Efficient PDE-constrained optimization under uncertainty using adaptive model reduction and sparse grids

Matthew J. Zahr[†], Kevin Carlberg, Drew Kouri

MS145: Reduced Order Modeling Techniques in Large Scale & Data-Driven PDE Problems
SIAM Conference on Computational Science and Engineering
Hilton Atlanta, Atlanta, Georgia, USA
February 27 - March 3, 2017

[†] Luis W. Alvarez Postdoctoral Fellow
Department of Mathematics
Lawrence Berkeley National Laboratory
University of California, Berkeley

Energetically optimal flapping motions

Goal: Find energetically optimal flapping motion that achieves zero thrust

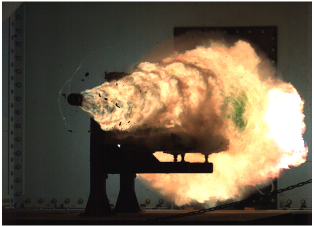
Energy = 1.4459e-01	Energy = 3.1378e-01
Thrust = -1.1192e-01	Thrust = 0.0000e+00

[Zahr and Persson, 2017]

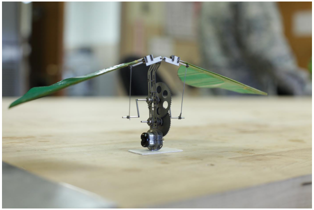
PDE optimization – a key player in next-gen problems

*Current interest in **computational physics** reaches far beyond analysis of a single configuration of a physical system into **design** (shape and topology) and **control** in an **uncertain** setting*

Engine System



EM Launcher



Micro-Aerial Vehicle

Repeated queries to **high-fidelity simulations** required by optimization and uncertainty quantification may be **prohibitively time-consuming**

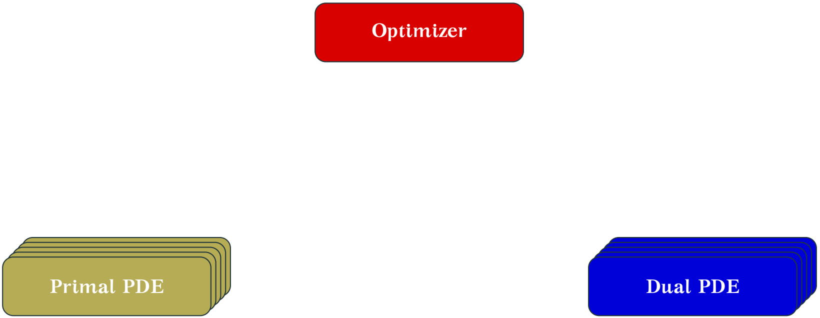
Stochastic PDE-constrained optimization formulation

$$\begin{aligned} \underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad & \mathbb{E}[\mathcal{J}(\mathbf{u}, \boldsymbol{\mu}, \cdot)] \\ \text{subject to} \quad & \mathbf{r}(\mathbf{u}; \boldsymbol{\mu}, \boldsymbol{\xi}) = 0 \quad \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \end{aligned}$$

- $\mathbf{r} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_\mu} \times \mathbb{R}^{n_\xi} \to \mathbb{R}^{n_u}$ — discretized stochastic PDE
- $\mathcal{J} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_\mu} \times \mathbb{R}^{n_\xi} \to \mathbb{R}$ — quantity of interest
- $\mathbf{u} \in \mathbb{R}^{n_u}$ — PDE state vector
- $\boldsymbol{\mu} \in \mathbb{R}^{n_\mu}$ — (deterministic) optimization parameters
- $\boldsymbol{\xi} \in \mathbb{R}^{n_\xi}$ — stochastic parameters
- $\mathbb{E}[\mathcal{F}] \equiv \int_{\boldsymbol{\Xi}} \mathcal{F}(\boldsymbol{\xi}) \rho(\boldsymbol{\xi}) \, d\boldsymbol{\xi}$

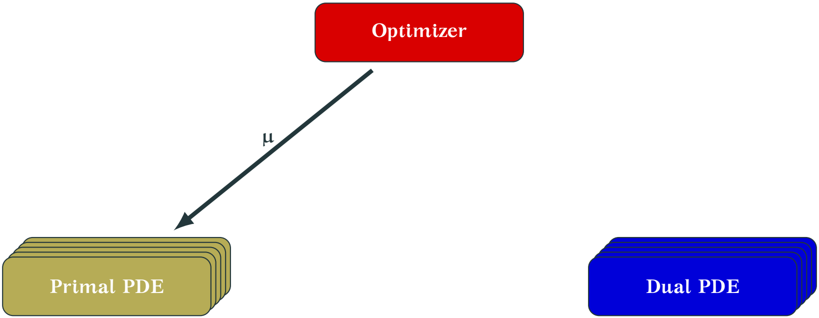
Nested approach to stochastic PDE-constrained optimization

Ensemble *of primal/dual PDE solves required at* ***every*** *optimization iteration*



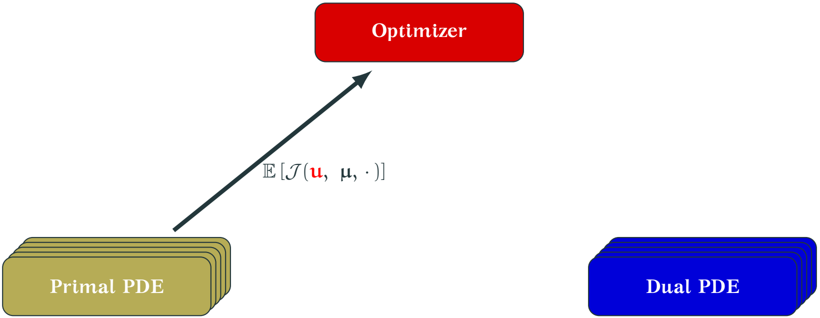
Nested approach to stochastic PDE-constrained optimization

Ensemble *of primal/dual PDE solves required at* ***every*** *optimization iteration*



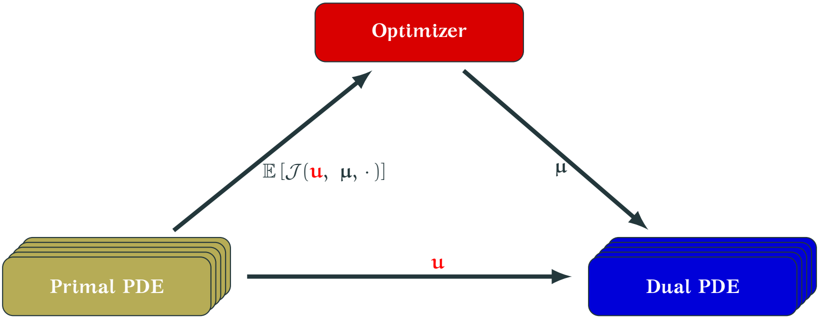
Nested approach to stochastic PDE-constrained optimization

Ensemble *of primal/dual PDE solves required at* ***every*** *optimization iteration*



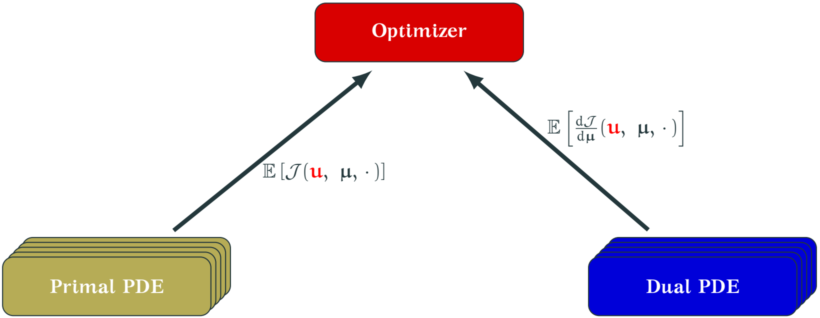
Nested approach to stochastic PDE-constrained optimization

Ensemble of primal/dual PDE solves required at *every* optimization iteration



Nested approach to stochastic PDE-constrained optimization

Ensemble *of primal/dual PDE solves required at* ***every*** *optimization iteration*



Proposed approach: managed inexactness

Replace expensive PDE with inexpensive approximation model

- **Anisotropic sparse grids** used for *inexact integration* of risk measures
- **Reduced-order models** used for *inexact PDE evaluations*

$$\underset{\boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad F(\boldsymbol{\mu}) \qquad \longrightarrow \qquad \underset{\boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad m(\boldsymbol{\mu})$$

Proposed approach: managed inexactness

Replace expensive PDE with inexpensive approximation model

- **Anisotropic sparse grids** used for *inexact integration* of risk measures
- **Reduced-order models** used for *inexact PDE evaluations*

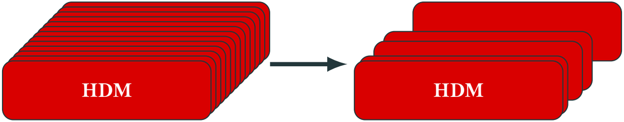
$$\underset{\boldsymbol{\mu} \in \mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad F(\boldsymbol{\mu}) \qquad \longrightarrow \qquad \underset{\boldsymbol{\mu} \in \mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad m(\boldsymbol{\mu})$$

Proposed approach: managed inexactness

Replace expensive PDE with inexpensive approximation model

- **Anisotropic sparse grids** used for *inexact integration* of risk measures
- **Reduced-order models** used for *inexact PDE evaluations*

$$\underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad F(\mu) \qquad \longrightarrow \qquad \underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad m(\mu)$$



Proposed approach: managed inexactness

Replace expensive PDE with inexpensive approximation model

- **Anisotropic sparse grids** used for *inexact integration* of risk measures
- **Reduced-order models** used for *inexact PDE evaluations*

$$\underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad F(\mu) \qquad \longrightarrow \qquad \underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad m(\mu)$$

Proposed approach: managed inexactness

Replace expensive PDE with inexpensive approximation model

- **Anisotropic sparse grids** used for *inexact integration* of risk measures
- **Reduced-order models** used for *inexact PDE evaluations*

$$\underset{\boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad F(\boldsymbol{\mu}) \qquad \longrightarrow \qquad \underset{\boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad m(\boldsymbol{\mu})$$

Manage inexactness with trust region method

- Embedded in globally convergent **trust region** method
- **Error indicators**[1] to account for *all* sources of inexactness
- **Refinement** of approximation model using *greedy algorithms*

$$\underset{\boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad F(\boldsymbol{\mu}) \qquad \longrightarrow \qquad \begin{aligned} &\underset{\boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} && m_k(\boldsymbol{\mu}) \\ &\text{subject to} && \|\boldsymbol{\mu}-\boldsymbol{\mu}_k\| \leq \Delta_k \end{aligned}$$

[1]Must be *computable* and apply to general, nonlinear PDEs

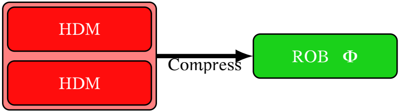
Trust region framework for optimization with ROMs

Schematic

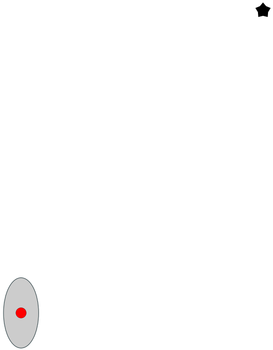
μ-space

Breakdown of Computational Effort

Trust region framework for optimization with ROMs



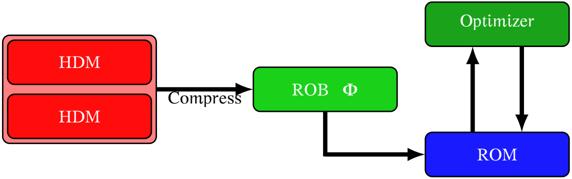
Schematic



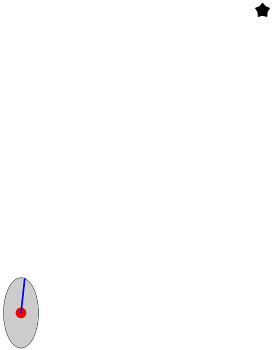
μ-space

Breakdown of Computational Effort

Trust region framework for optimization with ROMs



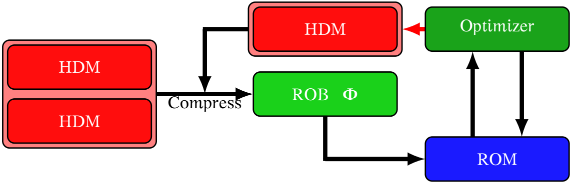
Schematic



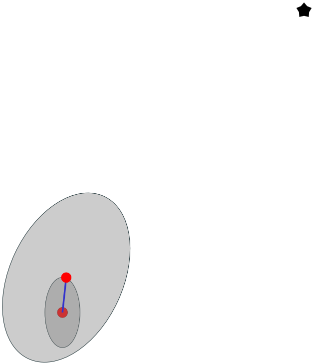
μ-space

Breakdown of Computational Effort

Trust region framework for optimization with ROMs



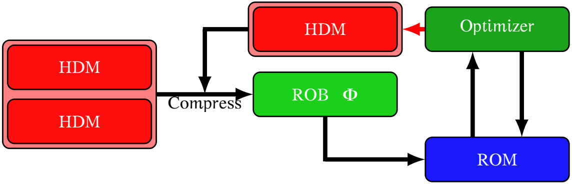
Schematic



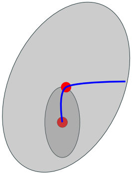
μ-space

Breakdown of Computational Effort

Trust region framework for optimization with ROMs



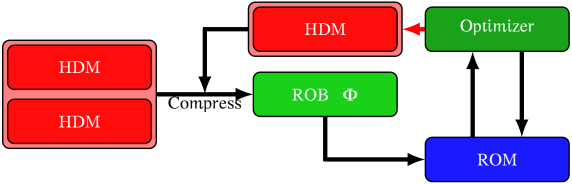
Schematic



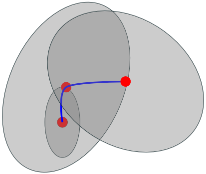
μ-space

Breakdown of Computational Effort

Trust region framework for optimization with ROMs



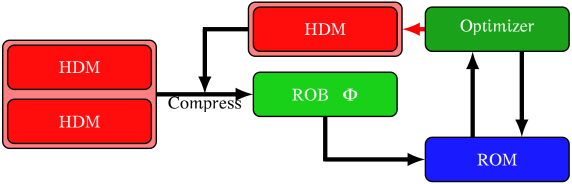
Schematic



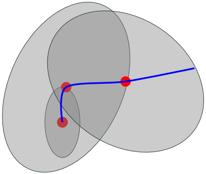
μ-space

Breakdown of Computational Effort

Trust region framework for optimization with ROMs



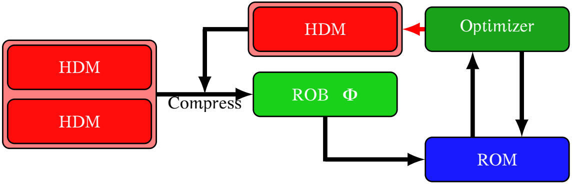
Schematic



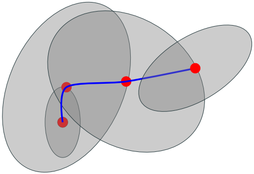
μ-space

Breakdown of Computational Effort

Trust region framework for optimization with ROMs



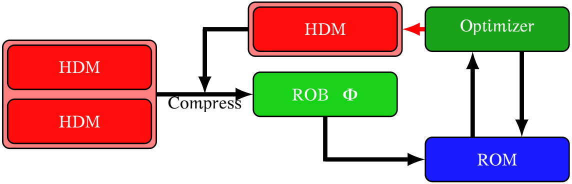
Schematic



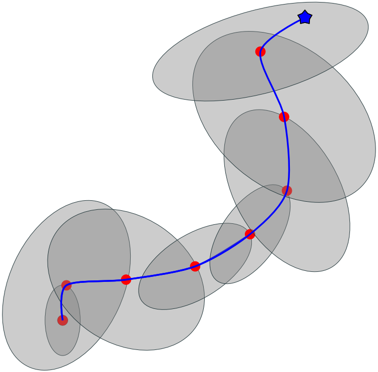
μ-space

Breakdown of Computational Effort

Trust region framework for optimization with ROMs



Schematic



μ-space

Breakdown of Computational Effort

Proposed approach: managed inexactness

Replace expensive PDE with inexpensive approximation model

- **Anisotropic sparse grids** used for *inexact integration* of risk measures
- **Reduced-order models** used for *inexact PDE evaluations*

$$\underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad F(\mu) \qquad \longrightarrow \qquad \underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad m(\mu)$$

Manage inexactness with trust region method

- Embedded in globally convergent **trust region** method
- **Error indicators**[2] to account for *all* sources of inexactness
- **Refinement** of approximation model using *greedy algorithms*

$$\underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad F(\mu) \qquad \longrightarrow \qquad \begin{array}{ll} \underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} & m_k(\mu) \\ \text{subject to} & \|\mu - \mu_k\| \leq \Delta_k \end{array}$$

[2] Must be *computable* and apply to general, nonlinear PDEs

First source of inexactness: anisotropic sparse grids

Stochastic collocation using anisotropic sparse grid nodes to approximate integral with summation

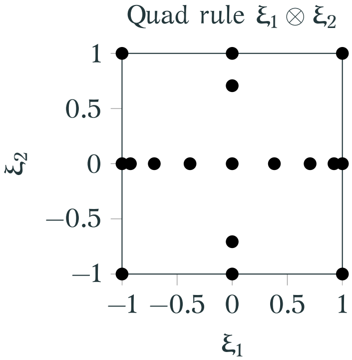
$$\begin{array}{ll}\underset{\mathbf{u}\in\mathbb{R}^{n_{\mathbf{u}}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} & \mathbb{E}[\mathcal{J}(\mathbf{u},\ \boldsymbol{\mu},\ \cdot\,)] \\ \text{subject to} & \mathbf{r}(\mathbf{u},\ \boldsymbol{\mu},\ \boldsymbol{\xi}) = 0 \quad \forall\boldsymbol{\xi}\in\boldsymbol{\Xi}\end{array}$$

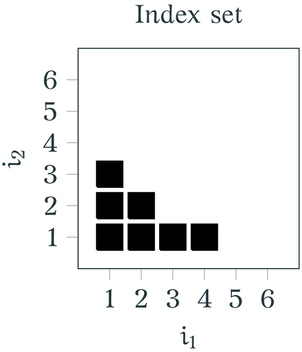
$$\Downarrow$$

$$\begin{array}{ll}\underset{\mathbf{u}\in\mathbb{R}^{n_{\mathbf{u}}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} & \mathbb{E}_{\mathcal{I}}[\mathcal{J}(\mathbf{u},\ \boldsymbol{\mu},\ \cdot\,)] \\ \text{subject to} & \mathbf{r}(\mathbf{u},\ \boldsymbol{\mu},\ \boldsymbol{\xi}) = 0 \quad \forall\boldsymbol{\xi}\in\boldsymbol{\Xi}_{\mathcal{I}}\end{array}$$

[Kouri et al., 2013, Kouri et al., 2014]

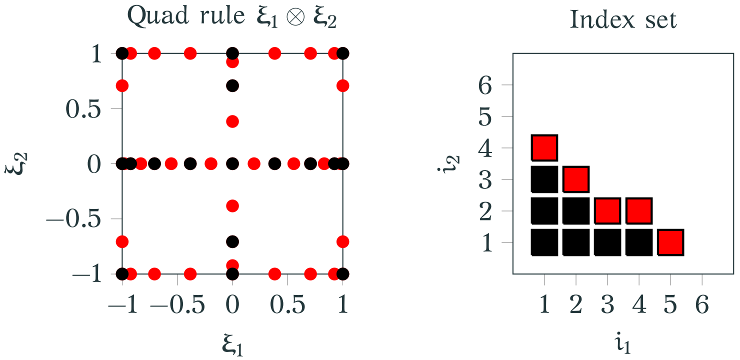
Source of inexactness: anisotropic sparse grids





Index set ($\mathcal{I}$) – ●   Neighbors ($\mathcal{N}(\mathcal{I})$) – ●

Source of inexactness: anisotropic sparse grids



Index set ($\mathcal{I}$) – ● Neighbors ($\mathcal{N}(\mathcal{I})$) – ●

Second source of inexactness: reduced-order models

Stochastic collocation of the reduced-order model over anisotropic sparse grid nodes used to approximate integral with cheap summation

$$
\begin{array}{ll}
\underset{\mathbf{u}\in\mathbb{R}^{n_\mathbf{u}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} & \mathbb{E}[\mathcal{J}(\mathbf{u},\ \boldsymbol{\mu},\ \cdot\ )] \\
\text{subject to} & \mathbf{r}(\mathbf{u},\ \boldsymbol{\mu},\ \boldsymbol{\xi}) = 0 \quad \forall \boldsymbol{\xi} \in \boldsymbol{\Xi}
\end{array}
$$

$$\Downarrow$$

$$
\begin{array}{ll}
\underset{\mathbf{u}\in\mathbb{R}^{n_\mathbf{u}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} & \mathbb{E}_{\mathcal{I}}[\mathcal{J}(\mathbf{u},\ \boldsymbol{\mu},\ \cdot\ )] \\
\text{subject to} & \mathbf{r}(\mathbf{u},\ \boldsymbol{\mu},\ \boldsymbol{\xi}) = 0 \quad \forall \boldsymbol{\xi} \in \boldsymbol{\Xi}_{\mathcal{I}}
\end{array}
$$

$$\Downarrow$$

$$
\begin{array}{ll}
\underset{\mathbf{u}_r\in\mathbb{R}^{k_\mathbf{u}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} & \mathbb{E}_{\mathcal{I}}[\mathcal{J}(\boldsymbol{\Phi}\mathbf{u}_r,\ \boldsymbol{\mu},\ \cdot\ )] \\
\text{subject to} & \boldsymbol{\Phi}^T\mathbf{r}(\boldsymbol{\Phi}\mathbf{u}_r,\ \boldsymbol{\mu},\ \boldsymbol{\xi}) = 0 \quad \forall \boldsymbol{\xi} \in \boldsymbol{\Xi}_{\mathcal{I}}
\end{array}
$$

Source of inexactness: projection-based model reduction

- Model reduction ansatz: *state vector lies in low-dimensional subspace*

$$\mathbf{u} \approx \mathbf{\Phi} \mathbf{u}_r$$

  - $\mathbf{\Phi} = \begin{bmatrix} \boldsymbol{\phi}^1 & \cdots & \boldsymbol{\phi}^{k_u} \end{bmatrix} \in \mathbb{R}^{n_u \times k_u}$ is the reduced (trial) basis ($n_u \gg k_u$)
  - $\mathbf{u}_r \in \mathbb{R}^{k_u}$ are the reduced coordinates of $\mathbf{u}$

- Substitute into $\mathbf{r}(\mathbf{u}, \boldsymbol{\mu}) = 0$ and perform Galerkin projection

$$\mathbf{\Phi}^T \mathbf{r}(\mathbf{\Phi} \mathbf{u}_r, \boldsymbol{\mu}) = 0$$

Proposed approach: managed inexactness

Replace expensive PDE with inexpensive approximation model

- **Anisotropic sparse grids** used for *inexact integration* of risk measures
- **Reduced-order models** used for *inexact PDE evaluations*

$$\underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad F(\mu) \qquad \longrightarrow \qquad \underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad m(\mu)$$

Manage inexactness with trust region method

- Embedded in globally convergent **trust region** method
- **Error indicators**[3] to account for *all* sources of inexactness
- **Refinement** of approximation model using *greedy algorithms*

$$\underset{\boldsymbol{\mu} \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad F(\boldsymbol{\mu}) \qquad \longrightarrow \qquad \begin{aligned} &\underset{\boldsymbol{\mu} \in \mathbb{R}^{n_\mu}}{\text{minimize}} && m_k(\boldsymbol{\mu}) \\ &\text{subject to} && \|\boldsymbol{\mu} - \boldsymbol{\mu}_k\| \leq \Delta_k \end{aligned}$$

[3] Must be *computable* and apply to general, nonlinear PDEs

Trust region ingredients for global convergence

$$\underset{\mu\in\mathbb{R}^{n_\mu}}{\text{minimize}} \quad F(\mu) \qquad \longrightarrow \qquad \begin{array}{ll} \underset{\mu\in\mathbb{R}^{n_\mu}}{\text{minimize}} & m_k(\mu) \\ \text{subject to} & \|\mu-\mu_k\| \leq \Delta_k \end{array}$$

Approximation models

$$m_k(\mu),\ \psi_k(\mu)$$

Error indicators

$$\begin{aligned} \|\nabla F(\mu) - \nabla m_k(\mu)\| &\leq \xi \varphi_k(\mu) \qquad && \xi > 0 \\ |F(\mu_k) - F(\mu) + \psi_k(\mu) - \psi_k(\mu_k)| &\leq \sigma \theta_k(\mu) \qquad && \sigma > 0 \end{aligned}$$

Adaptivity

$$\begin{aligned} \varphi_k(\mu_k) &\leq \kappa_\varphi \min\{\|\nabla m_k(\mu_k)\|,\ \Delta_k\} \\ \theta_k(\hat{\mu}_k)^\omega &\leq \eta \min\{m_k(\mu_k) - m_k(\hat{\mu}_k),\ r_k\} \end{aligned}$$

Trust region method with inexact gradients and objective

1: **Model update**: Choose model m_k and error indicator φ_k

$$\varphi_k(\mu_k) \leq \kappa_\varphi \min\{\|\nabla m_k(\mu_k)\|,\, \Delta_k\}$$

2: **Step computation**: Approximately solve the trust region subproblem

$$\hat{\mu}_k \quad = \quad \underset{\mu \in \mathbb{R}^{n_\mu}}{\arg\min}\; m_k(\mu) \quad \text{subject to} \quad \|\mu - \mu_k\| \leq \Delta_k$$

3: **Step acceptance**: Compute approximation of actual-to-predicted reduction

$$\rho_k = \frac{\psi_k(\mu_k) - \psi_k(\hat{\mu}_k)}{m_k(\mu_k) - m_k(\hat{\mu}_k)}$$

if $\quad \rho_k \geq \eta_1 \quad$ **then** $\quad \mu_{k+1} = \hat{\mu}_k \quad$ **else** $\quad \mu_{k+1} = \mu_k \quad$ **end if**

4: **Trust region update**:

if $\quad \rho_k \leq \eta_1 \quad$ **then** $\quad \Delta_{k+1} \in (0, \gamma\, \|\hat{\mu}_k - \mu_k\|)] \quad$ **end if**

if $\quad \rho_k \in (\eta_1, \eta_2) \quad$ **then** $\quad \Delta_{k+1} \in [\gamma\, \|\hat{\mu}_k - \mu_k\|\,, \Delta_k] \quad$ **end if**

if $\quad \rho_k \geq \eta_2 \quad$ **then** $\quad \Delta_{k+1} \in [\Delta_k, \Delta_{max}] \quad$ **end if**

Trust region ingredients for global convergence

Approximation models

$$m_k(\mu),\ \psi_k(\mu)$$

Error indicators

$$\|\nabla F(\mu) - \nabla m_k(\mu)\| \leq \xi \varphi_k(\mu) \qquad \xi > 0$$

$$|F(\mu_k) - F(\mu) + \psi_k(\mu) - \psi_k(\mu_k)| \leq \sigma \theta_k(\mu) \qquad \sigma > 0$$

Adaptivity

$$\varphi_k(\mu_k) \leq \kappa_\varphi \min\{\|\nabla m_k(\mu_k)\|, \Delta_k\}$$

$$\theta_k(\hat{\mu}_k)^\omega \leq \eta \min\{m_k(\mu_k) - m_k(\hat{\mu}_k), r_k\}$$

Global convergence

$$\boxed{\liminf_{k\to\infty} \|\nabla F(\mu_k)\| = 0}$$

Trust region method: ROM/SG approximation model

Approximation models built on two sources of inexactness

$$
\begin{aligned}
m_k(\mu) = &\quad \mathbb{E}_{\mathcal{I}_k}\left[\mathcal{J}(\boldsymbol{\Phi}_k \mathbf{u}_r(\mu, \cdot), \mu, \cdot)\right] \\
\psi_k(\mu) = &\quad \mathbb{E}_{\mathcal{I}'_k}\left[\mathcal{J}(\boldsymbol{\Phi}'_k \mathbf{u}_r(\mu, \cdot), \mu, \cdot)\right]
\end{aligned}
$$

Error indicators that account for both sources of error

$$
\begin{aligned}
\varphi_k(\mu) &= \alpha_1 \mathcal{E}_1(\mu; \mathcal{I}_k, \boldsymbol{\Phi}_k) + \alpha_2 \mathcal{E}_2(\mu; \mathcal{I}_k, \boldsymbol{\Phi}_k) + \alpha_3 \mathcal{E}_4(\mu; \mathcal{I}_k, \boldsymbol{\Phi}_k) \\
\theta_k(\mu) &= \beta_1(\mathcal{E}_1(\mu; \mathcal{I}'_k, \boldsymbol{\Phi}'_k) + \mathcal{E}_1(\mu_k; \mathcal{I}'_k, \boldsymbol{\Phi}'_k)) + \beta_2(\mathcal{E}_3(\mu; \mathcal{I}'_k, \boldsymbol{\Phi}'_k) + \mathcal{E}_3(\mu_k; \mathcal{I}'_k, \boldsymbol{\Phi}'_k))
\end{aligned}
$$

Reduced-order model errors

$$
\begin{aligned}
\mathcal{E}_1(\mu; \mathcal{I}, \boldsymbol{\Phi}) &= \mathbb{E}_{\mathcal{I} \cup \mathcal{N}(\mathcal{I})}\left[\|\mathbf{r}(\boldsymbol{\Phi}\mathbf{u}_r(\mu, \cdot), \mu, \cdot)\|\right] \\
\mathcal{E}_2(\mu; \mathcal{I}, \boldsymbol{\Phi}) &= \mathbb{E}_{\mathcal{I} \cup \mathcal{N}(\mathcal{I})}\left[\left\|\mathbf{r}^\lambda(\boldsymbol{\Phi}\mathbf{u}_r(\mu, \cdot), \boldsymbol{\Phi}\boldsymbol{\lambda}_r(\mu, \cdot), \mu, \cdot)\right\|\right]
\end{aligned}
$$

Sparse grid truncation errors

$$
\begin{aligned}
\mathcal{E}_3(\mu; \mathcal{I}, \boldsymbol{\Phi}) &= \mathbb{E}_{\mathcal{N}(\mathcal{I})}\left[|\mathcal{J}(\boldsymbol{\Phi}\mathbf{u}_r(\mu, \cdot), \mu, \cdot)|\right] \\
\mathcal{E}_4(\mu; \mathcal{I}, \boldsymbol{\Phi}) &= \mathbb{E}_{\mathcal{N}(\mathcal{I})}\left[\|\nabla\mathcal{J}(\boldsymbol{\Phi}\mathbf{u}_r(\mu, \cdot), \mu, \cdot)\|\right]
\end{aligned}
$$

Final requirement for convergence: Adaptivity

With the approximation model, $m_k(\mu)$, and gradient error indicator, $\varphi_k(\mu)$

$$
\begin{aligned}
m_k(\mu) &= \mathbb{E}_{\mathcal{I}_k}\left[\mathcal{J}(\boldsymbol{\Phi}_k u_r(\mu,\cdot),\, \mu,\, \cdot)\right] \\
\varphi_k(\mu) &= \alpha_1 \mathcal{E}_1(\mu; \mathcal{I}_k, \boldsymbol{\Phi}_k) + \alpha_2 \mathcal{E}_2(\mu; \mathcal{I}_k, \boldsymbol{\Phi}_k) + \alpha_3 \mathcal{E}_4(\mu; \mathcal{I}_k, \boldsymbol{\Phi}_k)
\end{aligned}
$$

the sparse grid $\mathcal{I}_k$ and reduced-order basis $\boldsymbol{\Phi}_k$ must be constructed such that the gradient condition holds

$$
\varphi_k(\mu_k) \leq \kappa_\varphi \min\{||\nabla m_k(\mu_k)||\,,\, \Delta_k\}
$$

Define dimension-adaptive greedy method to target each source of error such that the stronger conditions hold

$$
\begin{aligned}
\mathcal{E}_1(\mu_k; \mathcal{I}, \boldsymbol{\Phi}) &\leq \frac{\kappa_\varphi}{3\alpha_1} \min\{||\nabla m_k(\mu_k)||\,,\, \Delta_k\} \\
\mathcal{E}_2(\mu_k; \mathcal{I}, \boldsymbol{\Phi}) &\leq \frac{\kappa_\varphi}{3\alpha_2} \min\{||\nabla m_k(\mu_k)||\,,\, \Delta_k\} \\
\mathcal{E}_4(\mu_k; \mathcal{I}, \boldsymbol{\Phi}) &\leq \frac{\kappa_\varphi}{3\alpha_3} \min\{||\nabla m_k(\mu_k)||\,,\, \Delta_k\}
\end{aligned}
$$

Adaptivity: Dimension-adaptive greedy method

$$\textbf{while} \quad \mathcal{E}_4(\boldsymbol{\Phi}, \mathcal{I}, \boldsymbol{\mu}_k) > \frac{\kappa_\varphi}{3\alpha_3} \min\{\|\nabla m_k(\boldsymbol{\mu}_k)\|, \Delta_k\} \ \textbf{do}$$

Refine index set: Dimension-adaptive sparse grids

$$\mathcal{I}_k \leftarrow \mathcal{I}_k \cup \{\mathbf{j}^*\} \qquad \text{where} \qquad \mathbf{j}^* = \arg\max_{\mathbf{j} \in \mathcal{N}(\mathcal{I}_k)} \mathbb{E}_{\mathbf{j}}\left[\|\nabla \mathcal{J}(\boldsymbol{\Phi}\mathbf{u}_r(\boldsymbol{\mu}, \cdot), \boldsymbol{\mu}, \cdot)\|\right]$$

Adaptivity: Dimension-adaptive greedy method

while $\mathcal{E}_4(\boldsymbol{\Phi}, \mathcal{I}, \boldsymbol{\mu}_k) > \frac{\kappa_\varphi}{3\alpha_3} \min\{\|\nabla m_k(\boldsymbol{\mu}_k)\|, \Delta_k\}$ **do**

Refine index set: Dimension-adaptive sparse grids

$$\mathcal{I}_k \leftarrow \mathcal{I}_k \cup \{\mathbf{j}^*\} \qquad \text{where} \qquad \mathbf{j}^* = \operatorname*{arg\,max}_{\mathbf{j} \in \mathcal{N}(\mathcal{I}_k)} \mathbb{E}_{\mathbf{j}}\left[\|\nabla \mathcal{J}(\boldsymbol{\Phi}\mathbf{u}_r(\boldsymbol{\mu}, \cdot), \boldsymbol{\mu}, \cdot)\|\right]$$

Refine reduced-order basis: Greedy sampling

while $\mathcal{E}_1(\boldsymbol{\Phi}, \mathcal{I}, \boldsymbol{\mu}_k) > \frac{\kappa_\varphi}{3\alpha_1} \min\{\|\nabla m_k(\boldsymbol{\mu}_k)\|, \Delta_k\}$ **do**

$$\boldsymbol{\Phi}_k \leftarrow \begin{bmatrix} \boldsymbol{\Phi}_k & \mathbf{u}(\boldsymbol{\mu}_k, \boldsymbol{\xi}^*) & \boldsymbol{\lambda}(\boldsymbol{\mu}_k, \boldsymbol{\xi}^*) \end{bmatrix}$$

$$\boldsymbol{\xi}^* = \operatorname*{arg\,max}_{\boldsymbol{\xi} \in \Xi_{\mathbf{j}^*}} \rho(\boldsymbol{\xi}) \|\mathbf{r}(\boldsymbol{\Phi}_k \mathbf{u}_r(\boldsymbol{\mu}_k, \boldsymbol{\xi}), \boldsymbol{\mu}_k, \boldsymbol{\xi})\|$$

end while

Adaptivity: Dimension-adaptive greedy method

while $\mathcal{E}_4(\Phi, \mathcal{I}, \mu_k) > \frac{\kappa_\varphi}{3\alpha_3} \min\{\|\nabla m_k(\mu_k)\|, \Delta_k\}$ **do**

Refine index set: Dimension-adaptive sparse grids

$$\mathcal{I}_k \leftarrow \mathcal{I}_k \cup \{\mathbf{j}^*\} \qquad \text{where} \qquad \mathbf{j}^* = \underset{\mathbf{j} \in \mathcal{N}(\mathcal{I}_k)}{\arg\max}\ \mathbb{E}_{\mathbf{j}}\left[\|\nabla \mathcal{J}(\Phi \mathbf{u}_r(\mu, \cdot), \mu, \cdot)\|\right]$$

Refine reduced-order basis: Greedy sampling

while $\mathcal{E}_1(\Phi, \mathcal{I}, \mu_k) > \frac{\kappa_\varphi}{3\alpha_1} \min\{\|\nabla m_k(\mu_k)\|, \Delta_k\}$ **do**

$$\Phi_k \leftarrow \begin{bmatrix} \Phi_k & \mathbf{u}(\mu_k, \xi^*) & \lambda(\mu_k, \xi^*) \end{bmatrix}$$

$$\xi^* = \underset{\xi \in \Xi_{\mathbf{j}^*}}{\arg\max}\ \rho(\xi) \|\mathbf{r}(\Phi_k \mathbf{u}_r(\mu_k, \xi), \mu_k, \xi)\|$$

end while

while $\mathcal{E}_2(\Phi, \mathcal{I}, \mu_k) > \frac{\kappa_\varphi}{3\alpha_2} \min\{\|\nabla m_k(\mu_k)\|, \Delta_k\}$ **do**

$$\Phi_k \leftarrow \begin{bmatrix} \Phi_k & \mathbf{u}(\mu_k, \xi^*) & \lambda(\mu_k, \xi^*) \end{bmatrix}$$

$$\xi^* = \underset{\xi \in \Xi_{\mathbf{j}^*}}{\arg\max}\ \rho(\xi) \left\|\mathbf{r}^\lambda(\Phi_k \mathbf{u}_r(\mu_k, \xi), \Phi_k \lambda_r(\mu_k, \xi), \mu_k, \xi)\right\|$$

end while

Optimal control of steady Burgers' equation

- **Optimization problem:**

$$\underset{\mu \in \mathbb{R}^{n_\mu}}{\text{minimize}} \quad \int_\Xi \rho(\xi) \left[\int_0^1 \frac{1}{2} (u(\mu, \xi, x) - \bar{u}(x))^2 \, dx + \frac{\alpha}{2} \int_0^1 z(\mu, x)^2 \, dx \right] d\xi$$

where $u(\mu, \xi, x)$ solves

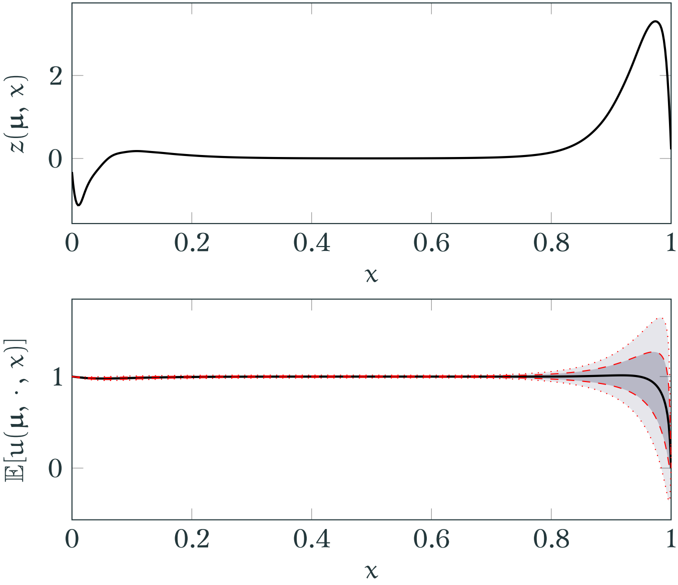
$$-\nu(\xi) \partial_{xx} u(\mu, \xi, x) + u(\mu, \xi, x) \partial_x u(\mu, \xi, x) = z(\mu, x) \quad x \in (0, 1), \quad \xi \in \Xi$$
$$u(\mu, \xi, 0) = d_0(\xi) \qquad u(\mu, \xi, 1) = d_1(\xi)$$

- **Target state**: $\bar{u}(x) \equiv 1$
- **Stochastic Space**: $\Xi = [-1, 1]^3$, $\rho(\xi) d\xi = 2^{-3} d\xi$

$$\nu(\xi) = 10^{\xi_1 - 2} \qquad d_0(\xi) = 1 + \frac{\xi_2}{1000} \qquad d_1(\xi) = \frac{\xi_3}{1000}$$

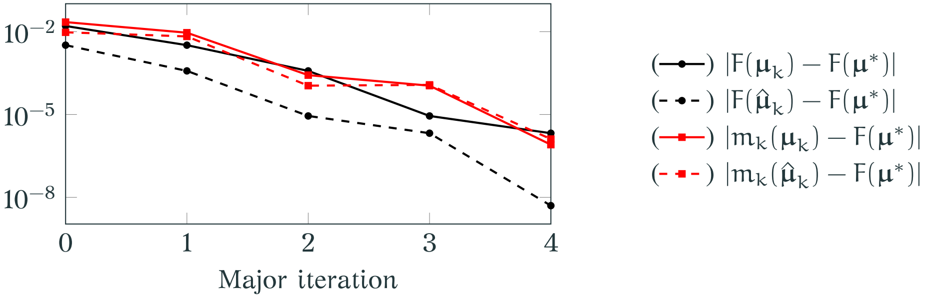
- **Parametrization**: $z(\mu, x)$ – cubic splines with 51 knots, $n_\mu = 53$

Optimal control and statistics



Optimal control and corresponding mean state (——) $\pm$ one (- - -) and two (⋯⋯) standard deviations

Global convergence without pointwise agreement

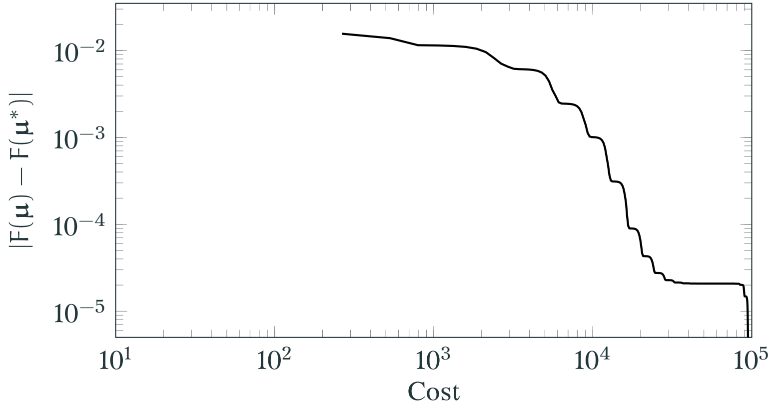


$F(\mu_k)$	$m_k(\mu_k)$	$F(\hat{\mu}_k)$	$m_k(\hat{\mu}_k)$	$\lVert\nabla F(\mu_k)\rVert$	ρ_k	Success?
6.6506e-02	7.2694e-02	5.3655e-02	5.9922e-02	2.2959e-02	1.0257e+00	1.0000e+00
5.3655e-02	5.9593e-02	5.0783e-02	5.7152e-02	2.3424e-03	9.7512e-01	1.0000e+00
5.0783e-02	5.0670e-02	5.0412e-02	5.0292e-02	1.9724e-03	9.8351e-01	1.0000e+00
5.0412e-02	5.0292e-02	5.0405e-02	5.0284e-02	9.2654e-05	8.7479e-01	1.0000e+00
5.0405e-02	5.0404e-02	5.0403e-02	5.0401e-02	8.3139e-05	9.9946e-01	1.0000e+00
5.0403e-02	5.0401e-02	-	-	2.2846e-06	-	-

Convergence history of trust region method built on two-level approximation

Significant reduction in cost, even if (largest) ROM only 10× faster than HDM

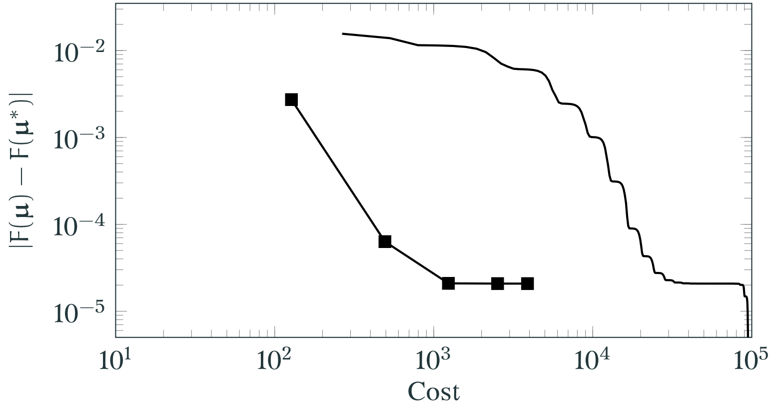
$$\text{Cost} = \text{nHdmPrim} + 0.5 \times \text{nHdmAdj} + \tau^{-1} \times (\text{nRomPrim} + 0.5 \times \text{nRomAdj})$$



5-level isotropic SG (——), dimension-adaptive SG [Kouri et al., 2014] ( ), and proposed ROM/SG for $\tau = 1$ ( ), $\tau = 10$ ( ), $\tau = 100$ ( ), $\tau = \infty$ ( )

Significant reduction in cost, even if (largest) ROM only $10\times$ faster than HDM

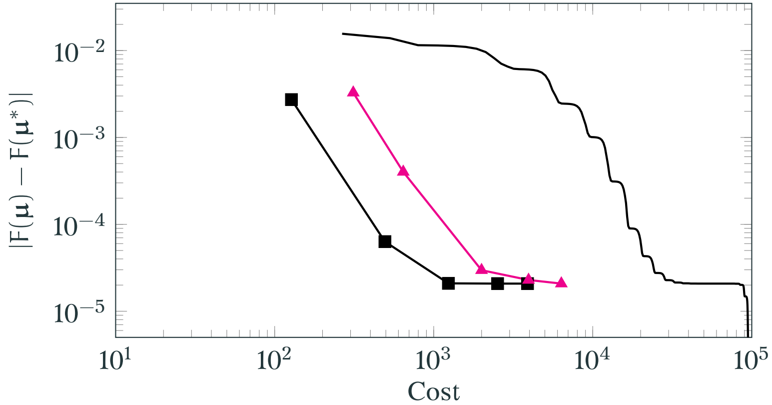
$$\text{Cost} = \text{nHdmPrim} + 0.5 \times \text{nHdmAdj} + \tau^{-1} \times (\text{nRomPrim} + 0.5 \times \text{nRomAdj})$$



5-level isotropic SG (——), dimension-adaptive SG [Kouri et al., 2014] (—■—), and proposed ROM/SG for $\tau = 1$ ( ), $\tau = 10$ ( ), $\tau = 100$ ( ), $\tau = \infty$ ( )

Significant reduction in cost, even if (largest) ROM only 10× faster than HDM

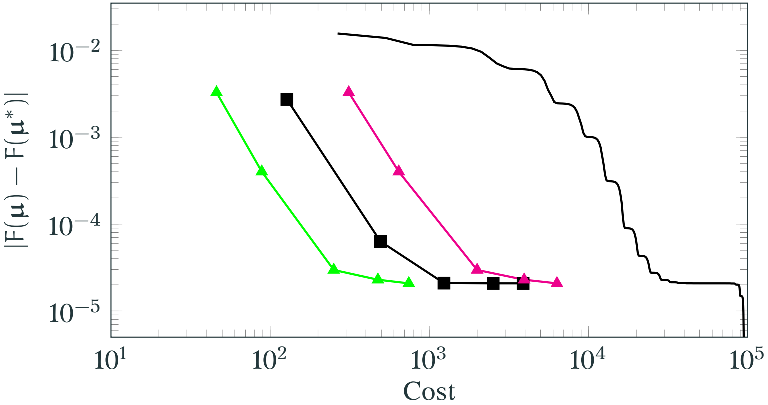
$$\text{Cost} = \text{nHdmPrim} + 0.5 \times \text{nHdmAdj} + \tau^{-1} \times (\text{nRomPrim} + 0.5 \times \text{nRomAdj})$$



5-level isotropic SG (——), dimension-adaptive SG [Kouri et al., 2014] (–■–), and proposed ROM/SG for $\tau = 1$ (–▲–), $\tau = 10$ ( ), $\tau = 100$ ( ), $\tau = \infty$ ( )

Significant reduction in cost, even if (largest) ROM only 10× faster than HDM

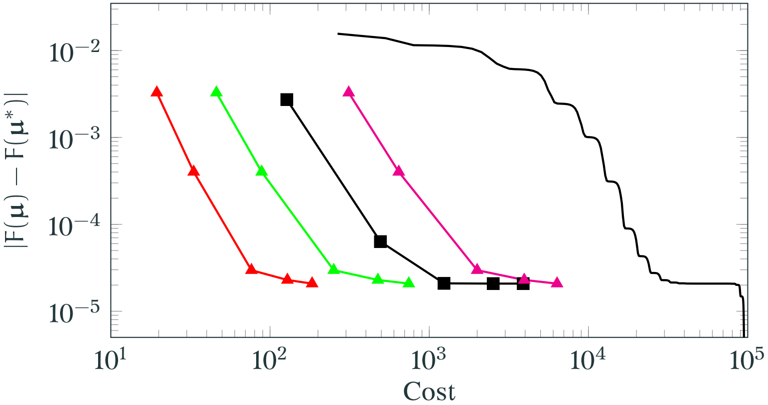
$$\text{Cost} = \text{nHdmPrim} + 0.5 \times \text{nHdmAdj} + \tau^{-1} \times (\text{nRomPrim} + 0.5 \times \text{nRomAdj})$$



5-level isotropic SG (——), dimension-adaptive SG [Kouri et al., 2014] (—■—), and proposed ROM/SG for $\tau = 1$ (—▲—), $\tau = 10$ (—▲—), $\tau = 100$ ( ), $\tau = \infty$ ( )

Significant reduction in cost, even if (largest) ROM only 10× faster than HDM

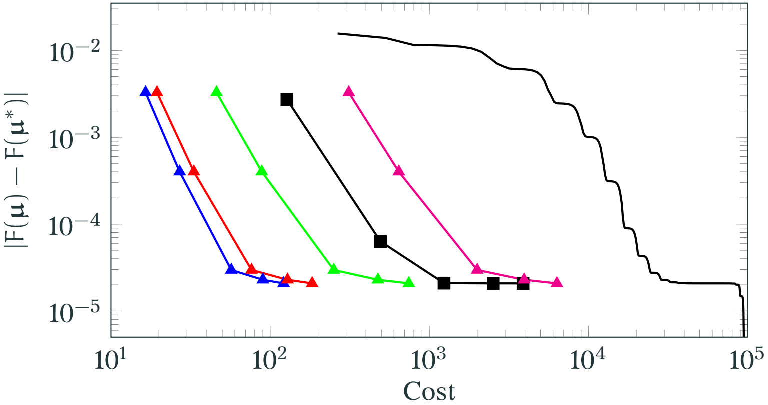
$$\text{Cost} = \text{nHdmPrim} + 0.5 \times \text{nHdmAdj} + \tau^{-1} \times (\text{nRomPrim} + 0.5 \times \text{nRomAdj})$$



5-level isotropic SG (——), dimension-adaptive SG [Kouri et al., 2014] (—■—), and proposed ROM/SG for $\tau = 1$ (—▲—), $\tau = 10$ (—▲—), $\tau = 100$ (—▲—), $\tau = \infty$ ( )

Significant reduction in cost, even if (largest) ROM only 10× faster than HDM

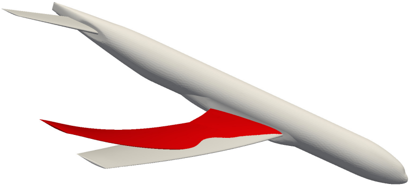
$$\text{Cost} = \text{nHdmPrim} + 0.5 \times \text{nHdmAdj} + \tau^{-1} \times (\text{nRomPrim} + 0.5 \times \text{nRomAdj})$$

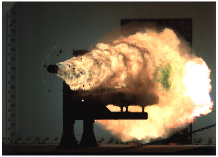


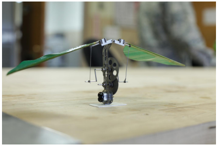
5-level isotropic SG (——), dimension-adaptive SG [Kouri et al., 2014] (—■—), and proposed ROM/SG for $\tau = 1$ (—▲—), $\tau = 10$ (—▲—), $\tau = 100$ (—▲—), $\tau = \infty$ (—▲—)

Leveraging inexactness to accelerate PDE-constrained optimization

- Framework introduced for accelerating **stochastic** PDE-constrained optimization problems
  - Adaptive *model reduction*
  - Dimension-adaptive *sparse grids*
- Inexactness **managed** with flexible **trust region** method
- **100**$\times$ speedup on (stochastic) optimal control of 1D flow







References I

Chen, P. and Quarteroni, A. (2014).
Weighted reduced basis method for stochastic optimal control problems with elliptic PDE constraint.
Siam/Asa Journal on Uncertainty Quantification, 2(1):364–396.

Chen, P. and Quarteroni, A. (2015).
A new algorithm for high–dimensional uncertainty quantification based on dimension–adaptive sparse grid approximation and reduced basis methods.
Journal of Computational Physics, 298:176–193.

Kouri, D. P., Heinkenschloss, M., Ridzal, D., and van Bloemen Waanders, B. G. (2013).
A trust-region algorithm with adaptive stochastic collocation for pde optimization under uncertainty.
SIAM Journal on Scientific Computing, 35(4):A1847–A1879.

References II

Kouri, D. P., Heinkenschloss, M., Ridzal, D., and van Bloemen Waanders, B. G. (2014).
Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty.
SIAM Journal on Scientific Computing, 36(6):A3011–A3029.

Tiesler, H., Kirby, R. M., Xiu, D., and Preusser, T. (2012).
Stochastic collocation for optimal control problems with stochastic PDE constraints.
SIAM Journal on Control and Optimization, 50(5):2659–2682.

Zahr, M. J. and Persson, P.-O. (2017).
Energetically optimal flapping wing motions via adjoint-based optimization and high-order discretizations.
In *Frontiers in PDE-Constrained Optimization*. Springer.