SAND2017-13377C

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)

DARPA EQUiPS Phase II Program Update

November 29, 2017

Target Optimization Formulations

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

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

$$\min_{z \in Z} p_\tau(S(z)) \quad \text{subject to} \quad \mathcal{R}(J(S(z), z)) \leq c_0$$

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

$$\min_{z \in Z} \mathcal{R}(J(S(z), z)) \quad \text{subject to} \quad p_\tau(S(z)) \leq p_0$$

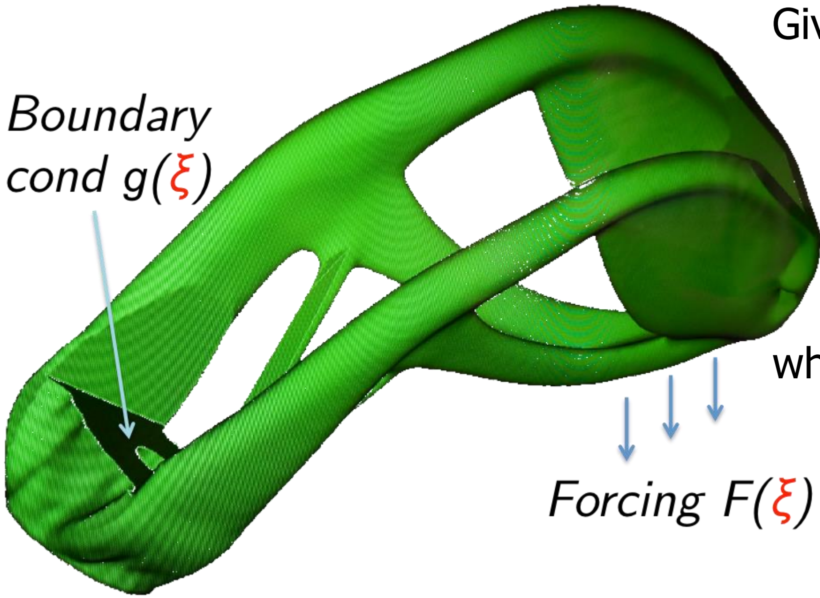
Notation: $z =$ control or design variable

$S(z) =$ solution to system of ODEs/PDEs/DAEs

Tie to EQUiPS Project: Potential to enable risk-averse design of hydrofoils and scram-jet engines

To Date: Risk-averse topology optimization, risk-averse control of chemical vapor deposition reactors

Structural Topology Optimization



Setup: The force 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 compliance tolerance τ, probability $p_0 \in (0,1)$,

$$\min_{0\le z\le 1} \int_D z\,dx =: \mathrm{vol}(z)$$

$$\text{subject to} \quad \mathrm{prob}\left(\left(\int_D F\cdot S(z)\,dx\right)\right) \le p_0$$

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

$$-\nabla\cdot(E(z):\varepsilon u) = F, \qquad \text{in D}$$

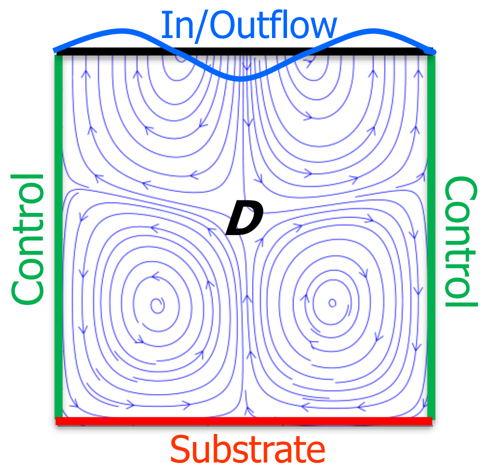
$$\varepsilon u = \tfrac{1}{2}(\nabla u + \nabla u^{\top}), \qquad \text{in } D$$

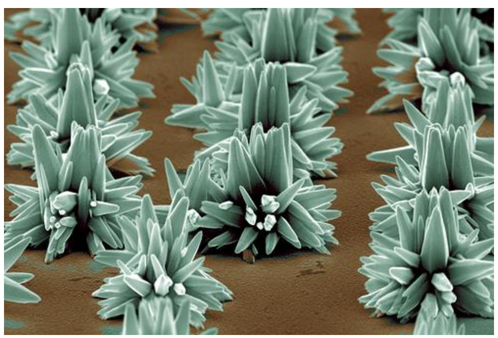
$$u = g, \qquad \text{on } \Gamma_D$$

$$\varepsilon u : n = t, \qquad \text{on } \partial D \setminus \Gamma_D$$

- Uncertain loads, materials, etc. imply displacements are uncertain.
- **Reliability formulation:** Determine light-weight design with *acceptable* probability of failure.
- **Challenge:** Constraints are nonsmooth and expensive to evaluate.

Control of Chemical Vapor Deposition Reactors





Consider the optimal control problem

$$\min_z \frac{1}{2}\mathcal{R}\left(\int_D (\nabla \times U(z))\, dx\right) + \frac{\gamma}{2}\int_{\Gamma_c} |z|^2\, dx$$

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

$$\begin{aligned}
-\nu\nabla^2 u + (u\cdot\nabla)u + \nabla p + \eta t g &= 0 && \text{in } D\\
\nabla\cdot u &= 0 && \text{in } D\\
-\kappa\Delta t + u\cdot\nabla t &= 0 && \text{in } D\\
u - u_i = 0, \qquad t &= 0 && \text{on } \Gamma_i\\
u - u_o = 0, \qquad \kappa\frac{\partial t}{\partial n} &= 0 && \text{on } \Gamma_o\\
u = 0, \qquad T &= T_b && \text{on } \Gamma_b\\
u = 0, \qquad \kappa\frac{\partial t}{\partial n} + h(z - t) &= 0 && \text{on } \Gamma_c
\end{aligned}$$

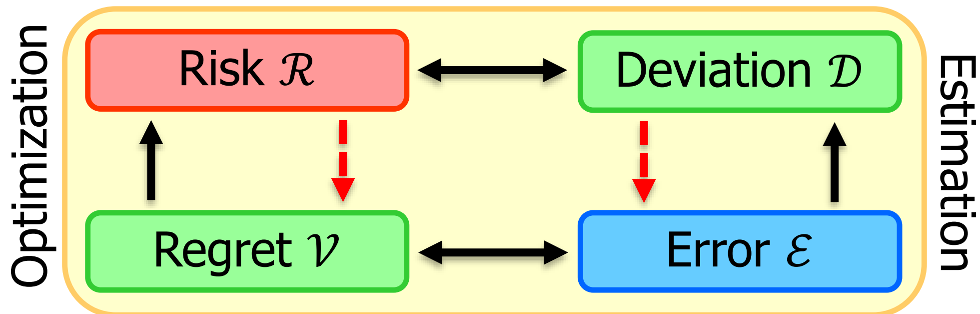
- Uncertain viscosity, thermal conductivity, substrate temperature, etc. imply flow velocity, pressure and temperature are uncertain.
- **Risk-Averse formulation:** Determine wall temperature that minimizes the average of *low-probability*, large vorticity scenarios.
- **Challenge:** Objective function is not differentiable.

Technical Overview

Risk Quadrangle: Connecting optimization and estimation

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

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

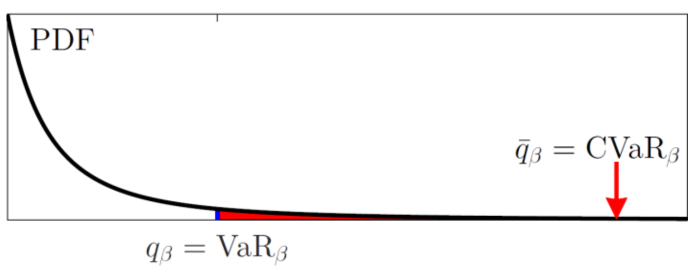


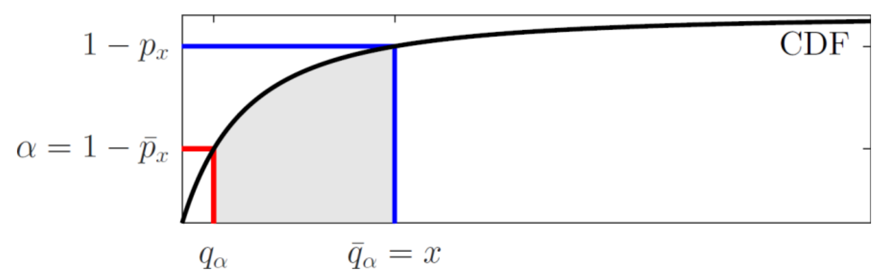
$$\mathcal{D}(X) = \mathcal{R}(X) - \mathbb{E}[X] = \min_t\{\mathcal{E}(X-t)\}$$

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

- $\mathcal{R}$ 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 positive values --- Used to define risk via *disutility*
- $\mathcal{D}$ quantifies **nonconstancy** --- Used to define risk via *variability*

Buffered Probability: Conservative surrogate for probability





- $\mathrm{prob}(X > \tau)$ *counts* failures regardless of degree (i.e., magnitude) of failure
- Buffered probability adds *buffer* region to $\{X > \tau\}$ that accounts for **tail weight**

⟶ Has exceptional mathematical properties and is easy to compute

Phase I Accomplishments

1. **Solved a variety of optimization problems with PDE constraints using risk measures, buffered probability and distributionally robust optimization:**
   - Optimal control of incompressible fluid flow with uncertain viscosity and inflow;
   - Topological design of elastic structures subject to random external loads;
   - Optimal control of a counter flow heat exchanger with random conductivities;
   - Optimal contaminant mitigation subject to random sources and advection.
2. **Our research bridges stochastic programming (financial mathematics) and PDE-constrained optimization (engineering design and control):**
   - Developed application-inspired risk measures and applied to the examples above;
   - Established mathematical theory for buffered probability.
3. **Risk-averse and probabilistic optimization is computationally feasible for large-scale multiphysics applications:**
   - Demonstrated with our PDE-OPT application development kit and the risk-averse optimization capabilities in our optimization libraries (ROL and PSG);
   - Exploited high degree of concurrency in the optimization process by parallelizing over samples and linear algebra;
   - Efficiently and accurately evaluated risk of PDE solution dependent quantities of interest using adaptive reduced basis and risk-informed sampling;
   - Efficiently minimized risk using adaptive discrete density estimation, optimization-based sampling and optimization algorithms that exploit inexactness.

Challenge: Risk and BP are *Not* Differentiable

Example: Optimal control of Burger's equation using CVaR

– **Problem size is small:** 1D spatial domain, 4D stochastic domain
– PDE is nonlinear $\implies$ Objective function is not convex
– CVaR risk measure quantifies *tail weight* and is **not** differentiable

Application of an *off-the-shelf* nonsmooth opt. algorithm:

β	0.1	0.5	0.9
# iter	9,740	10,035	10,128

Required $\mathcal{O}(10^8)$ nonlinear and $\mathcal{O}(10^8)$ linearized PDE solves!

Application of smoothed $\mathcal{R}$ with globalized Newton's Method:

Required $\mathcal{O}(10^6)$ nonlinear and $\mathcal{O}(10^7)$ linearized PDE solves!

Solving real world problems is intractable without…

- Better **nonsmooth** optim. Algorithms or **differentiable** $\mathcal{R}$
- **Adaptive/variable fidelity** approx. in physical/stochastic space

Phase II Accomplishments

A risk measure $\mathcal{R}: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is **coherent** if for $X, X' \in \mathcal{X}$ and $t \in \mathbb{R}$

(C1) **Subadditivity:** $\mathcal{R}(X + X') \le \mathcal{R}(X) + \mathcal{R}(X')$, *i.e., diversifying decreases risk*

(C2) **Monotonicity:** $\mathcal{R}(X) \ge \mathcal{R}(X')$ when $X \ge X'$ a.e., *i.e., small value preference*

(C3) **Translation Equivariance:** $\mathcal{R}(X + t) = \mathcal{R}(X) + t$, *i.e., deterministic is riskless*

(C4) **Positive Homogeneity:** $\mathcal{R}(tX) = t\mathcal{R}(X), \forall t > 0$, *i.e., permit change of units*

$\mathcal{R}$ is **coherent** $\iff$ $\mathcal{R}(X) = \max_{\vartheta \in \mathfrak{A}} \mathbb{E}[\vartheta X]$

$\mathcal{R}(X)$ is a **worst-case** expected value $\longrightarrow$ Often **not** differentiable!

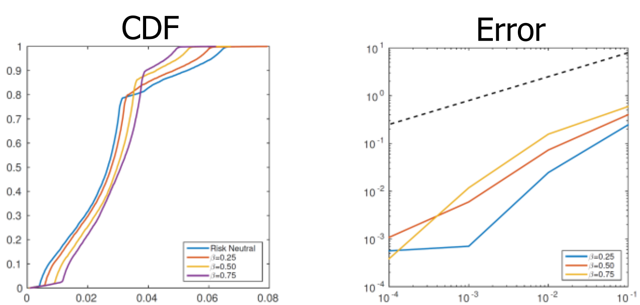
Epi-Regularized Risk Measures:

1. Smooth risk measure by regularizing the worst-case expectation

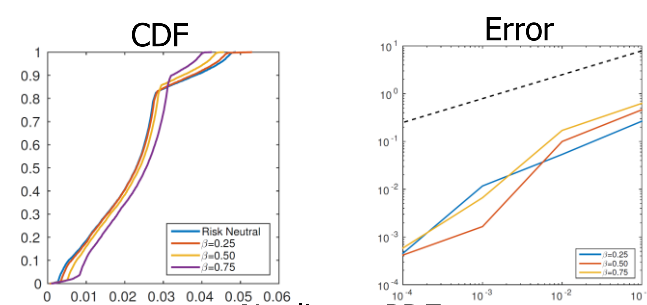
$$\mathcal{R}_\varepsilon(X) := \min_{Y \in \mathcal{X}} \{\mathcal{R}(X - Y) + \varepsilon \Phi(X/\varepsilon)\} = \max_{\vartheta \in \mathfrak{A}} \{\mathbb{E}[\vartheta X] - \varepsilon \Phi^*(\vartheta)\}, \qquad \varepsilon > 0$$

2. $\mathcal{R}_\varepsilon$ satisfies (C1), (C2), (C3), but **not** (C4), i.e., $\mathcal{R}_\varepsilon(tX) = t\mathcal{R}_{\varepsilon/t}(X)$
3. $\mathcal{R}_\varepsilon$ is differentiable $\longrightarrow$ Can minimize using derivative-based optimization

Results: Provable $\mathcal{O}(\sqrt{\varepsilon})$ error, confirmed on nonconvex application!



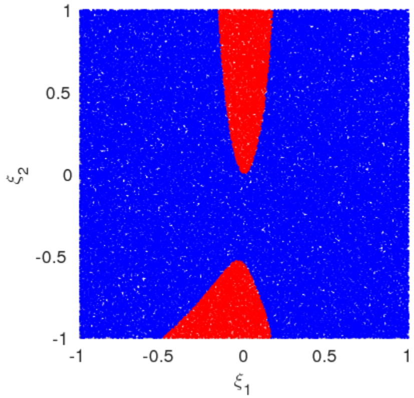
Linear PDE



Nonlinear PDE

Associated with each risk measure $\mathcal{R}$ and r.v. X is a *risk identifier*

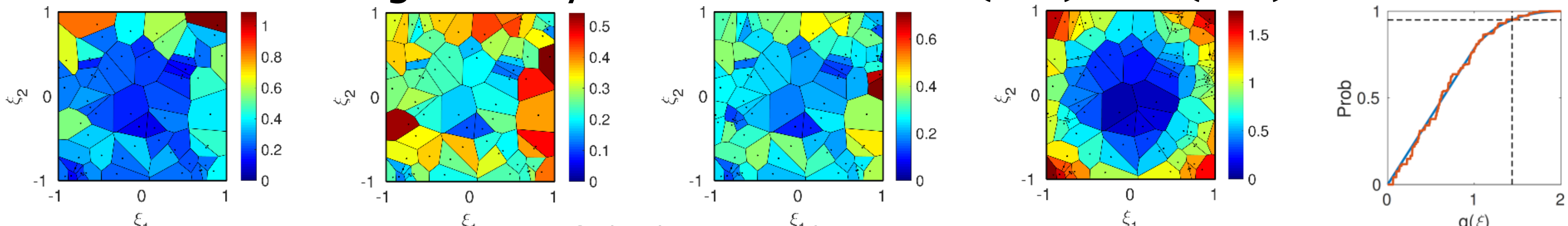
- Illuminates *risky* region in parameter space (red)
- Need accurate approximation of X in risky region
- Use risk identifier to guide adaptive sampling



Adaptive Risk-Informed Modeling:

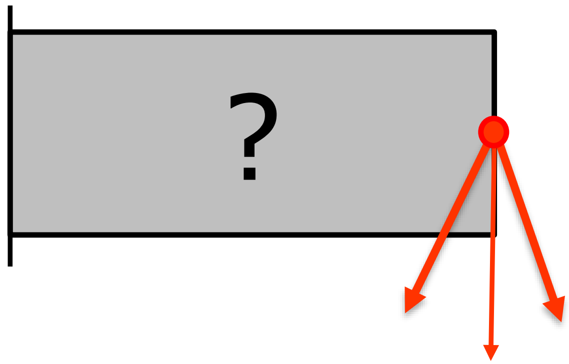
1. Construct local reduced basis models around current samples
$$\bar{u} = \sum_k \mathbb{1}_{\Xi_k} u_k, \text{ where } u_k \text{ is a reduced basis model}$$
2. Evaluate risk-informed QoI error within cell containing sample
$$\left|\mathcal{R}\left(g(u^h)\right) - \mathcal{R}(g(\bar{u}))\right| \leq K \sum_k \max_{\vartheta\in\mathfrak{A}} \mathbb{E}[\vartheta \mathbb{1}_{\Xi_k} \varepsilon_u^\alpha], \ \vartheta \in \mathfrak{A} \text{ are } \textit{risk identifiers}$$
3. Choose new samples based on largest local QoI errors
4. Combine with inexact trust-region algorithm for efficient optimization

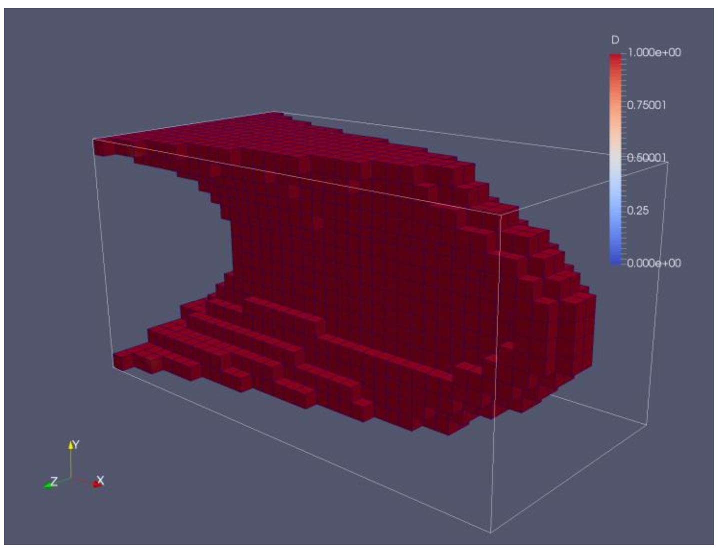
Result: Reduced high-fidelity PDE solves from $\mathcal{O}(10^6)$ to $\mathcal{O}(10^3)$!



3D Results

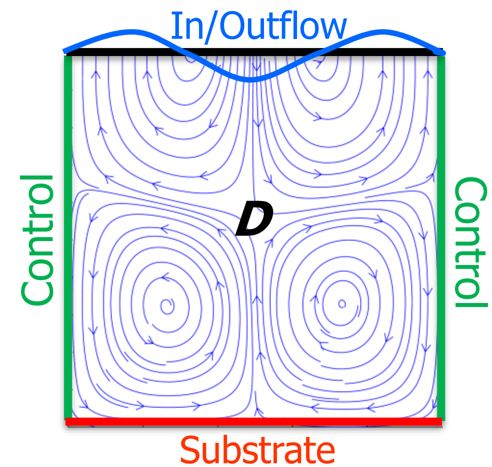
Topology Optimization

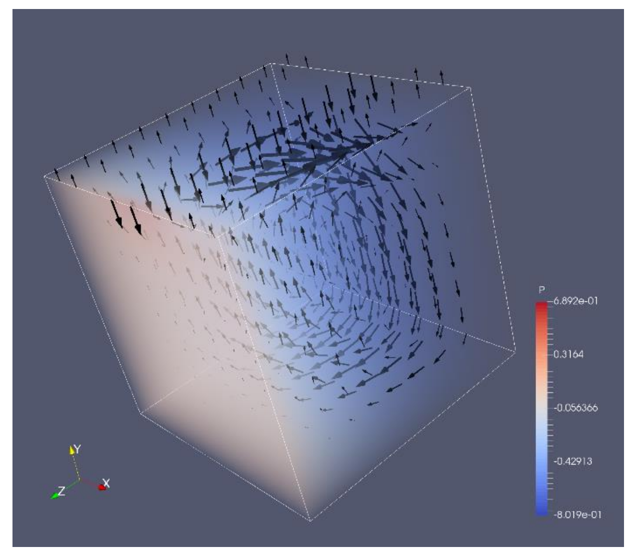




Minimal volume density that satisfies constraint on risk of compliance

Control of CVD Reactor





Controlled velocity (arrows) and pressure (color bar) using entropic risk.

Publications

1. A. Shapiro. Distributionally Robust Stochastic Programming. Technical report, Georgia Institute of Technology, 2016.
2. V. Guigues, V. Krätschmer, and A. Shapiro. Statistical inference and hypotheses testing of risk averse stochastic programs. Technical report, Georgia Institute of Technology, 2016.
3. A. Mafusalov and S. Uryasev. Buffered probability of Exceedance: Mathematical Properties and Optimization Algorithms. Technical Report 2014-1, University of Florida, 2016.
4. A. Mafusalov, A. Shapiro, and S. Uryasev. Estimation and Asymptotics for Buffered Probability of Exceedance. Technical Report 2015-5, University of Florida, 2016.
5. A. Shapiro and K. Ugurlu. Decomposability and time consistency of risk averse multistage programs. Technical report, Georgia Institute of Technology, 2016.
6. D. P. Kouri and T. M. Surowiec. Existence and optimality conditions for risk-averse PDE-constrained optimization. Submitted to SIAM/ASA Journal on Uncertainty Quantification, 2016.
7. A. Shapiro and A. Pichler. Time and Dynamic Consistency of Risk Averse Stochastic Programs. Technical report, Georgia Institute of Technology, 2016.
8. Z. Zou, D. P. Kouri, and W. Aquino. An adaptive sampling approach for solving PDEs with uncertain inputs and evaluating risk. In AIAA SciTech Proceedings, 2016.
9. A. Shapiro. Statistical Inference of Semidefinite Programming. Technical report, Georgia Institute of Technology, 2017.
10. R. T. Rockafellar. Solving Stochastic Programming Problems with Risk Measures by Progressive Hedging. Technical report, University of Washington, 2017.
11. D. P. Kouri and A. Shapiro. Optimization of PDEs with uncertain inputs. In H. Antil, D. P. Kouri, M. Lacasse, and D. Ridzal, editors, IMA Volumes in Mathematics and its Applications. Springer-Verlag, 2017.
12. D. P. Kouri and D. Ridzal. Inexact trust-region methods for PDE-constrained optimization. In H. Antil, D. P. Kouri, M. Lacasse, and D. Ridzal, editors, IMA Volumes in Mathematics and its Applications. Springer-Verlag, 2017.
13. A. Shapiro. Interchangeability principle and dynamic equations in risk averse stochastic programming. Operations Research Letters, 2017.
14. D. P. Kouri and T. M. Surowiec. Epi-regularization of risk measures for PDE-constrained optimization. Submitted to Mathematics of Operations Research, 2017.
15. D. P. Kouri. A measure approximation for distributionally robust PDE-constrained optimization. To appear in SIAM Journal on Numerical Analysis, 2017.
16. D. P. Kouri. Spectral risk measures: The risk quadrangle and optimal approximation. Submitted to Math Programming Series B, 2017.