PDE-Constrained Optimization using Progressively-Constructed Reduced-Order Models

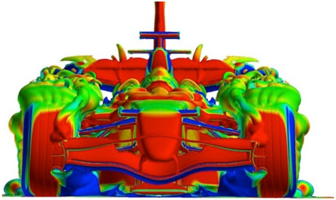
Matthew J. Zahr and Charbel Farhat

Institute for Computational and Mathematical Engineering
Farhat Research Group
Stanford University

World Congress on Computational Mechanics XI
July 20 - 25, 2014
Barcelona, Spain
Advanced Reduced-Order Modeling Strategies for Parametrized PDEs and Applications II

Motivation

- PDE-constrained is ubiquitous in engineering
  - Design optimization
  - Optimal control
  - Parameter estimation (inverse problems)
- Notoriously expensive as many calls to a PDE solver may be required
  - CFD, structural dynamics, acoustic models
- Good candidate for model reduction
  - Many-query application

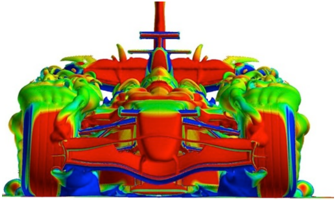


Problem Formulation

Goal: Rapidly solve PDE-constrained optimization problems of the form

$$\begin{aligned} \underset{\mathbf{w}\in\mathbb{R}^N,\ \boldsymbol{\mu}\in\mathbb{R}^p}{\text{minimize}} \quad & f(\mathbf{w}, \boldsymbol{\mu}) \\ \text{subject to} \quad & \mathbf{R}(\mathbf{w}, \boldsymbol{\mu}) = 0 \end{aligned}$$

where $\mathbf{R} : \mathbb{R}^N \times \mathbb{R}^p \to \mathbb{R}^N$ is the discretized (nonlinear) PDE, $\mathbf{w}$ is the PDE state vector, $\boldsymbol{\mu}$ is the vector of parameters, and N is assumed to be very large.



Reduced-Order Model

- Model Order Reduction (MOR) assumption: *state vector lies in low-dimensional affine subspace*

$$\mathbf{w} \approx \mathbf{w}_r = \bar{\mathbf{w}} + \boldsymbol{\Phi}\mathbf{y}$$

  where $\mathbf{y} \in \mathbb{R}^n$ are the reduced coordinates of $\mathbf{w}$ in the basis $\boldsymbol{\Phi} \in \mathbb{R}^{N\times n}$, $\bar{\mathbf{w}}$ piecewise constant in $\boldsymbol{\mu}$, and $n \ll N$
- Substitute assumption into High-Dimensional Model (HDM), $\mathbf{R}(\mathbf{w}, \boldsymbol{\mu}) = 0$

$$\mathbf{R}(\bar{\mathbf{w}} + \boldsymbol{\Phi}\mathbf{y}, \boldsymbol{\mu}) \approx 0$$

- Require projection of residual in low-dimensional *left subspace*, with basis $\boldsymbol{\Psi} \in \mathbb{R}^{N\times n}$ to be zero

$$\boxed{\mathbf{R}_r(\mathbf{y}, \boldsymbol{\mu}) = \boldsymbol{\Psi}^T\mathbf{R}(\bar{\mathbf{w}} + \boldsymbol{\Phi}\mathbf{y}, \boldsymbol{\mu}) = 0}$$

Reduced Optimization Problem

- *ROM-constrained optimization* problem

$$\begin{aligned} \underset{\mathbf{y}\in\mathbb{R}^n,\ \boldsymbol{\mu}\in\mathbb{R}^p}{\text{minimize}} \quad & f(\bar{\mathbf{w}} + \mathbf{\Phi}\mathbf{y}, \boldsymbol{\mu}) \\ \text{subject to} \quad & \mathbf{\Psi}^T\mathbf{R}(\bar{\mathbf{w}} + \mathbf{\Phi}\mathbf{y}, \boldsymbol{\mu}) = 0 \end{aligned}$$

- Issues that must be considered
  - Basis construction
  - Reduced sensitivity derivation
  - Training

State-Sensitivity POD

- MOR assumption: $\mathbf{w} \approx \mathbf{w}_r = \bar{\mathbf{w}} + \mathbf{\Phi}\mathbf{y} \implies \frac{\partial \mathbf{w}_r}{\partial \boldsymbol{\mu}} = \mathbf{\Phi}\frac{\partial \mathbf{y}}{\partial \boldsymbol{\mu}}$
- Collect state and sensitivity snapshots by sampling HDM

$$\mathbf{X} = \begin{bmatrix} \mathbf{w}(\boldsymbol{\mu}_1) - \bar{\mathbf{w}} & \mathbf{w}(\boldsymbol{\mu}_2) - \bar{\mathbf{w}} & \cdots & \mathbf{w}(\boldsymbol{\mu}_n) - \bar{\mathbf{w}} \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} \frac{\partial \mathbf{w}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}_1) & \frac{\partial \mathbf{w}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}_2) & \cdots & \frac{\partial \mathbf{w}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}_n) \end{bmatrix}$$

- Use Proper Orthogonal Decomposition to generate reduced bases from each *individually*

$$\mathbf{\Phi}_{\mathbf{X}} = \mathrm{POD}(\mathbf{X})$$

$$\mathbf{\Phi}_{\mathbf{Y}} = \mathrm{POD}(\mathbf{Y})$$

- Concatenate to get ROB

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{\Phi}_{\mathbf{X}} & \mathbf{\Phi}_{\mathbf{Y}} \end{bmatrix}$$

Sensitivities

For gradient-based optimization, sensitivities are required

- HDM sensitivities

$$\mathbf{R}(\mathbf{w}(\boldsymbol{\mu}), \boldsymbol{\mu}) = 0 \implies \frac{\partial \mathbf{w}}{\partial \boldsymbol{\mu}} = -\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]^{-1} \frac{\partial \mathbf{R}}{\partial \boldsymbol{\mu}}$$

- ROM sensitivities

$$\mathbf{R}_r(\mathbf{y}(\boldsymbol{\mu}), \boldsymbol{\mu}) = 0 \implies \frac{\partial \mathbf{w}_r}{\partial \boldsymbol{\mu}} = \boldsymbol{\Phi} \frac{\partial \mathbf{y}}{\partial \boldsymbol{\mu}} = \boldsymbol{\Phi} \mathbf{A}^{-1} \mathbf{B}$$

$$\mathbf{A} = \sum_{j=1}^{N} \mathbf{R}_j \frac{\partial \left(\boldsymbol{\Psi}^T \mathbf{e}_j\right)}{\partial \mathbf{w}} \boldsymbol{\Phi} + \boldsymbol{\Psi}^T \frac{\partial \mathbf{R}}{\partial \mathbf{w}} \boldsymbol{\Phi}$$

$$\mathbf{B} = -\left(\sum_{j=1}^{N} \mathbf{R}_j \frac{\partial \left(\boldsymbol{\Psi}^T \mathbf{e}_j\right)}{\partial \boldsymbol{\mu}} + \boldsymbol{\Psi}^T \frac{\partial \mathbf{R}}{\partial \boldsymbol{\mu}}\right)$$

Minimum-Error Reduced Sensitivities

- True sensitivities of the ROM
  - May be difficult to compute if $\boldsymbol{\Psi} = \boldsymbol{\Psi}(\boldsymbol{\mu})$
    - LSPG [Bui-Thanh *et al* 2008, Carlberg *et al* 2011]
  - May not represent HDM sensitivities well
    - Gradients of reduced optimization functions may not be close to the true gradients
- Define quantity that *minimizes* the sensitivity error in some norm $\boldsymbol{\Theta} \succ 0$

$$\widehat{\frac{\partial \mathbf{y}}{\partial \boldsymbol{\mu}}} = \arg\min_{\mathbf{a}} \left|\left| \frac{\partial \mathbf{w}}{\partial \boldsymbol{\mu}} - \boldsymbol{\Phi}\mathbf{a} \right|\right|_{\boldsymbol{\Theta}}$$

$$\Longrightarrow \quad \widehat{\frac{\partial \mathbf{y}}{\partial \boldsymbol{\mu}}} = -\left(\boldsymbol{\Theta}^{1/2}\boldsymbol{\Phi}\right)^{\dagger}\boldsymbol{\Theta}^{1/2}\frac{\partial \mathbf{R}}{\partial \mathbf{w}}^{-1}\frac{\partial \mathbf{R}}{\partial \boldsymbol{\mu}}$$

Minimum-Error Reduced Sensitivities

- Select $\boldsymbol{\Theta}^{1/2} = \frac{\partial \mathbf{R}}{\partial \mathbf{w}}$

$$\widehat{\frac{\partial \mathbf{y}}{\partial \boldsymbol{\mu}}} = -\left(\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\boldsymbol{\Phi}\right)^{\dagger}\frac{\partial \mathbf{R}}{\partial \boldsymbol{\mu}}$$

- *Exactly* reproduce sensitivities at training points if sensitivity basis not truncated
- May cause convergence issues for *reduced* optimization problem

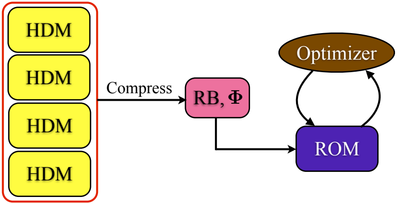
Training: Offline-Online (Database) Approach

- Identify samples in *offline* phase to be used for training
- Collect snapshots by running HDM (state vector and sensitivities)
- Build ROB $\mathbf{\Phi}$
- Solve optimization problem

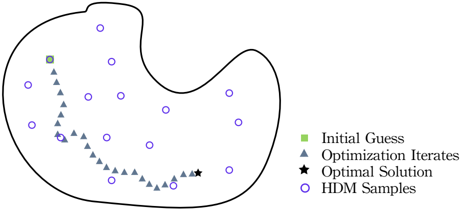
$$\begin{aligned} \underset{\mathbf{y}\in\mathbb{R}^n,\ \boldsymbol{\mu}\in\mathbb{R}^p}{\text{minimize}} \quad & f(\bar{\mathbf{w}} + \mathbf{\Phi}\mathbf{y}, \boldsymbol{\mu}) \\ \text{subject to} \quad & \mathbf{\Psi}^T\mathbf{R}(\bar{\mathbf{w}} + \mathbf{\Phi}\mathbf{y}, \boldsymbol{\mu}) = 0 \end{aligned}$$

[Lassila *et al* 2010, Rozza *et al* 2010, Manzoni *et al* 2012]

Offline-Online Approach



(a) Schematic of Algorithm



(b) Idealized Optimization Trajectory in Parameter Space

(c) Breakdown of Computational Effort

Training: Progressive Approach

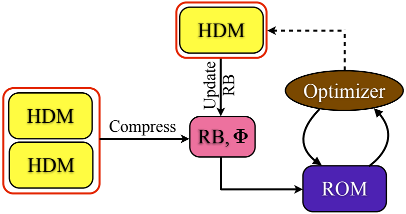
- Collect snapshots by running HDM (state vector and sensitivities) at initial guess for optimization problem
- Build ROB $\mathbf{\Phi}$ from sparse training
- Solve optimization problem

$$\begin{aligned} \underset{\mathbf{y}\in\mathbb{R}^n,\ \boldsymbol{\mu}\in\mathbb{R}^p}{\text{minimize}} \quad & f(\bar{\mathbf{w}} + \mathbf{\Phi}\mathbf{y}, \boldsymbol{\mu}) \\ \text{subject to} \quad & \mathbf{\Psi}^T\mathbf{R}(\bar{\mathbf{w}} + \mathbf{\Phi}\mathbf{y}, \boldsymbol{\mu}) = 0 \\ & \frac{1}{2}||\mathbf{R}(\bar{\mathbf{w}} + \mathbf{\Phi}\mathbf{y}, \boldsymbol{\mu})||_2^2 \leq \epsilon \end{aligned}$$

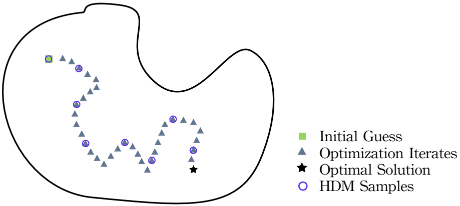
- Use solution of above problem to enrich training and repeat until convergence

Similar approaches found in: [Arian *et al* 2000, Afanasiev *et al* 2001, Fahl 2001]

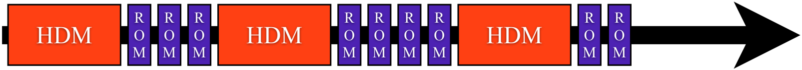
Progressive Approach



(a) Schematic of Algorithm

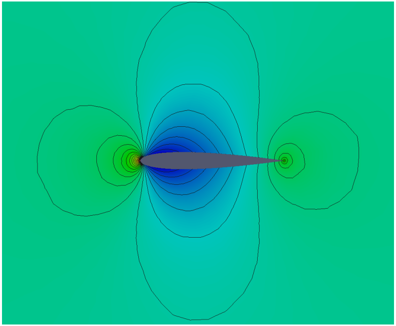


(b) Idealized Optimization Trajectory in Parameter Space

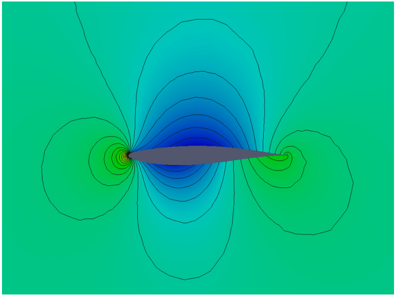


(c) Breakdown of Computational Effort

Compressible, Inviscid Airfoil Inverse Design



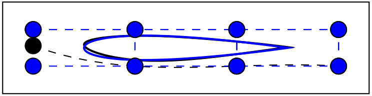
(a) NACA0012: Pressure field ($M_\infty = 0.5$, $\alpha = 0.0°$)



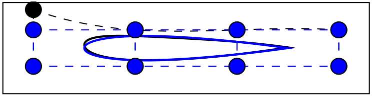
(b) RAE2822: Pressure field ($M_\infty = 0.5$, $\alpha = 0.0°$)

- Pressure discrepancy minimization (Euler equations)
  - Initial Configuration: NACA0012
  - Target Configuration: RAE2822

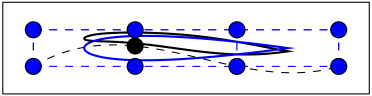
Shape Parametrization



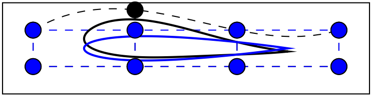
(a) $\boldsymbol{\mu}(1) = 0.1$



(b) $\boldsymbol{\mu}(2) = 0.1$



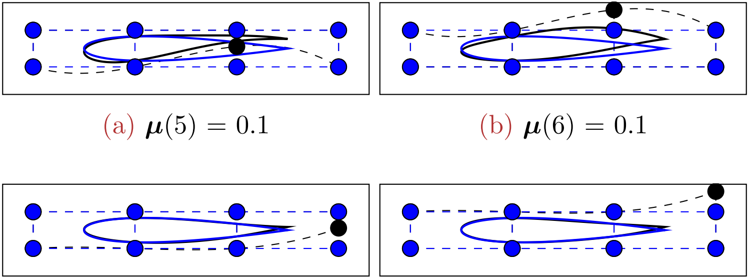
(c) $\boldsymbol{\mu}(3) = 0.1$



(d) $\boldsymbol{\mu}(4) = 0.1$

Figure : Shape parametrization of a NACA0012 airfoil using a *cubic* design element

Shape Parametrization

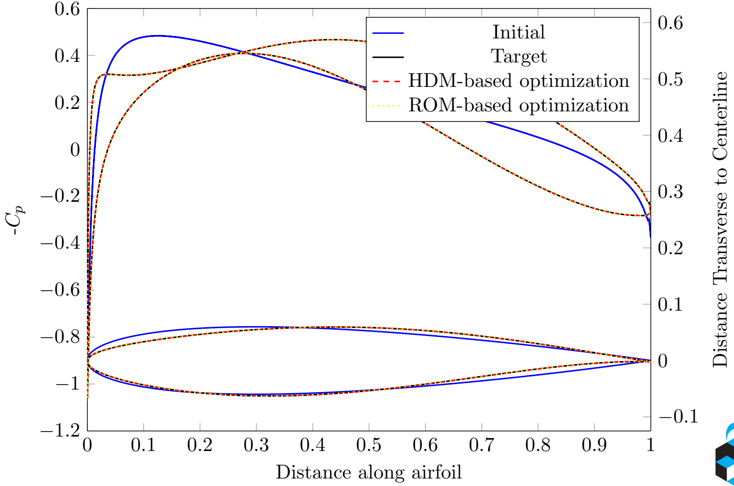


(a) $\boldsymbol{\mu}(5) = 0.1$ (b) $\boldsymbol{\mu}(6) = 0.1$

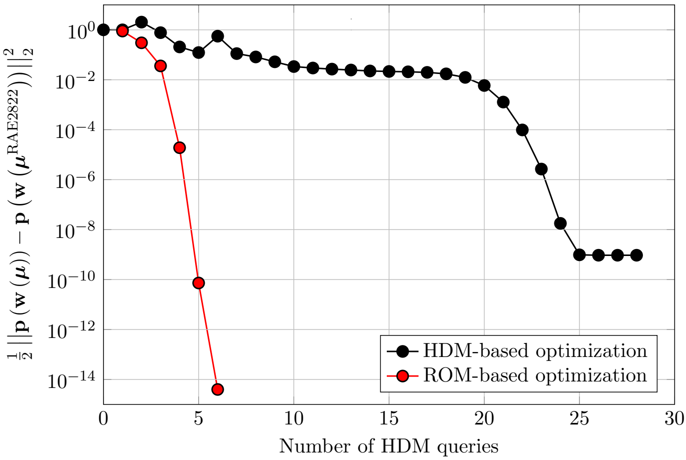
(c) $\boldsymbol{\mu}(7) = 0.1$ (d) $\boldsymbol{\mu}(8) = 0.1$

Figure : Shape parametrization of a NACA0012 airfoil using a *cubic* design element

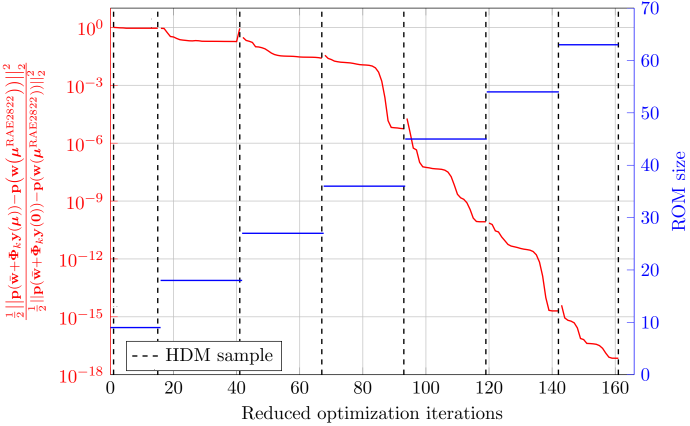
Optimization Results



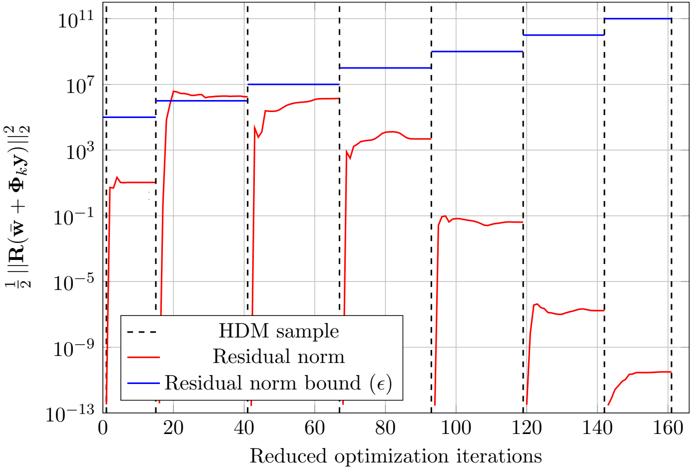
Optimization Results



Optimization Results



Optimization Results



Optimization Results

	HDM-based optimization	ROM-based optimization
# of HDM Evaluations	29	7
# of ROM Evaluations	-	346
$\frac{\|\|\boldsymbol{\mu}^* - \boldsymbol{\mu}^{RAE2822}\|\|}{\|\|\boldsymbol{\mu}^{RAE2822}\|\|}$	$2.28 \times 10^{-3}\%$	$4.17 \times 10^{-6}\%$

Table : Performance of the HDM- and ROM-based optimization methods

Conclusion

Overview

- Introduced progressive, nonlinear trust region framework for reduced optimization
- Proposed minimum-error reduced sensitivity analysis
- Demonstrated approach on canonical problem from aerodynamic shape optimization

Future work

- Extend to hyper-reduced models (i.e. reduce nonlinear term) to achieve significant speedup
- Application to large-scale, 3D problems