Scalable Algorithms for Optimal Control of Systems Governed by PDEs Under Uncertainty

Alen Alexanderian¹, Omar Ghattas², Noémi Petra³, Georg Stadler⁴

¹Department of Mathematics
North Carolina State University

²Institute for Computational Engineering and Sciences
The University of Texas at Austin

³School of Natural Sciences
University of California, Merced

³Courant Institute of Mathematical Sciences
New York University

Advances in Uncertainty Quantification Methods, Algorithms and Applications
Annual Meeting of KAUST SRI-UQ Center
King Abdullah University of Science and Technology, Saudi Arabia
January 5–10, 2016
Uncertain parameter $m$

Mathematical Model

$A(m, u) = f$

Quantity of Interest (QoI) $q(m)$
From data to decisions under uncertainty

Uncertain parameter $m$

Bayesian inversion

$$\pi_{\text{post}}(m|y) \propto \pi_{\text{like}}(y|m)\pi_0(m)$$

Experimental data $y$

Mathematical Model

$$A(m,u) = f$$

Quantity of Interest (QoI)

$$q(m)$$

prior

posterior

$$\pi_{\text{like}}(y|m) = \pi_{\text{noise}}(Bu - y)$$

Optimal control under uncertainty

Optimal control/design under uncertainty e.g.

$$\min_{\xi} \int \Psi \left[ \pi_{\text{post}}(m|y;\xi) \right] \pi(y) \, dy$$

$$\xi$$: experimental design

$$\Psi$$: design objective/criterion

Omar Ghattas (ICES, UT Austin)
From data to decisions under uncertainty

**Uncertain parameter** \( m \)

**Bayesian inversion**

\[
\pi_{\text{post}}(m|y) \propto \pi_{\text{like}}(y|m)\pi_0(m)
\]

**Mathematical Model**

\[ A(m, u) = f \]

**Experimental data** \( y \)

**Design of experiments**

\[
\min_{\xi} \int \Psi[\pi_{\text{post}}(m|y; \xi)] \pi(y) dy
\]

\( \xi \): experimental design

\( \Psi \): design objective/criterion

**Quantity of Interest (QoI)**

\[ q(m) \]
From data to decisions under uncertainty

Uncertain parameter $m$

Bayesian inversion

$$\pi_{post}(m|y) \propto \pi_{like}(y|m)\pi_0(m)$$

Experimental data $y$

Mathematical Model

$$A(m,u) = f(\cdot; z)$$

$$z := \text{control function}$$

Quantity of Interest (QoI)

$$q(z,m)$$

Prior

Posterior

$$\pi_{like}(y|m) = \pi_{noise}(Bu - y)$$

Design of experiments

$$\min_{\xi} \int \Psi[\pi_{post}(m|y; \xi)] \pi(y) dy$$

$$\xi: \text{experimental design}$$

$$\Psi: \text{design objective/criterion}$$

Optimal control/design under uncertainty

$$\text{e.g.} \min_{z} \int q(z,m)\mu(dm)$$
Example: Groundwater contaminant remediation

Source: Reed Maxwell, CSM
Example: Groundwater contaminant remediation

- **Inverse problem**
  - Infer (uncertain) soil permeability from (uncertain) measurements of pressure head at wells and from a (uncertain) model of subsurface flow and transport

- **Prediction (or forward) problem**
  - Predict (uncertain) evolution of contaminant concentration at municipal wells from (uncertain) permeability and (uncertain) subsurface flow/transport model

- **Optimal experimental design problem**
  - Where should new observation wells be placed so that permeability is inferred with the least uncertainty?

- **Optimal design problem**
  - Where should new remediation wells be placed so that (uncertain) contaminant concentrations at municipal wells are minimized?

- **Optimal control problem**
  - What should the rates of extraction/injection at remediation wells be so that (uncertain) contaminant concentrations at municipal wells are minimized?
Optimal control of systems governed by PDEs with uncertain parameter fields

PDE-constrained control objective

\[ q = q(u(z, m)) \]

where

\[ \mathcal{A}(u, m) = f(z) \]

- \( q \): control objective
- \( \mathcal{A} \): forward operator
- \( m \): uncertain parameter field
- \( z \): control function

**Problem:** given the uncertainty model for \( m \), find \( z \) that “optimizes” \( q(u(z, m)) \)
Optimization under uncertainty (OUU)

- $\mathcal{H}$: parameter space, infinite-dimensional separable Hilbert space
- $q(z, m)$: control objective functional
- $m \in \mathcal{H}$: uncertain model parameter field, $z$: control function

Optimization under uncertainty (OUU):

$$\min_z q(z, m)$$
Optimization under uncertainty (OUU)

- $\mathcal{H}$: parameter space, infinite-dimensional separable Hilbert space
- $q(z,m)$: control objective functional

  \[ m \in \mathcal{H}: \text{uncertain model parameter field}, \quad z: \text{control function} \]

- Risk-neutral optimization under uncertainty (OUU):

  \[
  \min_z E_m \{ q(z, m) \}
  \]

  \[
  E_m \{ q(z, m) \} = \int_{\mathcal{H}} q(z, m) \mu(dm)
  \]

Main challenges:
Integration over infinite/high-dimensional parameter space
Evaluation of $q$ requires PDE solves
Standard Monte Carlo approach (Sample Average Approximation) is prohibitive
Numerous ($n_{mc}$) samples required, each requires PDE solve
Resulting PDE-constrained optimization problem has $n_{mc}$ PDE constraints
Optimization under uncertainty (OUU)

- $\mathcal{H}$: parameter space, infinite-dimensional separable Hilbert space
- $q(z, m)$: control objective functional
  
  $$m \in \mathcal{H}: \text{uncertain model parameter field}, \quad z: \text{control function}$$

- Risk-averse optimization under uncertainty (OUU):
  
  $$\min_z E_m \{q(z, m)\} + \beta \text{var}_m \{q(z, m)\}$$

  $$E_m \{q(z, m)\} = \int_{\mathcal{H}} q(z, m) \mu(dm)$$

  $$\text{var}_m \{q(z, m)\} = E_m \{q(z, m)^2\} - E_m \{q(z, m)\}^2$$
Optimization under uncertainty (OUU)

- $\mathcal{H}$: parameter space, infinite-dimensional separable Hilbert space
- $q(z, m)$: control objective functional

$$m \in \mathcal{H}: \text{uncertain model parameter field, } z: \text{control function}$$

Risk-averse optimization under uncertainty (OUU):

$$\min_z \mathbb{E}_m \{ q(z, m) \} + \beta \text{var}_m \{ q(z, m) \}$$

$$\mathbb{E}_m \{ q(z, m) \} = \int_{\mathcal{H}} q(z, m) \mu(dm)$$

$$\text{var}_m \{ q(z, m) \} = \mathbb{E}_m \{ q(z, m)^2 \} - \mathbb{E}_m \{ q(z, m) \}^2$$

Main challenges:

- Integration over infinite/high-dimensional parameter space
- Evaluation of $q$ requires PDE solves

Standard Monte Carlo approach (Sample Average Approximation) is prohibitive

- Numerous ($n_{mc}$) samples required, each requires PDE solve
- Resulting PDE-constrained optimization problem has $n_{mc}$ PDE constraints
Some existing approaches for PDE-constrained OUU

Methods based on stochastic collocation, sparse/adaptive sampling, POD, ...

Control of injection wells in a porous medium flow

$m = \text{mean of log permeability field}$

$q = \text{target pressure at production wells}$

- **state PDE:** single phase flow in a porous medium
  \[ -\nabla \cdot (e^m \nabla u) = \sum_{i=1}^{n_c} z_i f_i(x) \]
  with Dirchlet lateral & Neumann top/bottom BCs

- **uncertain parameter:** log permeability field $m$:

- **control variables:** $z_i$, mass source at injection wells; $f_i$, mollified Dirac deltas

- **control objective:** $q(z, m) := \frac{1}{2} ||Qu(z, m) - \bar{q}||^2$, $\bar{q}$: target pressure

- **dimensions:** $n_s = n_m = 3242$, $n_c = 20$, $n_q = 12$
Porous medium with random permeability field

- Distribution law of $m$:

$$\mu = \mathcal{N}(\bar{m}, \mathcal{C}) \quad \text{(Gaussian measure on Hilbert space } \mathcal{H})$$

- Take covariance operator as square of inverse of Poisson-like operator:

$$\mathcal{C} = (-\kappa \Delta + \alpha I)^{-2} \quad \kappa, \alpha > 0$$

- $\mathcal{C}$ is positive, self-adjoint, of trace-class; $\mu$ well-defined on $\mathcal{H}$ (Stuart ’10)

- $\frac{\kappa}{\alpha} \propto$ correlation length; the larger $\alpha$, the smaller the variance

Random draws for $\kappa = 2 \times 10^{-2}, \alpha = 4$
OUU with linearized parameter-to-objective map

- Risk-averse optimal control problem (including cost of controls)

\[
\min_z E_m \{ q(z, m) \} + \beta \text{var}_m \{ q(z, m) \} + \gamma \| z \|^2
\]

- Linear approximation to parameter-to-objective map

\[
q_{\text{lin}}(z, m) = q(z, \bar{m}) + \langle g_m(z, \bar{m}), m - \bar{m} \rangle
\]

- \( g_m(z, \cdot) := \frac{dq(z, \cdot)}{dm} \) is the gradient with respect to \( m \)

- The moments of the linearized objective:

\[
E_m \{ q_{\text{lin}}(z, \cdot) \} = q(z, \bar{m}),
\]

\[
\text{var}_m \{ q_{\text{lin}}(z, \cdot) \} = \langle g_m(z, \bar{m}), C[g_m(z, \bar{m})] \rangle
\]

\[
q_{\text{lin}}(z, \cdot) \sim \mathcal{N} \left( q(z, \bar{m}), \langle g_m(z, \bar{m}), C[g_m(z, \bar{m})] \rangle \right)
\]
Risk-averse optimal control problem with linearized parameter-to-objective map

State-and-adjoint-PDE constrained optimization problem (quartic in $z$):

$$
\min_{z \in Z} \mathcal{J}(z) := \frac{1}{2} \|Q u - \bar{q}\|^2 + \frac{\beta}{2} \langle g_m(\bar{m}), C[g_m(\bar{m})]\rangle + \frac{\gamma}{2} \|z\|^2
$$

with $g_m(\bar{m}) = e^m \nabla u \cdot \nabla p$, where

$$
- \nabla \cdot (e^m \nabla u) = \sum_{i=1}^{n_c} z_i f_i \quad \text{state equation}
$$

$$
- \nabla \cdot (e^m \nabla p) = -Q^*(Q u - \bar{q}) \quad \text{adjoint equation}
$$

Lagrangian of the risk-averse optimal control problem with $q_{\text{lin}}$:

$$
\mathcal{L}(z, u, p, u^*, p^*) = \frac{1}{2} \|Q u - \bar{q}\|^2 + \frac{\beta}{2} \langle e^m \nabla u \cdot \nabla p, C[e^m \nabla u \cdot \nabla p]\rangle + \frac{\gamma}{2} \|z\|^2
$$

$$
+ \langle e^m \nabla u, \nabla u^*\rangle - \sum_{i=1}^{n_c} z_i \langle f_i, u^*\rangle
$$

$$
+ \langle e^m \nabla p, \nabla p^*\rangle + \langle Q^*(Q u - \bar{q}), p^*\rangle
$$
Gradient computation for risk averse optimal control

- “State problem” for risk-averse optimal control problem with $q_{\text{lin}}$:

$$\langle e \tilde{m} \nabla u, \nabla \tilde{u} \rangle = \sum_{i=1}^{n_c} z_i \langle f_i, \tilde{u} \rangle$$

$$\langle e \tilde{m} \nabla p, \nabla \tilde{p} \rangle = -\langle Q^* (Q u - \bar{q}), \tilde{p} \rangle$$

for all test functions $\tilde{p}$ and $\tilde{u}$

- “Adjoint problem” for risk-averse optimal control problem with $q_{\text{lin}}$:

$$\langle e \tilde{m} \nabla p^*, \nabla \tilde{p} \rangle = -\beta \langle e \tilde{m} \nabla u \cdot \nabla \tilde{p}, C[e \tilde{m} \nabla u \cdot \nabla p] \rangle$$

$$\langle e \tilde{m} \nabla u^*, \nabla \tilde{u} \rangle = -\langle Q^* (Q u - \bar{q}), \tilde{u} \rangle - \beta \langle e \tilde{m} \nabla \tilde{u} \cdot \nabla p, C[e \tilde{m} \nabla u \cdot \nabla p] \rangle - \langle Q^* Q p^*, \tilde{u} \rangle$$

for all test functions $\tilde{p}$ and $\tilde{u}$

- Gradient: $ \frac{\partial J}{\partial z_j} = \gamma z_j - \langle f_j, u^* \rangle, \quad j = 1, \ldots, n_c$

- Cost of objective = 2 PDE solves; cost of gradient = 2 PDE solves
Risk-averse optimal control with linearized objective

initial (suboptimal) control $z^0$

distrib. of exact & approx objectives at $z^0$
Risk-averse optimal control with linearized objective

Initial (suboptimal) control \( z^0 \)

Distribution of exact & approx objectives at \( z^0 \)

Optimal control \( z_{\text{lin}}^{\text{opt}} \) based on \( q_{\text{lin}} \)

Distribution of exact & approx objectives at \( z_{\text{lin}}^{\text{opt}} \)
Quadratic approximation to parameter-to-objective map

Quadratic approximation to the parameter-to-control-objective map:

\[ q_{\text{quad}}(z, m) = q(z, \bar{m}) + \langle g_m(z, \bar{m}), m - \bar{m} \rangle + \frac{1}{2} \langle \mathcal{H}_m(z, \bar{m})(m - \bar{m}), m - \bar{m} \rangle \]

- \( g_m \): gradient of parameter-to-objective map
- \( \mathcal{H}_m \): Hessian of parameter-to-objective map

Observations:

- Quadratic approximation does not lead to a Gaussian control objective
- However, can derive analytic formulas for the moments of \( q_{\text{quad}} \) in the infinite-dimensional Hilbert space setting
Mean:
\[ \mathbb{E}_m \{ q_{\text{quad}}(z, \cdot) \} = q(z, \bar{m}) + \frac{1}{2} \text{tr}[\tilde{\mathcal{H}}_m(z, \bar{m})] \]

Variance:
\[ \text{var}_m \{ q_{\text{quad}}(z, \cdot) \} = \langle g_m(z, \bar{m}), \mathbb{C}[g_m(z, \bar{m})] \rangle + \frac{1}{2} \text{tr}[\tilde{\mathcal{H}}_m(z, \bar{m})^2] \]

where \( \tilde{\mathcal{H}}_m = \mathbb{C}^{1/2} \mathcal{H}_m \mathbb{C}^{1/2} \) is the covariance-preconditioned Hessian

Risk averse optimal control objective with \( q_{\text{quad}} \):
\[
J(z) = q(z, \bar{m}) + \frac{1}{2} \text{tr}[\tilde{\mathcal{H}}_m(z, \bar{m})] \\
+ \frac{\beta}{2} \left\{ \langle g_m(z, \bar{m}), \mathbb{C}[g_m(z, \bar{m})] \rangle + \frac{1}{2} \text{tr}[\tilde{\mathcal{H}}_m(z, \bar{m})^2] \right\}
\]
Randomized trace estimator

- **Randomized trace estimation:**

  \[
  \text{tr}(\tilde{H}_m) \approx \frac{1}{n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \tilde{H}_m \xi_j, \xi_j \rangle = \frac{1}{n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle H_m \xi_j, \xi_j \rangle
  \]

  \[
  \text{tr}(\tilde{H}_m^2) \approx \frac{1}{n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle H_m \xi_j, C[H_m \xi_j] \rangle
  \]

  where \( \xi_j = C^{1/2} \xi_j \)

- In computations, we use draws \( \xi_j \sim \mathcal{N}(0, C) =: \nu \)

- Straightforward to show:

  \[
  \int_{\mathcal{H}} \langle H_m \zeta, \zeta \rangle \nu(d\zeta) = \text{tr}(\tilde{H}_m), \quad \int_{\mathcal{H}} \langle H_m \zeta, C[H_m \zeta] \rangle \nu(d\zeta) = \text{tr}(\tilde{H}_m^2)
  \]
Risk-averse optimal control with quadraticized objective

\[
\min_{z \in \mathcal{Z}} \frac{1}{2} \| Qu - \bar{q} \|_2^2 + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, \eta_j \rangle + \frac{\beta}{2} \left\{ \langle g_m(\bar{m}), C[g_m(\bar{m})] \rangle + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \| C^{1/2} \eta_j \|_2^2 \right\}
\]

with

\[
g_m(\bar{m}) = e^\bar{m} \nabla u \cdot \nabla p
\]

\[
\eta_j = e^\bar{m} (\zeta_j \nabla u \cdot \nabla p + \nabla \nu_j \cdot \nabla p + \nabla u \cdot \nabla \rho_j)
\]

\[
\mathcal{H}_m \zeta_j
\]

where

\[
- \nabla \cdot (e^\bar{m} \nabla u) = \sum_{i=1}^{N} z_i f_i
\]

\[
- \nabla \cdot (e^\bar{m} \nabla p) = -Q^* (Qu - \bar{q})
\]

\[
- \nabla \cdot (e^\bar{m} \nabla \nu_j) = \nabla \cdot (\zeta_j e^\bar{m} \nabla u)
\]

\[
- \nabla \cdot (e^\bar{m} \nabla \rho_j) = -Q^* Q \nu_j + \nabla \cdot (\zeta_j e^\bar{m} \nabla p)
\]
Risk-averse optimal control with quadraticized objective

\[
\min_{z \in Z} \frac{1}{2} \|Qu - \bar{q}\|_2^2 + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, \eta_j \rangle + \beta \left\{ \langle g_m(\bar{m}), C[g_m(\bar{m})] \rangle + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \|C^{1/2} \eta_j\|^2 \right\}
\]

with
\[
g_m(\bar{m}) = e^{\bar{m}} \nabla u \cdot \nabla p
\]
\[
\eta_j = e^{\bar{m}} (\zeta_j \nabla u \cdot \nabla p + \nabla \nu_j \cdot \nabla p + \nabla u \cdot \nabla \rho_j)
\]
\[j \in \{1, \ldots, n_{\text{tr}}\}\]

where
\[
- \nabla \cdot (e^{\bar{m}} \nabla u) = \sum_{i=1}^{N} z_i f_i
\]
\[
- \nabla \cdot (e^{\bar{m}} \nabla p) = -Q^* (Qu - \bar{q})
\]
\[
- \nabla \cdot (e^{\bar{m}} \nabla \nu_j) = \nabla \cdot (\zeta_j e^{\bar{m}} \nabla u)
\]
\[
- \nabla \cdot (e^{\bar{m}} \nabla \rho_j) = -Q^* Qu + \nabla \cdot (\zeta_j e^{\bar{m}} \nabla p)
\]
Risk-averse optimal control with quadraticized objective

\[
\min_{z \in Z} \frac{1}{2} \|Qu - \bar{q}\|_2^2 + \frac{1}{2n_{tr}} \sum_{j=1}^{n_{tr}} \langle \xi_j, \eta_j \rangle + \frac{\beta}{2} \left\{ \langle g_m(\bar{m}), C[g_m(\bar{m})] \rangle + \frac{1}{2n_{tr}} \sum_{j=1}^{n_{tr}} \|C^{1/2} \eta_j\|^2 \right\}
\]

with

\[
g_m(\bar{m}) = e^{\bar{m}} \nabla u \cdot \nabla p
\]
\[
\eta_j = e^{\bar{m}} (\xi_j \nabla u \cdot \nabla p + \nabla \nu_j \cdot \nabla p + \nabla u \cdot \nabla \rho_j) \quad j \in \{1, \ldots, n_{tr}\}
\]

where

\[
- \nabla \cdot (e^{\bar{m}} \nabla u) = \sum_{i=1}^{N} z_i \bar{f}_i
\]
\[
- \nabla \cdot (e^{\bar{m}} \nabla p) = -Q^* (Qu - \bar{q})
\]
\[
- \nabla \cdot (e^{\bar{m}} \nabla \nu_j) = \nabla \cdot (\xi_j e^{\bar{m}} \nabla u)
\]
\[
- \nabla \cdot (e^{\bar{m}} \nabla \rho_j) = -Q^* Q \nu_j + \nabla \cdot (\xi_j e^{\bar{m}} \nabla p)
\]
Risk-averse optimal control with quadraticized objective

\[
\min_{z \in Z} \frac{1}{2} \|Qu - \bar{q}\|_2^2 + \frac{1}{2n_{tr}} \sum_{j=1}^{n_{tr}} \langle \zeta_j, \eta_j \rangle + \frac{\beta}{2} \left\{ \langle g_m(\bar{m}), C[g_m(\bar{m})] \rangle + \frac{1}{2n_{tr}} \sum_{j=1}^{n_{tr}} \|C^{1/2} \eta_j\|_2^2 \right\}
\]

with

\[
g_m(\bar{m}) = e^\bar{m} \nabla u \cdot \nabla p
\]

\[
\eta_j = e^\bar{m} (\zeta_j \nabla u \cdot \nabla p + \nabla \nu_j \cdot \nabla p + \nabla u \cdot \nabla \rho_j) \quad j \in \{1, \ldots, n_{tr}\}
\]

where

\[
- \nabla \cdot (e^\bar{m} \nabla u) = \sum_{i=1}^{N} z_i f_i
\]

\[
- \nabla \cdot (e^\bar{m} \nabla p) = -Q^*(Qu - \bar{q})
\]

\[
- \nabla \cdot (e^\bar{m} \nabla \nu_j) = \nabla \cdot (\zeta_j e^\bar{m} \nabla u)
\]

\[
- \nabla \cdot (e^\bar{m} \nabla \rho_j) = -Q^* Q \nu_j + \nabla \cdot (\zeta_j e^\bar{m} \nabla p)
\]
Risk-averse optimal control with quadraticized objective

\[
\min_{z \in Z} \frac{1}{2} \|Q u - \bar{q}\|_2^2 + \frac{1}{2n_{tr}} \sum_{j=1}^{n_{tr}} \langle \zeta_j, \eta_j \rangle + \beta \frac{1}{2} \left\{ \langle g_m(\bar{m}), C[g_m(\bar{m})] \rangle + \frac{1}{2n_{tr}} \sum_{j=1}^{n_{tr}} \|C^{1/2} \eta_j\|^2 \right\}
\]

with

\[
g_m(\bar{m}) = e^{\bar{m}} \nabla u \cdot \nabla p
\]

\[
\eta_j = e^{\bar{m}} (\zeta_j \nabla u \cdot \nabla p + \nabla \nu_j \cdot \nabla p + \nabla u \cdot \nabla \rho_j)
\]

\[\mathcal{H}_m \zeta_j\]

where

\[
- \nabla \cdot (e^{\bar{m}} \nabla u) = \sum_{i=1}^{N} z_i f_i
\]

\[
- \nabla \cdot (e^{\bar{m}} \nabla p) = -Q^* (Qu - \bar{q})
\]

\[
- \nabla \cdot (e^{\bar{m}} \nabla \nu_j) = \nabla \cdot (\zeta_j e^{\bar{m}} \nabla u)
\]

\[
- \nabla \cdot (e^{\bar{m}} \nabla \rho_j) = -Q^* Q \nu_j + \nabla \cdot (\zeta_j e^{\bar{m}} \nabla p)
\]
Lagrangian for risk-averse optimal control with $q_{\text{quad}}$

\[ \mathcal{L}(z, u, p, \{v_j\}_{j=1}^{n_{\text{tr}}}, \{\rho_j\}_{j=1}^{n_{\text{tr}}}, u^*, p^*, \{v_j^*\}_{j=1}^{n_{\text{tr}}}, \{\rho_j^*\}_{j=1}^{n_{\text{tr}}}) \]
\[ = \frac{1}{2} \|Qu - \bar{q}\|_2^2 \]
\[ + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, [e^m(\zeta_j \nabla u \cdot \nabla p + \nabla v_j \cdot \nabla p + \nabla u \cdot \nabla \rho_j)] \rangle \]
\[ + \frac{\beta}{2} \langle e^m \nabla u \cdot \nabla p, C[e^m \nabla u \cdot \nabla p] \rangle \]
\[ + \frac{\beta}{4n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \left\| C^{1/2} \left[ e^m(\zeta_j \nabla u \cdot \nabla p + \nabla v_j \cdot \nabla p + \nabla u \cdot \nabla \rho_j) \right] \right\|^2 \]
\[ + \langle e^m \nabla u, \nabla u^* \rangle - \sum_{i=1}^{N} z_i \langle f_i, u^* \rangle \]
\[ + \langle e^m \nabla p, \nabla p^* \rangle + \langle Q^*(Qu - \bar{q}), p^* \rangle \]
\[ + \sum_{j=1}^{n_{\text{tr}}} \left[ \langle e^m \nabla v_j, \nabla v_j^* \rangle + \langle \zeta_j e^m \nabla u, \nabla v_j^* \rangle \right] \]
\[ + \sum_{j=1}^{n_{\text{tr}}} \left[ \langle e^m \nabla \rho_j, \nabla \rho_j^* \rangle + \langle Q^* Q v_j, \rho_j^* \rangle + \langle \zeta_j e^m \nabla p, \nabla \rho_j^* \rangle \right] \]
Adjoint & gradient for risk-averse optimal control w/ $q_{\text{quad}}$

**Adjoint problem for $q_{\text{quad}}$ approximation**

\[- \nabla \cdot (e\bar{m} \nabla \rho^*_j) = b_1^{(j)} \]
\[- \nabla \cdot (e\bar{m} \nabla v^*_j) + Q^* Q \rho^*_j = b_2^{(j)} \]
\[- \nabla \cdot (e\bar{m} \nabla p^*) - \sum_{j=1}^{n_{tr}} \nabla \cdot (\zeta_j e\bar{m} \nabla \rho^*_j) = b_3 \]
\[- \nabla \cdot (e\bar{m} \nabla u^*) + Q^* Q p^* - \sum_{j=1}^{n_{tr}} \nabla \cdot (\zeta_j e\bar{m} \nabla v^*_j) = b_4 \]

**Gradient for $q_{\text{quad}}$ approximation**

\[ \frac{\partial \mathcal{L}}{\partial z_j} = \gamma z_j - \langle f_j, u^* \rangle, \quad j = 1, \ldots, n_c \]

Cost of objective = $2 + 2n_{tr}$ PDE solves; cost of gradient = $2 + 2n_{tr}$ PDE solves
Risk-averse optimal control with quadraticized objective

initial (suboptimal) control $z^0$

distribution of exact & approx objectives at $z^0$
Risk-averse optimal control with quadraticized objective

initial (suboptimal) control $z^0$

distrib. of exact & approx objectives at $z^0$

optimal control $z_{\text{quad}}^{\text{opt}}$ based on $q_{\text{quad}}$

distrib. of exact & approx objectives at $z_{\text{quad}}^{\text{opt}}$
Effect of number of trace estimator vectors on distribution of control objective evaluated at optimal controls

- Optimal controls $z_{\text{quad}}^{\text{opt}}$ computed for each value of trace estimator using quadratic approximation of control objective, $q_{\text{quad}}$
- Each curve based on 10,000 samples of distribution of $q(z_{\text{quad}}^{\text{opt}}, m)$ (control objective evaluated at optimal control)
Comparison of the distributions of $q(z_{\text{lin}}^{\text{opt}}, m)$ with $q(z_{\text{quad}}^{\text{opt}}, m)$

- $\beta = 1$, $\gamma = 10^{-5}$ and $n_{\text{tr}} = 40$ trace estimation vectors
- KDE results are based on 10,000 samples
- Inserts show Monte Carlo sample convergence for mean and variance
Some concluding remarks:

- Optimal control of PDEs with infinite-dimensional parameters
- Use of local approximations to parameter-to-objective map
- Quadratic approximation is effective in capturing the distribution of the control objective for the target problem
- Computational complexity of objective & gradient evaluation is independent of problem dimension \( n_m = 3242 \)
  - objective + gradient cost = \( \sim 125 \) PDE solves
  - objective/gradient cost for Monte Carlo/SAA: \( \sim 10,000 \) PDE solves
  - differences even more striking for nonlinear PDE state problems

Possible extensions:

- Higher moments (e.g., for forward UQ)
- Alternative risk measures
- Higher order approximations, e.g., third order Taylor expansion with proper tensor contraction of third order derivative tensor