Abstract operator equation

$$M(S, X) = 0, \quad S \in \mathcal{S},\ X \in \mathcal{X}.$$

- The probability law of X is known and image of X is high-dimensional.
  We **need** to estimate the probability law of S.

Our Approach

- We seek a discrete random variable $\widetilde{X}$ that is "close" to X with atoms $\mathbf{x} = \{x_1, \ldots, x_m\}$ and corresponding probabilities $\mathbf{p} = \{p_1, \ldots, p_m\}$ satisfying

$$p_k \geq 0 \quad \forall\, k \qquad \text{and} \qquad p_1 + \cdots + p_m = 1.$$

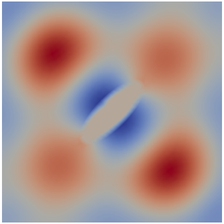
- The atoms **x** and probabilities **p** solve:

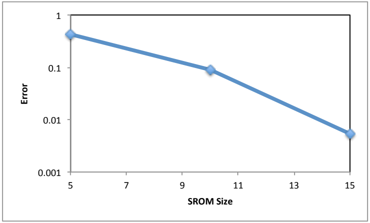
$$\begin{aligned} \underset{\mathbf{x},\,\mathbf{p}}{\text{Minimize}} \quad & \mathcal{H}_X(\mathbf{x}, \mathbf{p}) \\ \text{subject to} \quad & p_k \geq 0 \quad \forall\, k \\ & p_1 + \cdots + p_m = 1 \\ & \mathcal{G}_X(\mathbf{x}, \mathbf{p}) \in \mathcal{F}_X. \end{aligned}$$

Preliminary Results

Problem Setup

- Optimal control of elastic structure in acoustic medium, e.g., DFAT.
- Excitation over four sides using a single frequency: 2Hz.
- Wave speed in fluid $c = 1$.
- Assumed shear and buld moduli with iid beta random variables.
- Minimize CVaR with probability level $p = 0.95$.
- Optimization-based samples: $m \in \{5, 10, 15\}$.





Error: Calculated with respect to 2000 sample Monte Carlo solution.
Clock Time: 10 minutes versus 18 hours for Monte Carlo.

Optimization-Based Sampling

Preliminary Results

m	Objective	CPU Time (sec)	# of iterations
5	0.00551	0.91	5
10	0.00321	1.71	7
50	0.00079	3.92	9
100	0.00034	10.53	13

Table: Approx. of a distribution with 448 atoms by optimizing both atom positions and probabilities.

- B. Ergashev, K. Pavlikov, **S. Uryasev**, and E. Sekeris, "Estimation of Truncated Data Samples in Operational Risk Modeling". Journal of Risk and Insurance, 2015.
- J. E. Warner, M. D. Grigoriu, and **W. Aquino**, "Stochastic Reduced Order Models for Random Vectors: Application to Random Eigenvalue Problems", Probabilistic Engineering Mechanics, vol. 31, 2013.

Research Tasks

- Develop $\mathcal{H}_X$, $\mathcal{G}_X$, $\mathcal{F}_X$ to approx., e.g., tail info. in risk and bPOE computations.
- A priori and a posteriori error estimation will enable adaptive approximation.
- Incorporate adaptive optimization-based sampling with trust-region algorithm to efficiently solve large-scale thermal-fluid control problems.