Accelerating PDE-Constrained Optimization Problems using Adaptive Reduced-Order Models

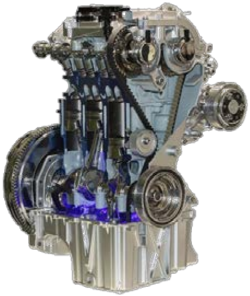
Matthew J. Zahr

Advisor: Charbel Farhat
Computational and Mathematical Engineering
Stanford University

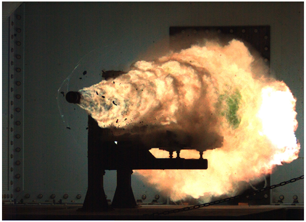
Department of Aerospace and Mechanical Engineering
University of Notre Dame
South Bend, IN
March 3, 2016

Multiphysics Optimization Key Player in Next-Gen Problems

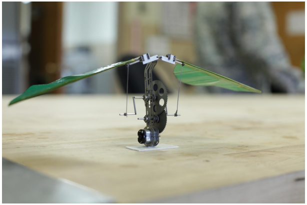
Current interest in computational physics reaches far beyond analysis of a single configuration of a physical system into ***design*** *(shape and topology[1]),* ***control****, and* ***uncertainty quantification***



Engine System



EM Launcher

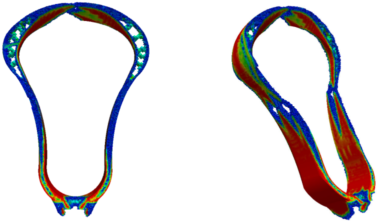


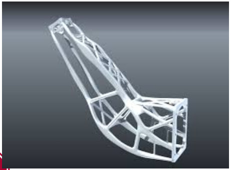
Micro-Aerial Vehicle

[1] Emergence of additive manufacturing technologies has made topology optimization increasingly relevant, particularly in DOE.

Topology Optimization and Additive Manufacturing[2]

- Emergence of AM has made TO an increasingly relevant topic
- AM+TO lead to highly efficient designs that could not be realized previously
- Challenges: smooth topologies require **very fine** meshes and modeling of complex **manufacturing process**





[2] *MIT Technology Review*, **Top 10 Technological Breakthrough 2013**

PDE-Constrained Optimization I

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

$$\begin{aligned} \underset{\boldsymbol{u}\in\mathbb{R}^{n_{\boldsymbol{u}}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{u},\ \boldsymbol{\mu}) \\ \text{subject to} \quad & \boldsymbol{r}(\boldsymbol{u},\ \boldsymbol{\mu}) = 0 \end{aligned}$$

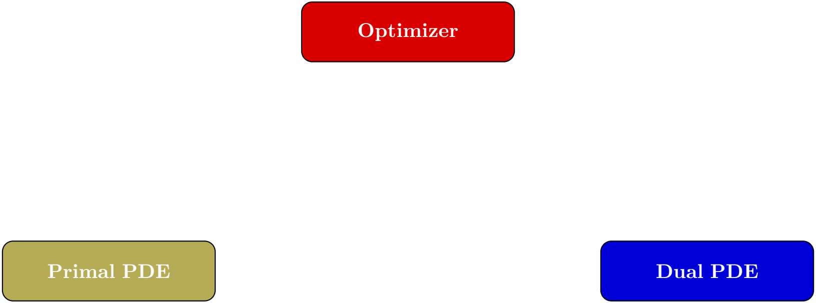
where

- $\boldsymbol{r} : \mathbb{R}^{n_{\boldsymbol{u}}} \times \mathbb{R}^{n_{\boldsymbol{\mu}}} \to \mathbb{R}^{n_{\boldsymbol{u}}}$ is the discretized partial differential equation
- $\mathcal{J} : \mathbb{R}^{n_{\boldsymbol{u}}} \times \mathbb{R}^{n_{\boldsymbol{\mu}}} \to \mathbb{R}$ is the objective function
- $\boldsymbol{u} \in \mathbb{R}^{n_{\boldsymbol{u}}}$ is the PDE state vector
- $\boldsymbol{\mu} \in \mathbb{R}^{n_{\boldsymbol{\mu}}}$ is the vector of parameters

red indicates a large-scale quantity, $\mathcal{O}(mesh)$

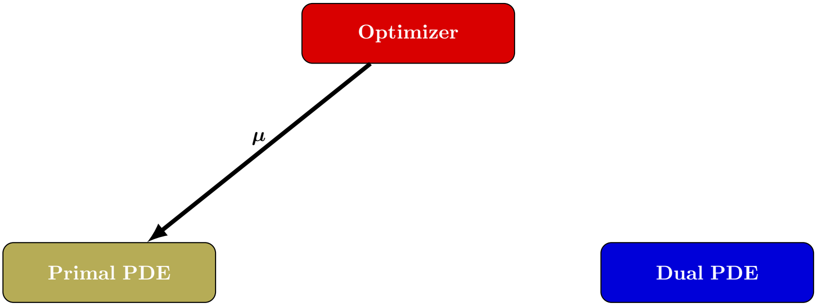
Nested Approach to PDE-Constrained Optimization

Virtually all expense emanates from primal/dual PDE solvers



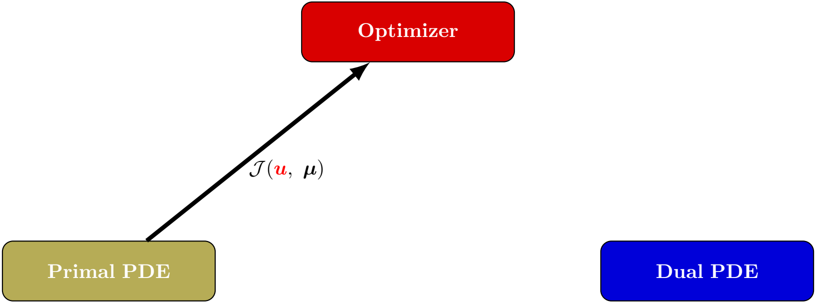
Nested Approach to PDE-Constrained Optimization

Virtually all expense emanates from primal/dual PDE solvers



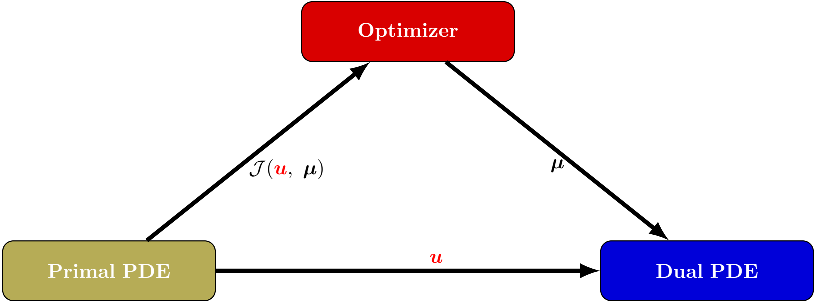
Nested Approach to PDE-Constrained Optimization

Virtually all expense emanates from primal/dual PDE solvers



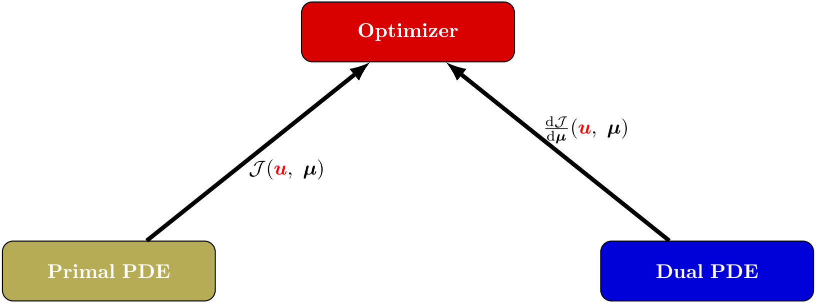
Nested Approach to PDE-Constrained Optimization

Virtually all expense emanates from primal/dual PDE solvers



Nested Approach to PDE-Constrained Optimization

Virtually all expense emanates from primal/dual PDE solvers



Projection-Based Model Reduction to Reduce PDE Size

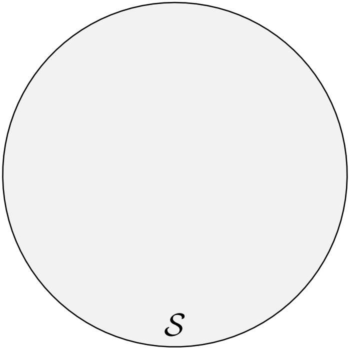
- Model Order Reduction (MOR) assumption: *state vector lies in low-dimensional subspace*

$$\boldsymbol{u} \approx \boldsymbol{\Phi_u} \boldsymbol{u}_r \qquad \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\mu}} \approx \boldsymbol{\Phi_u} \frac{\partial \boldsymbol{u}_r}{\partial \boldsymbol{\mu}}$$

  where
  - $\boldsymbol{\Phi_u} = \begin{bmatrix} \boldsymbol{\phi_u}^1 & \cdots & \boldsymbol{\phi_u}^{k_{\boldsymbol{u}}} \end{bmatrix} \in \mathbb{R}^{n_{\boldsymbol{u}} \times k_{\boldsymbol{u}}}$ is the reduced basis
  - $\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}}$ are the reduced coordinates of $\boldsymbol{u}$
  - $n_{\boldsymbol{u}} \gg k_{\boldsymbol{u}}$
- Substitute assumption into High-Dimensional Model (HDM), $\boldsymbol{r}(\boldsymbol{u}, \ \boldsymbol{\mu}) = 0$, and project onto test subspace $\boldsymbol{\Psi_u} \in \mathbb{R}^{n_{\boldsymbol{u}} \times k_{\boldsymbol{u}}}$

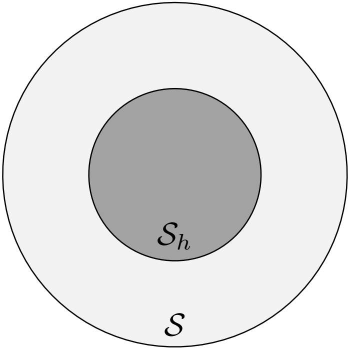
$$\boldsymbol{\Psi_u}^T \boldsymbol{r}(\boldsymbol{\Phi_u} \boldsymbol{u}_r, \ \boldsymbol{\mu}) = 0$$

Connection to Finite Element Method: Hierarchical Subspaces



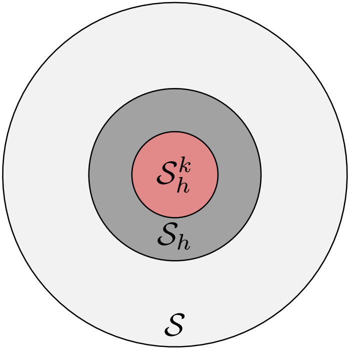
- $\mathcal{S}$ - infinite-dimensional trial space

Connection to Finite Element Method: Hierarchical Subspaces



- $\mathcal{S}$ - infinite-dimensional trial space
- $\mathcal{S}_h$ - (large) finite-dimensional trial space

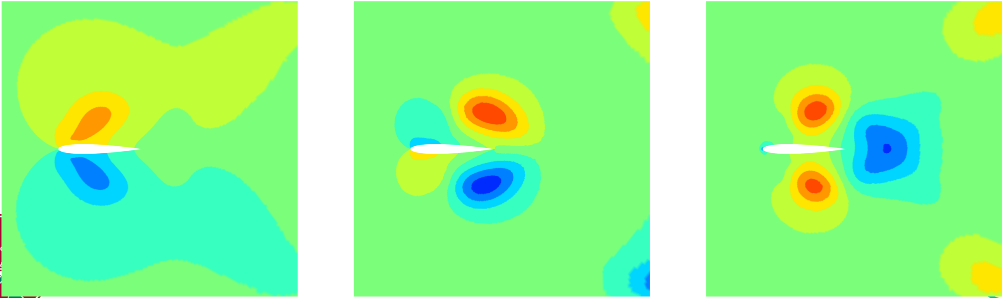
Connection to Finite Element Method: Hierarchical Subspaces



- $\mathcal{S}$ - infinite-dimensional trial space
- $\mathcal{S}_h$ - (large) finite-dimensional trial space
- $\mathcal{S}_h^k$ - (small) finite-dimensional trial space
- $\mathcal{S}_h^k \subset \mathcal{S}_h \subset \mathcal{S}$

Few Global, Data-Driven Basis Functions v. Many Local Ones

- Instead of using traditional *local* shape functions (e.g., FEM), use *global* shape functions
- Instead of a-priori, analytical shape functions, leverage data-rich computing environment by using *data-driven* modes



Definition of $\boldsymbol{\Phi}_{\boldsymbol{u}}$: Data-Driven Reduction

State-Sensitivity Proper Orthogonal Decomposition (POD)

- Collect state and sensitivity snapshots by sampling HDM

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

- Use Proper Orthogonal Decomposition to generate reduced basis for each individually

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

- Concatenate and orthogonalize to get reduced-order basis

$$\boldsymbol{\Phi}_{\boldsymbol{u}} = \mathrm{QR}\left(\begin{bmatrix} \boldsymbol{u}(\boldsymbol{\mu}^*) & \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}^*) & \boldsymbol{\Phi}_{\boldsymbol{X}} & \boldsymbol{\Phi}_{\boldsymbol{Y}} \end{bmatrix}\right)$$

Definition of $\boldsymbol{\Psi}_{\boldsymbol{u}}$: Minimum-Residual ROM

Least-Squares Petrov-Galerkin (LSPG)[3] projection

$$\boldsymbol{\Psi}_{\boldsymbol{u}} = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{u}} \boldsymbol{\Phi}_{\boldsymbol{u}}$$

Minimum-Residual Property

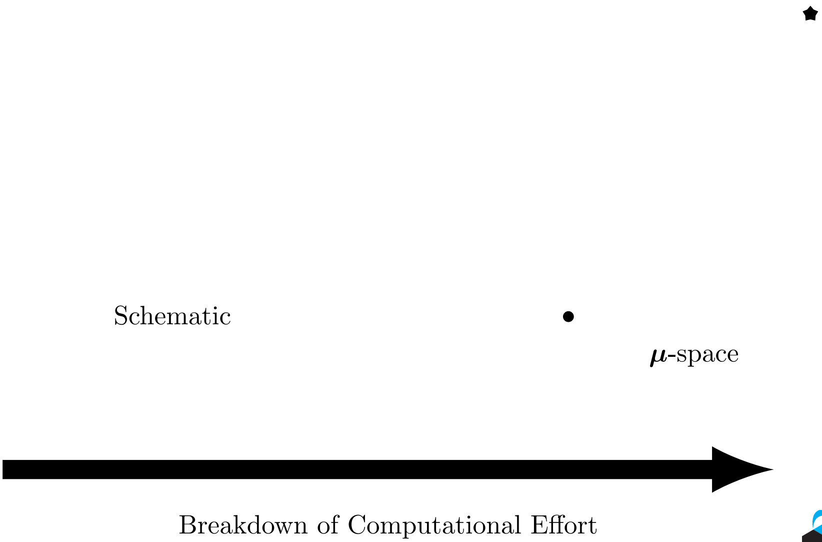
A ROM possesses the minimum-residual property if $\boldsymbol{\Psi}_{\boldsymbol{u}} \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu}) = 0$ is equivalent to the optimality condition of $(\Theta \succ 0)$

$$\underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}}}{\text{minimize}} \quad ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu})||_{\Theta}$$

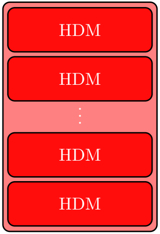
- Implications
  - Recover exact solution when basis not truncated (consistent[3])
  - Monotonic improvement of solution as basis size increases
  - Ensures sensitivity information in $\boldsymbol{\Phi}_{\boldsymbol{u}}$ cannot degrade state approximation[4]
- LSPG possesses minimum-residual property

[3] [Bui-Thanh et al., 2008]
[4] [Fahl, 2001]

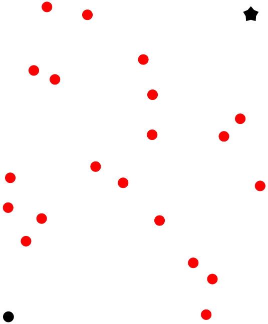
Offline-Online Approach to Optimization



Offline-Online Approach to Optimization



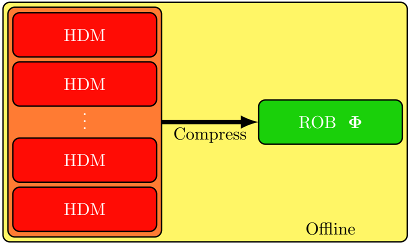
Schematic



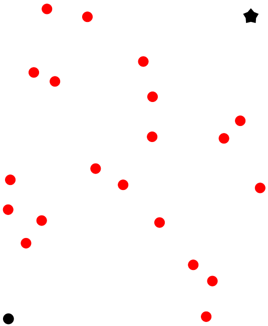
μ-space

Breakdown of Computational Effort

Offline-Online Approach to Optimization



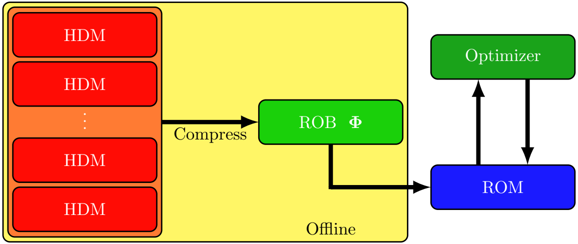
Schematic



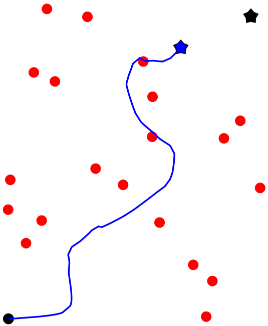
μ-space

Breakdown of Computational Effort

Offline-Online Approach to Optimization



Schematic

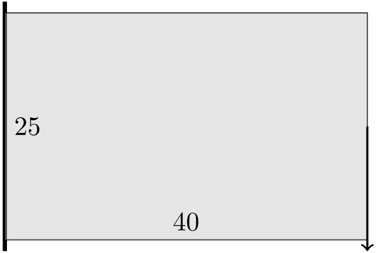


μ-space

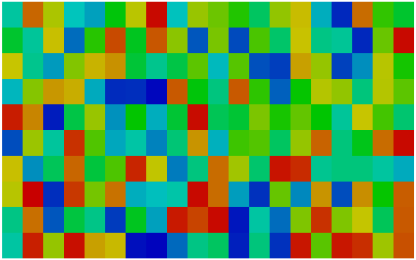
Breakdown of Computational Effort

Numerical Demonstration: Offline-Online Breakdown

- Parameter reduction (Φ_{μ})
  - *apriori spatial clustering*
  - $k_{\mu} = 200$
- *Greedy* Training
  - 5000 candidate points (LHS)
  - 50 snapshots
  - Error indicator: $||\mathbf{r}(\Phi_{u}\boldsymbol{u}_r, \ \Phi_{\mu}\boldsymbol{\mu}_r||$
- State reduction (Φ_{u})
  - POD
  - $k_{u} = 25$
  - Polynomialization acceleration

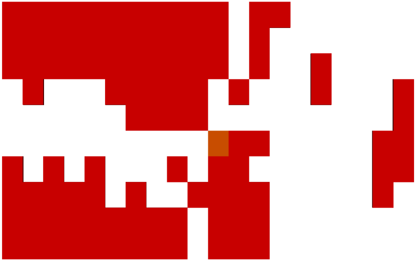


Stiffness maximization, volume constraint

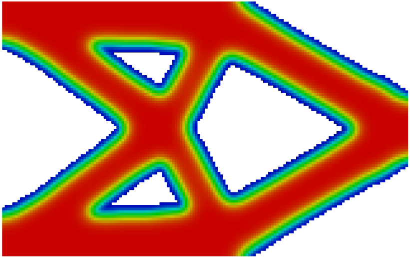


Parametrization with $k_{\mu} = 200$

Numerical Demonstration: Offline-Online Breakdown



Optimal Solution (ROM)



Optimal Solution (HDM)

HDM Solution	**ROB** Construction	**Greedy** Algorithm	**ROM** Optimization
2.84×10^3 s	5.48×10^4 s	1.67×10^5 s	30 s
1.26%	24.36%	74.37%	0.01%

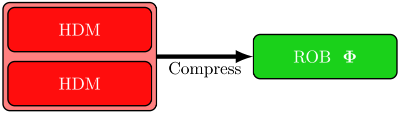
HDM Optimization: 1.97×10^4 s

ROM-Based Trust-Region Framework for Optimization

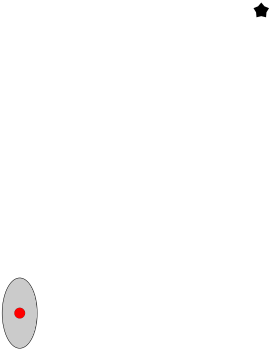
Schematic

Breakdown of Computational Effort

ROM-Based Trust-Region Framework for Optimization



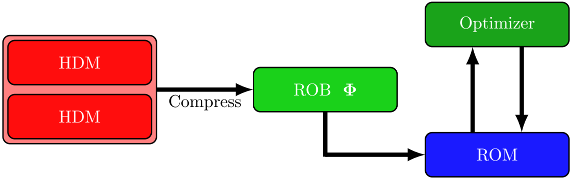
Schematic



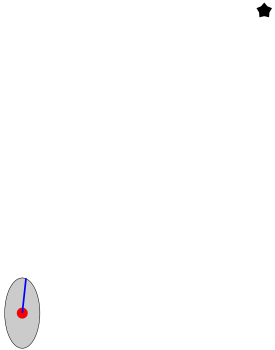
$\boldsymbol{\mu}$-space

Breakdown of Computational Effort

ROM-Based Trust-Region Framework for Optimization



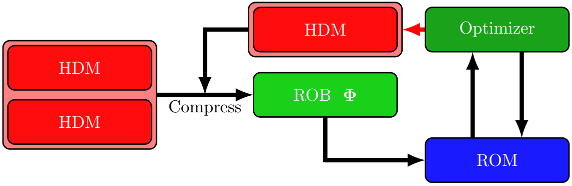
Schematic



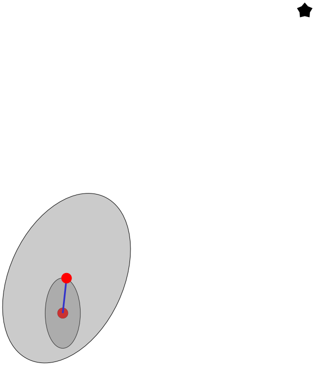
$\boldsymbol{\mu}$-space

Breakdown of Computational Effort

ROM-Based Trust-Region Framework for Optimization



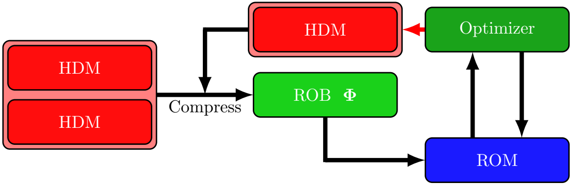
Schematic



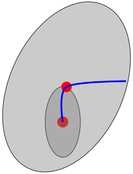
μ-space

Breakdown of Computational Effort

ROM-Based Trust-Region Framework for Optimization



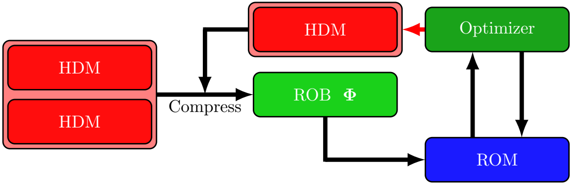
Schematic



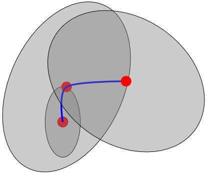
μ-space

Breakdown of Computational Effort

ROM-Based Trust-Region Framework for Optimization



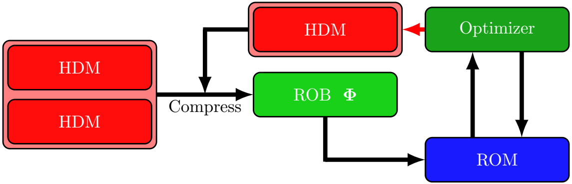
Schematic



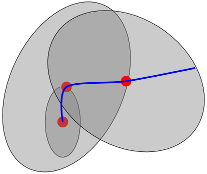
$\boldsymbol{\mu}$-space

Breakdown of Computational Effort

ROM-Based Trust-Region Framework for Optimization



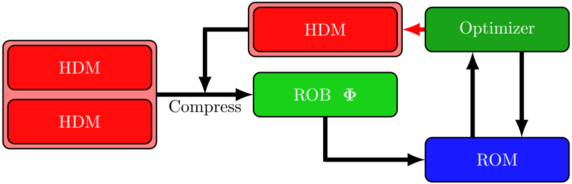
Schematic



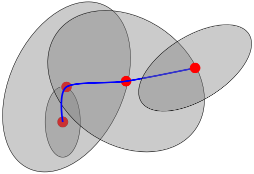
μ-space

Breakdown of Computational Effort

ROM-Based Trust-Region Framework for Optimization



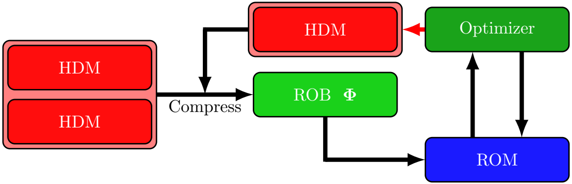
Schematic



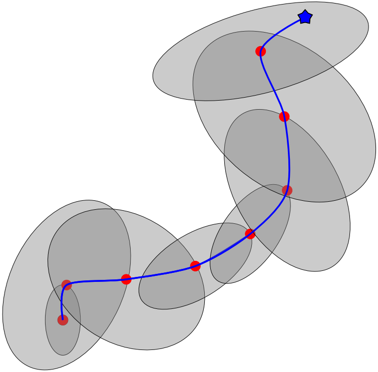
$\boldsymbol{\mu}$-space

Breakdown of Computational Effort

ROM-Based Trust-Region Framework for Optimization



Schematic



$\boldsymbol{\mu}$-space

Breakdown of Computational Effort

Non-Quadratic Trust-Region Method with Adaptive Reduced-Order Models

1: **Initialization**: Build $\boldsymbol{\Phi}_{\boldsymbol{u}}$ from *sparse* training

Non-Quadratic Trust-Region Method with Adaptive Reduced-Order Models

1: **Initialization**: Build $\boldsymbol{\Phi}_{\boldsymbol{u}}$ from *sparse* training

2: **Step computation**: Approximately solve the reduced optimization problem with non-quadratic trust-region for a candidate, $\hat{\boldsymbol{\mu}}_k$

$$\underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu} \in \mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu}) \quad \text{subject to} \quad \boldsymbol{\Psi}_{\boldsymbol{u}}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu}) = 0$$

$$||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu})|| \leq \Delta_k$$

Non-Quadratic Trust-Region Method with Adaptive Reduced-Order Models

1: **Initialization**: Build $\boldsymbol{\Phi}_{\boldsymbol{u}}$ from *sparse* training

2: **Step computation**: Approximately solve the reduced optimization problem with non-quadratic trust-region for a candidate, $\hat{\boldsymbol{\mu}}_k$

$$\underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu} \in \mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu}) \qquad \text{subject to} \quad \boldsymbol{\Psi}_{\boldsymbol{u}}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu}) = 0$$

$$||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu})|| \leq \Delta_k$$

3: **Step acceptance**: Compute

$$\rho_k = \frac{\mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k) - \mathcal{J}(\boldsymbol{u}(\hat{\boldsymbol{\mu}}_k), \hat{\boldsymbol{\mu}}_k)}{\mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k) - \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\hat{\boldsymbol{\mu}}_k), \hat{\boldsymbol{\mu}}_k)}$$

if $\rho_k \geq \eta_0$ **then** $\boldsymbol{\mu}_{k+1} = \hat{\boldsymbol{\mu}}_k$ **else** $\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k$ **end if**

Non-Quadratic Trust-Region Method with Adaptive Reduced-Order Models

1: **Initialization**: Build $\boldsymbol{\Phi}_{\boldsymbol{u}}$ from *sparse* training

2: **Step computation**: Approximately solve the reduced optimization problem with non-quadratic trust-region for a candidate, $\hat{\boldsymbol{\mu}}_k$

$$\underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu} \in \mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \ \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu}) \qquad \text{subject to} \quad \boldsymbol{\Psi}_{\boldsymbol{u}}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu}) = 0$$

$$||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu})|| \leq \Delta_k$$

3: **Step acceptance**: Compute

$$\rho_k = \frac{\mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k) - \mathcal{J}(\boldsymbol{u}(\hat{\boldsymbol{\mu}}_k), \hat{\boldsymbol{\mu}}_k)}{\mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k) - \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\hat{\boldsymbol{\mu}}_k), \hat{\boldsymbol{\mu}}_k)}$$

if $\rho_k \geq \eta_0$ **then** $\boldsymbol{\mu}_{k+1} = \hat{\boldsymbol{\mu}}_k$ **else** $\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k$ **end if**

4: **Trust-region update**:

if $\rho_k \leq \eta_1$ **then** $\Delta_{k+1} \in (0, \gamma||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\hat{\boldsymbol{\mu}}_k), \hat{\boldsymbol{\mu}}_k)||]$ **end if**

if $\rho_k \in (\eta_1, \eta_2)$ **then** $\Delta_{k+1} \in [\gamma||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\hat{\boldsymbol{\mu}}_k), \hat{\boldsymbol{\mu}}_k)||, \Delta_k]$ **end if**

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

Non-Quadratic Trust-Region Method with Adaptive Reduced-Order Models

1: **Initialization**: Build $\boldsymbol{\Phi}_{\boldsymbol{u}}$ from *sparse* training

2: **Step computation**: Approximately solve the reduced optimization problem with non-quadratic trust-region for a candidate, $\hat{\boldsymbol{\mu}}_k$

$$\underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu} \in \mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \ \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu}) \qquad \text{subject to} \quad \boldsymbol{\Psi}_{\boldsymbol{u}}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu}) = 0$$

$$||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\mu})|| \leq \Delta_k$$

3: **Step acceptance**: Compute

$$\rho_k = \frac{\mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k) - \mathcal{J}(\boldsymbol{u}(\hat{\boldsymbol{\mu}}_k), \hat{\boldsymbol{\mu}}_k)}{\mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k) - \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\hat{\boldsymbol{\mu}}_k), \hat{\boldsymbol{\mu}}_k)}$$

if $\rho_k \geq \eta_0$ **then** $\boldsymbol{\mu}_{k+1} = \hat{\boldsymbol{\mu}}_k$ **else** $\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k$ **end if**

4: **Trust-region update**:

if $\rho_k \leq \eta_1$ **then** $\Delta_{k+1} \in (0, \gamma||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\hat{\boldsymbol{\mu}}_k), \hat{\boldsymbol{\mu}}_k)||]$ **end if**

if $\rho_k \in (\eta_1, \eta_2)$ **then** $\Delta_{k+1} \in [\gamma||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\hat{\boldsymbol{\mu}}_k), \hat{\boldsymbol{\mu}}_k)||, \Delta_k]$ **end if**

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

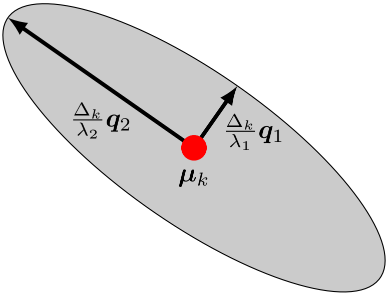
5: **Model update**: Enrich $\boldsymbol{\Phi}_{\boldsymbol{u}}$ with $\boldsymbol{u}(\hat{\boldsymbol{\mu}}_k)$ and $\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\mu}}(\hat{\boldsymbol{\mu}}_k)$

Residual-Based Trust-Region Interpretation

Let $\hat{\boldsymbol{r}}(\boldsymbol{\mu}) = \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r(\boldsymbol{\mu}), \boldsymbol{\mu})$ and $\boldsymbol{A}_k = \frac{\partial \hat{\boldsymbol{r}}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}_k)^T \frac{\partial \hat{\boldsymbol{r}}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}_k) = \boldsymbol{Q}_k \boldsymbol{\Lambda}_k^2 \boldsymbol{Q}_k^T$.

Then, to first order[5],

$$||\hat{\boldsymbol{r}}(\boldsymbol{\mu})||_2 = ||\frac{\partial \hat{\boldsymbol{r}}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}_k)(\boldsymbol{\mu} - \boldsymbol{\mu}_k)||_2 = ||\boldsymbol{\mu} - \boldsymbol{\mu}_k||_{\boldsymbol{A}_k} \leq \Delta_k$$



Annotated schematic of trust-region: $\boldsymbol{q}_i = \boldsymbol{Q}_k \boldsymbol{e}_i$ and $\lambda_i = \boldsymbol{e}_i^T \boldsymbol{\Lambda}_k \boldsymbol{e}_i$

[5] assuming $\hat{\boldsymbol{r}}(\boldsymbol{\mu}_k) = 0$, i.e., ROM exact at trust-region center

Convergence to Critical Point of *Unreduced* Problem

Lim-Inf Convergence to Critical Point of Unreduced Optimization Problem

Let $\{\boldsymbol{\mu}_k\}$ be a sequence of iterations produced by the Algorithm and suppose

- $\mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k) = \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k)$
- There exists $\xi > 0$ such that

$$||\nabla \mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k) - \nabla \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k)|| \leq \xi ||\nabla \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k)||$$

- There exists $\zeta > 0$ such that for all $\boldsymbol{\mu} \in \{\boldsymbol{\mu} \mid ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}), \boldsymbol{\mu})|| \leq \Delta_k\}$

$$|\mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}), \boldsymbol{\mu}) - \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}), \boldsymbol{\mu})| \leq \zeta ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}), \boldsymbol{\mu})||.$$

Then

$$\liminf_{k \to \infty} ||\nabla \mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}_k), \ \boldsymbol{\mu}_k)|| = 0$$

Assumptions of Convergence Theory Hold

If $\boldsymbol{\mu}_k$ is a *training* point, then

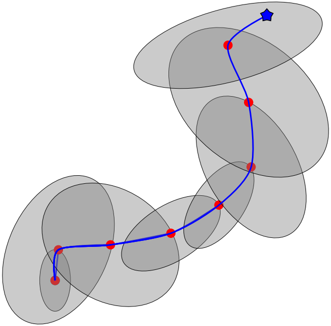
- **Minimum-residual** formulation for the **primal** reduced-order model implies

$$\mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k) = \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k)$$

- **Minimum-residual** formulation for the reduced-order model **sensitivity** implies

$$\nabla \mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k) = \nabla \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}_k), \boldsymbol{\mu}_k)$$

- Standard **residual-based error estimation** implies, for some $\zeta > 0$,

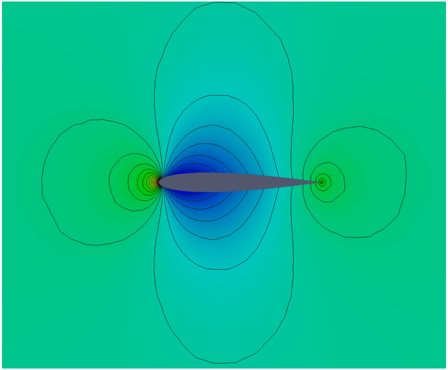


$\boldsymbol{\mu}$-space

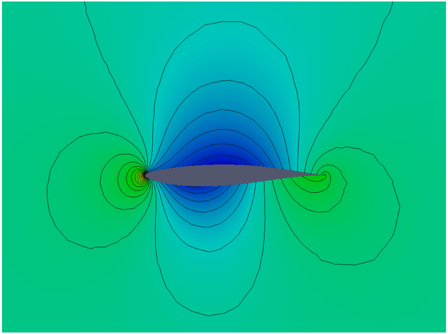
$$|\mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}), \boldsymbol{\mu}) - \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}), \boldsymbol{\mu})| \leq \zeta ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r(\boldsymbol{\mu}), \boldsymbol{\mu})||$$

Compressible, Inviscid Airfoil Inverse Design

Pressure discrepancy minimization (Euler equations)



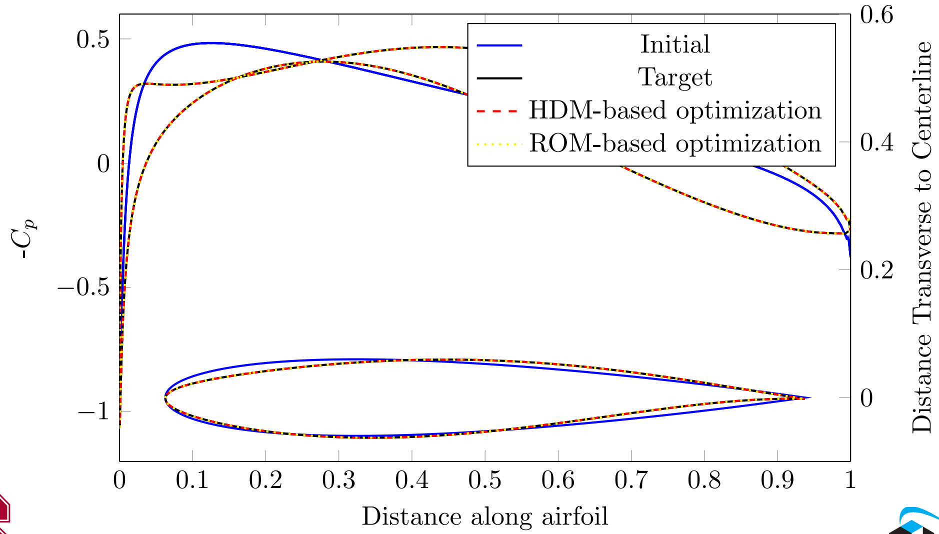
NACA0012: Initial



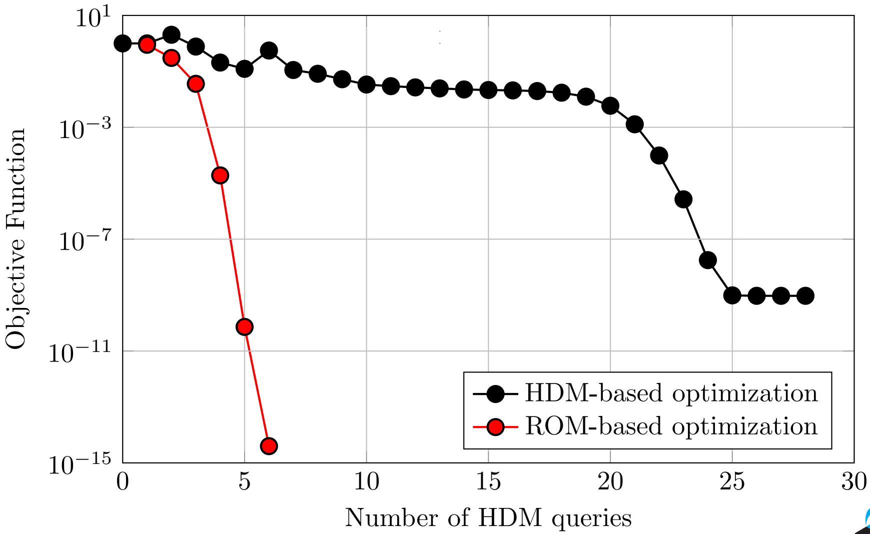
RAE2822: Target

Pressure field for airfoil configurations at $M_\infty = 0.5$, $\alpha = 0.0°$

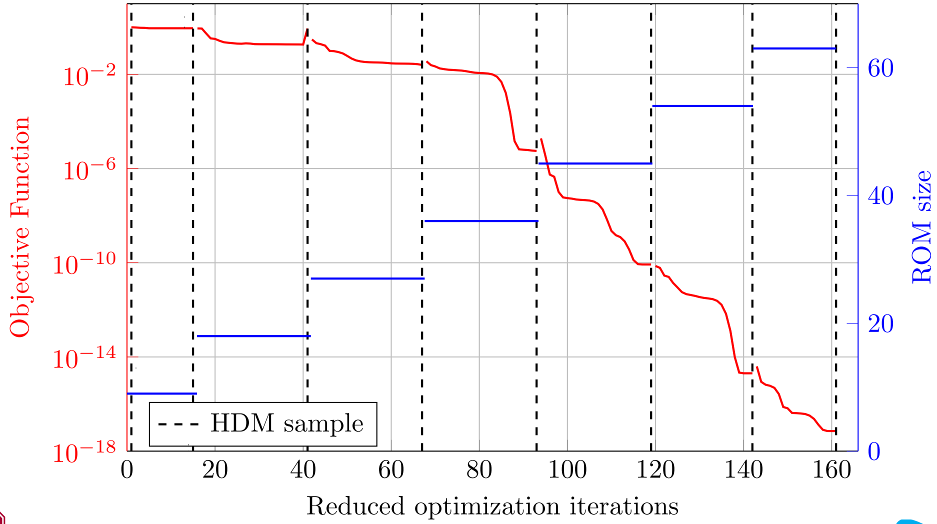
ROM-Constrained Optimization Solver Recovers Target



ROM Solver Requires 4× Fewer HDM Queries

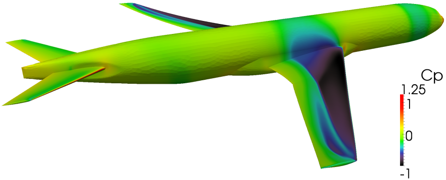


At the Cost of ROM Queries

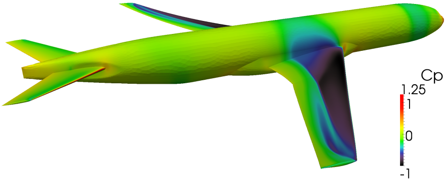


Next: Shape Optimization of Full Aircraft (CRM)

ROMs are fast, accurate, and require limited resources



HDM solution (Drag $= 142.336kN$)



ROM solution (Drag $= 142.304kN$)

- HDM: 70×10^6 DOF, **2hr on 1024** Intel Xeon E5-2698 v3 cores (2.3GHz)
- ROM: **170s on 2** Intel i7 cores (1.8GHz)
- Relative error in drag 0.022%
- CPU-time speedup greater than 2.15×10^4
- Wall-time speedup greater than 42
- *Washabaugh, Zahr, Farhat (AIAA, 2016)*

PDE-Constrained Optimization II

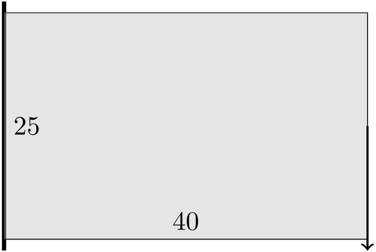
Goal: Rapidly solve PDE-constrained optimization problem of the form

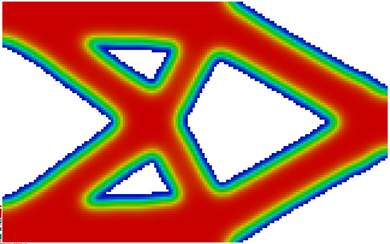
$$\begin{aligned} \underset{\boldsymbol{u}\in\mathbb{R}^{n_{\boldsymbol{u}}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{u},\ \boldsymbol{\mu}) \\ \text{subject to} \quad & \boldsymbol{r}(\boldsymbol{u},\ \boldsymbol{\mu}) = 0 \\ & \boldsymbol{c}(\boldsymbol{u},\ \boldsymbol{\mu}) \geq 0 \end{aligned}$$

where

- $\boldsymbol{r} : \mathbb{R}^{n_{\boldsymbol{u}}} \times \mathbb{R}^{n_{\boldsymbol{\mu}}} \to \mathbb{R}^{n_{\boldsymbol{u}}}$ is the discretized partial differential equation
- $\mathcal{J} : \mathbb{R}^{n_{\boldsymbol{u}}} \times \mathbb{R}^{n_{\boldsymbol{\mu}}} \to \mathbb{R}$ is the objective function
- $\boldsymbol{c} : \mathbb{R}^{n_{\boldsymbol{u}}} \times \mathbb{R}^{n_{\boldsymbol{\mu}}} \to \mathbb{R}^{n_{\boldsymbol{c}}}$ are the side constraints
- $\boldsymbol{u} \in \mathbb{R}^{n_{\boldsymbol{u}}}$ is the PDE state vector
- $\boldsymbol{\mu} \in \mathbb{R}^{n_{\boldsymbol{\mu}}}$ is the vector of parameters

Problem Setup





- 16000 8-node brick elements, 77760 dofs
- Total Lagrangian form, finite strain, StVK[6]
- St. Venant-Kirchhoff material
- Sparse Cholesky linear solver (CHOLMOD[7])
- Newton-Raphson nonlinear solver
- Minimum compliance optimization problem

$$\begin{aligned} \underset{\boldsymbol{u}\in\mathbb{R}^{n_{\boldsymbol{u}}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad & \boldsymbol{f}_{\text{ext}}{}^T \boldsymbol{u} \\ \text{subject to} \quad & V(\boldsymbol{\mu}) \leq \frac{1}{2} V_0 \\ & \boldsymbol{r}(\boldsymbol{u},\ \boldsymbol{\mu}) = 0 \end{aligned}$$

- Gradient computations: Adjoint method
- Optimizer: SNOPT [Gill et al., 2002]

[6] [Bonet and Wood, 1997, Belytschko et al., 2000]
[7] [Chen et al., 2008]

Restrict Parameter Space to Low-Dimensional Subspace

- *Restrict parameter to a low-dimensional subspace*

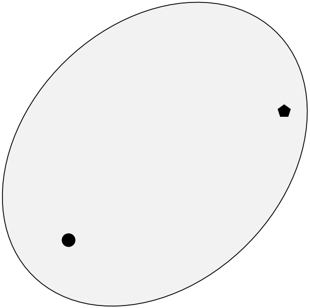
$$\boldsymbol{\mu} \approx \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r$$

  - $\boldsymbol{\Phi}_{\boldsymbol{\mu}} = \begin{bmatrix} \phi_{\boldsymbol{\mu}}^1 & \cdots & \phi_{\boldsymbol{\mu}}^{k_{\boldsymbol{\mu}}} \end{bmatrix} \in \mathbb{R}^{n_{\boldsymbol{\mu}} \times k_{\boldsymbol{\mu}}}$ is the reduced basis
  - $\boldsymbol{\mu}_r \in \mathbb{R}^{k_{\boldsymbol{\mu}}}$ are the reduced coordinates of $\boldsymbol{\mu}$
  - $n_{\boldsymbol{\mu}} \gg k_{\boldsymbol{\mu}}$
- Substitute restriction into reduced-order model to obtain

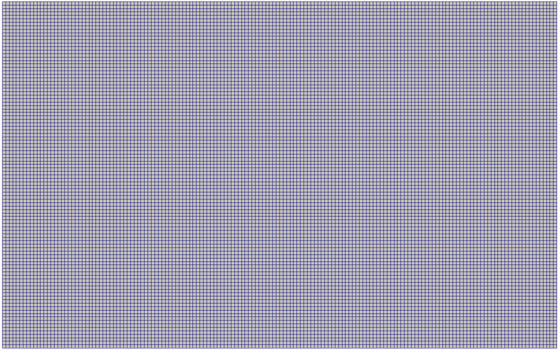
$$\boldsymbol{\Phi}_{\boldsymbol{u}}{}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r, \ \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r) = 0$$

- Related work:
[Maute and Ramm, 1995, Lieberman et al., 2010, Constantine et al., 2014]

Restrict Parameter Space to Low-Dimensional Subspace

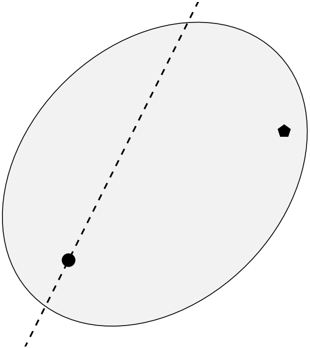


$\boldsymbol{\mu}$-space

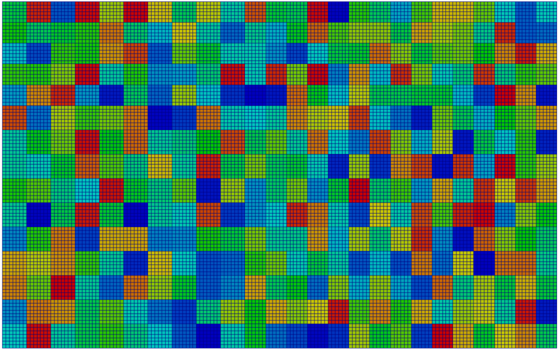


Background mesh

Restrict Parameter Space to Low-Dimensional Subspace



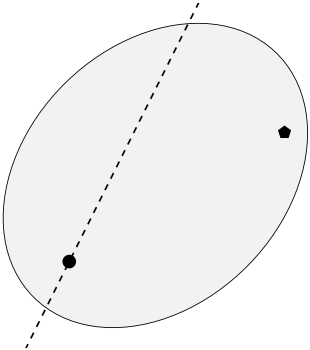
$\boldsymbol{\mu}$-space



Macroelements

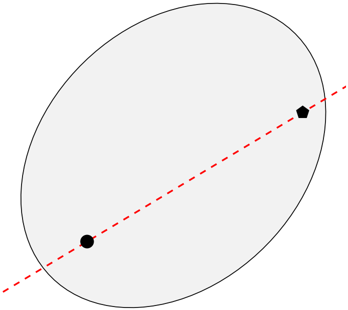
Optimality Conditions to Adapt Reduced-Order Basis, Φ_μ

- Selection of Φ_μ amounts to a *restriction* of the parameter space



Optimality Conditions to Adapt Reduced-Order Basis, Φ_μ

- Selection of Φ_μ amounts to a *restriction* of the parameter space
- Adaptation of Φ_μ should attempt to include the optimal solution in the restricted parameter space, i.e. $\mu^* \in \text{col}(\Phi_\mu)$
- Adaptation based on **first-order optimality conditions** of HDM optimization problem



Optimality Conditions to Adapt Reduced-Order Basis, Φ_{μ}

Lagrangian

$$\mathcal{L}(\boldsymbol{\mu},\ \boldsymbol{\lambda}) = \mathcal{J}(\boldsymbol{u}(\boldsymbol{\mu}),\ \boldsymbol{\mu}) - \boldsymbol{\lambda}^T \boldsymbol{c}(\boldsymbol{u}(\boldsymbol{\mu}),\ \boldsymbol{\mu})$$

Karush-Kuhn Tucker (KKT) Conditions[8]

$$\begin{aligned}
\nabla_{\boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\mu},\ \boldsymbol{\lambda}) &= 0 \\
\boldsymbol{\lambda} &\geq 0 \\
\boldsymbol{\lambda}_i \boldsymbol{c}_i(\boldsymbol{u}(\boldsymbol{\mu}),\ \boldsymbol{\mu}) &= 0 \\
\boldsymbol{c}(\boldsymbol{u}(\boldsymbol{\mu}), \boldsymbol{\mu}) &\geq 0
\end{aligned}$$

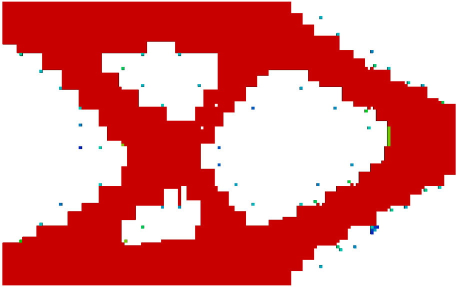
[8] [Nocedal and Wright, 2006]

Lagrangian Gradient Refinement Indicator

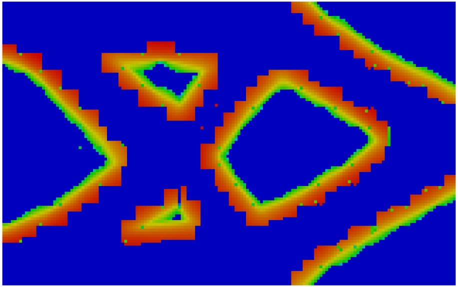
- From Lagrange multiplier estimates, only KKT condition not satisfied automatically:

$$\nabla_{\boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\lambda}) = 0$$

- Use $|\nabla_{\boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\lambda})|$ as indicator for **refinement** of discretization of $\boldsymbol{\mu}$-space



$\boldsymbol{\mu}$

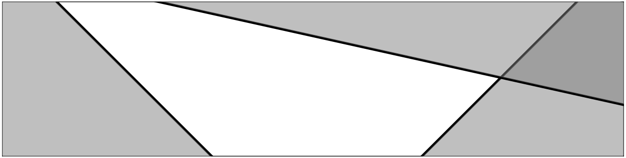


$|\nabla_{\boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\lambda})|$

Constraints may lead to infeasible sub-problems

Non-Quadratic Trust-Region MOR [Zahr and Farhat, 2014]

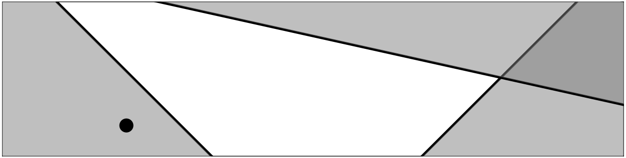
$$\begin{aligned}
\underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu}_r \in \mathbb{R}^{k_{\boldsymbol{\mu}}}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) \\
\text{subject to} \quad & \boldsymbol{c}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) \geq 0 \\
& \boldsymbol{\Psi}_{\boldsymbol{u}}{}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) = 0 \\
& ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r)|| \leq \Delta
\end{aligned}$$



Constraints may lead to infeasible sub-problems

Non-Quadratic Trust-Region MOR [Zahr and Farhat, 2014]

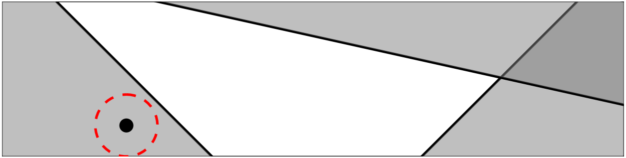
$$\begin{aligned} \underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu}_r \in \mathbb{R}^{k_{\boldsymbol{\mu}}}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) \\ \text{subject to} \quad & \boldsymbol{c}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) \geq 0 \\ & \boldsymbol{\Psi}_{\boldsymbol{u}}{}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) = 0 \\ & ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r)|| \leq \Delta \end{aligned}$$



Constraints may lead to infeasible sub-problems

Non-Quadratic Trust-Region MOR [Zahr and Farhat, 2014]

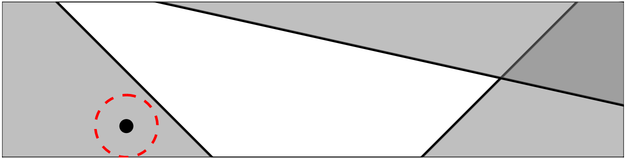
$$\begin{aligned}
\underset{\boldsymbol{u}_r\in\mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu}_r\in\mathbb{R}^{k_{\boldsymbol{\mu}}}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) \\
\text{subject to} \quad & \boldsymbol{c}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) \geq 0 \\
& \boldsymbol{\Psi}_{\boldsymbol{u}}{}^T\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) = 0 \\
& ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r)|| \leq \Delta
\end{aligned}$$



Elastic constraints to circumvent infeasible subproblems

Constrained Non-Quadratic Trust-Region MOR (CNQTR-MOR)

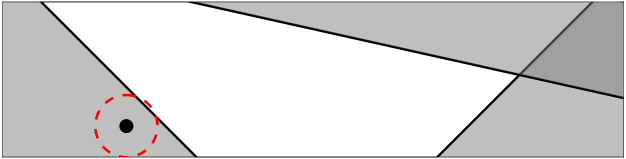
$$
\begin{aligned}
\underset{\boldsymbol{u}_r\in\mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu}_r\in\mathbb{R}^{k_{\boldsymbol{\mu}}},\ \mathbf{t}\in\mathbb{R}^{n_{\boldsymbol{c}}}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) - \gamma \mathbf{t}^T\mathbf{1} \\
\text{subject to} \quad & \boldsymbol{c}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) \geq \mathbf{t} \\
& \boldsymbol{\Psi}_{\boldsymbol{u}}{}^T\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) = 0 \\
& ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r)|| \leq \Delta \\
& \mathbf{t} \leq 0
\end{aligned}
$$



Elastic constraints to circumvent infeasible subproblems

Constrained Non-Quadratic Trust-Region MOR (CNQTR-MOR)

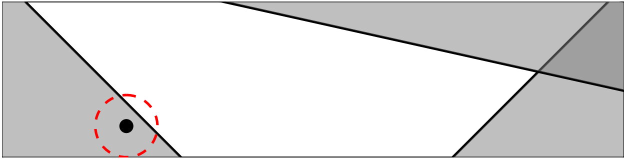
$$
\begin{aligned}
\underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu}_r \in \mathbb{R}^{k_{\boldsymbol{\mu}}},\ \mathbf{t} \in \mathbb{R}^{n_{\boldsymbol{c}}}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r) - \gamma \mathbf{t}^T \mathbf{1} \\
\text{subject to} \quad & \boldsymbol{c}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r) \geq \mathbf{t} \\
& \boldsymbol{\Psi}_{\boldsymbol{u}}{}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r) = 0 \\
& ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r)|| \leq \Delta \\
& \mathbf{t} \leq 0
\end{aligned}
$$



Elastic constraints to circumvent infeasible subproblems

Constrained Non-Quadratic Trust-Region MOR (CNQTR-MOR)

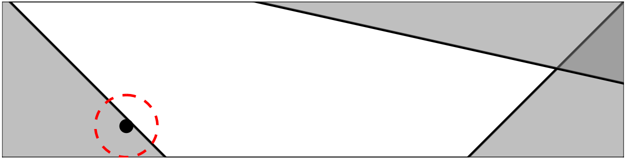
$$
\begin{aligned}
\underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu}_r \in \mathbb{R}^{k_{\boldsymbol{\mu}}},\ \mathbf{t} \in \mathbb{R}^{n_{\boldsymbol{c}}}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) - \gamma \mathbf{t}^T \mathbf{1} \\
\text{subject to} \quad & \boldsymbol{c}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) \geq \mathbf{t} \\
& \boldsymbol{\Psi}_{\boldsymbol{u}}{}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) = 0 \\
& ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r)|| \leq \Delta \\
& \mathbf{t} \leq 0
\end{aligned}
$$



Elastic constraints to circumvent infeasible subproblems

Constrained Non-Quadratic Trust-Region MOR (CNQTR-MOR)

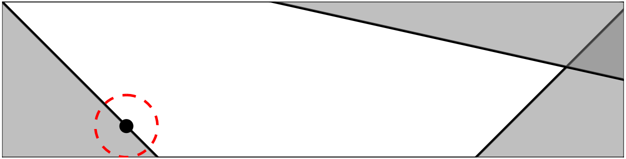
$$\begin{aligned}
\underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu}_r \in \mathbb{R}^{k_{\boldsymbol{\mu}}},\ \mathbf{t} \in \mathbb{R}^{n_{\boldsymbol{c}}}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) - \gamma \mathbf{t}^T \mathbf{1} \\
\text{subject to} \quad & \boldsymbol{c}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) \geq \mathbf{t} \\
& \boldsymbol{\Psi}_{\boldsymbol{u}}{}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) = 0 \\
& ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r)|| \leq \Delta \\
& \mathbf{t} \leq 0
\end{aligned}$$



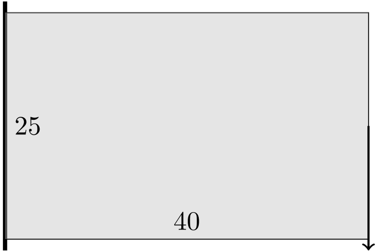
Elastic constraints to circumvent infeasible subproblems

Constrained Non-Quadratic Trust-Region MOR (CNQTR-MOR)

$$\begin{aligned} \underset{\boldsymbol{u}_r \in \mathbb{R}^{k_{\boldsymbol{u}}},\ \boldsymbol{\mu}_r \in \mathbb{R}^{k_{\boldsymbol{\mu}}},\ \mathbf{t} \in \mathbb{R}^{n_{\boldsymbol{c}}}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r) - \gamma \mathbf{t}^T \mathbf{1} \\ \text{subject to} \quad & \boldsymbol{c}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r) \geq \mathbf{t} \\ & \boldsymbol{\Psi}_{\boldsymbol{u}}{}^T \boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r) = 0 \\ & ||\boldsymbol{r}(\boldsymbol{\Phi}_{\boldsymbol{u}} \boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r)|| \leq \Delta \\ & \mathbf{t} \leq 0 \end{aligned}$$



Compliance Minimization: 2D Cantilever



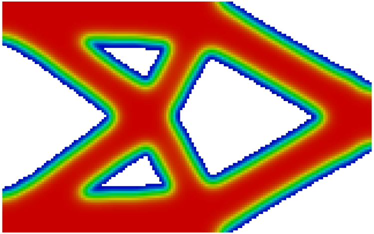
- 16000 8-node brick elements, 77760 dofs
- Total Lagrangian form, finite strain, StVK[9]
- St. Venant-Kirchhoff material
- Sparse Cholesky linear solver (CHOLMOD[10])
- Newton-Raphson nonlinear solver
- Minimum compliance optimization problem

$$\begin{aligned} \underset{\boldsymbol{u}\in\mathbb{R}^{n_{\boldsymbol{u}}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad & {\boldsymbol{f}_{\text{ext}}}^T \boldsymbol{u} \\ \text{subject to} \quad & V(\boldsymbol{\mu}) \leq \frac{1}{2} V_0 \\ & \boldsymbol{r}(\boldsymbol{u},\ \boldsymbol{\mu}) = 0 \end{aligned}$$

- Gradient computations: Adjoint method
- Optimizer: SNOPT [Gill et al., 2002]
- Maximum ROM size: $k_{\boldsymbol{u}} \leq 5$

[9] [Bonet and Wood, 1997, Belytschko et al., 2000]
[10] [Chen et al., 2008]

Order of Magnitude Speedup to Suboptimal Solution



HDM

CNQTR-MOR + $\boldsymbol{\Phi}_\mu$ adaptivity

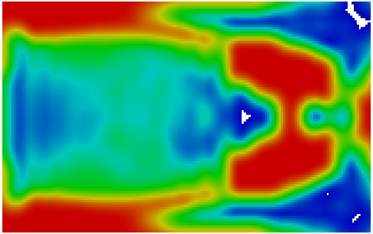
HDM Solution	**HDM** Gradient	**HDM** Optimization
7458s (450)	4018s (411)	8284s

HDM
Elapsed time = 19761s

HDM Solution	**HDM** Gradient	**ROB** Construction	**ROM** Optimization
1049s (64)	88s (9)	727s (56)	39s (3676)

CNQTR-MOR + $\boldsymbol{\Phi}_\mu$ adaptivity
Elapsed time = 2197s, Speedup ≈ 9x

Better Solution after 64 HDM Evaluations

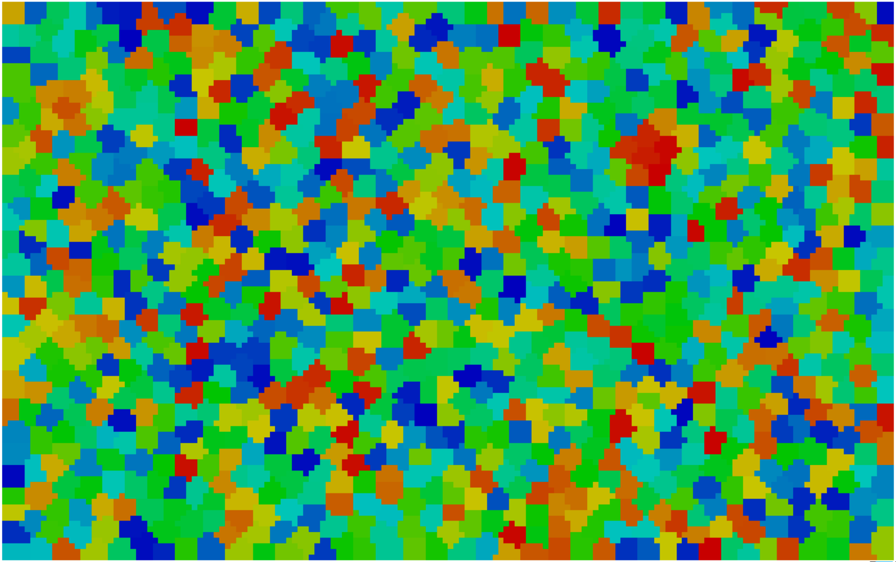


HDM

CNQTR-MOR + $\boldsymbol{\Phi}_\mu$ adaptivity

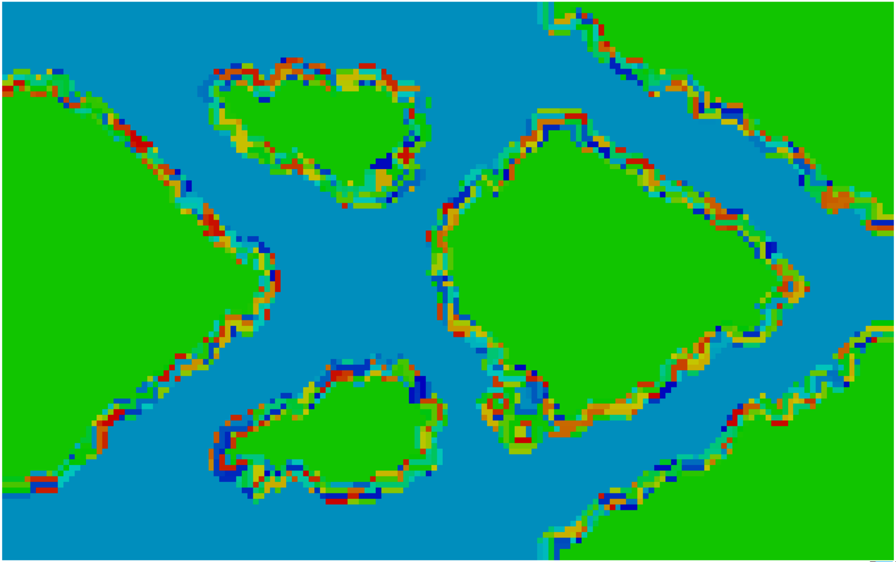
- CNQTR-MOR + $\boldsymbol{\Phi}_\mu$ adaptivity: superior approximation to optimal solution than HDM approach after fixed number of HDM solves (64)
- Reasonable option to *warm-start* HDM topology optimization

Macro-element Evolution



Iteration 0 (1000)

Macro-element Evolution



Iteration 1 (977)

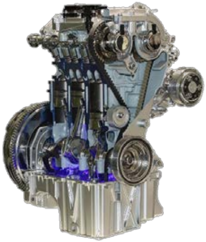
CNQTR-MOR + Φ_μ adaptivity

An Adaptive Reduction Framework for Optimization under Uncertainty

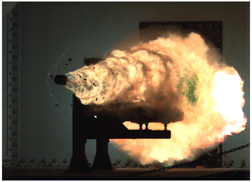
- Highly volatile systems tend to be plagued by uncertainties, which must be quantified for meaningful problem formulation
- Optimize *moments* of quantities of interest of stochastic partial differential equation

$$\begin{aligned} \underset{\boldsymbol{u}\in\mathbb{R}^{n_{\boldsymbol{u}}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad & \int_{\boldsymbol{\Xi}} \mathcal{J}(\boldsymbol{u},\ \boldsymbol{\mu};\ \boldsymbol{\xi})\, d\boldsymbol{\xi} \\ \text{subject to} \quad & \boldsymbol{r}(\boldsymbol{u},\ \boldsymbol{\mu};\ \boldsymbol{\xi}) = 0 \qquad \boldsymbol{\xi} \in \boldsymbol{\Xi} \end{aligned}$$

- Combine adaptive model reduction framework with dimension-adaptive sparse grids to **enable** stochastic optimization



Engine System



EM Launcher

Collaborators: Drew Kouri (Sandia NM), Kevin Carlberg (Sandia CA)

High-Order Methods for Optimization of Conservation Laws

- Derived, implemented fully discrete adjoint method for globally high-order discretization of conservation laws on deforming domains
- *Incorporation of time-periodicity constraints*

Energy = 9.4096e+00	Energy = 4.9476e+00	Energy = 4.6110e+00
Thrust = 1.7660e-01	Thrust = 2.5000e+00	Thrust = 2.5000e+00

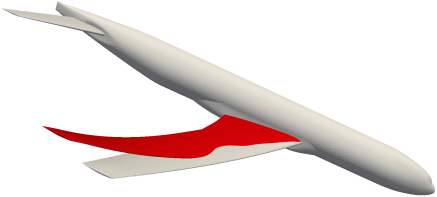
Initial	Optimal Control	Optimal Shape/Control

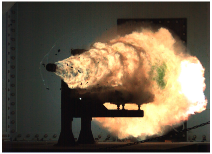
Collaborators: Per-Olof Persson (UCB, LBNL), Jon Wilkening (UCB, LBNL)

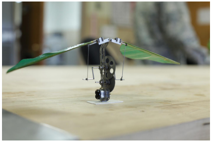
Approaching Many-Query, Extreme-Scale Computational Physics

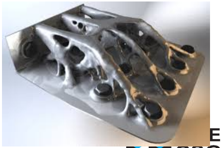
Leveraging Inexactness For Acceleration of Many-Query Multiphysics Problems

- Framework introduced for accelerating PDE-constrained optimization problems with **side constraints** and **large-dimensional parameter space**
  - Adaptive reduction of state and parameter spaces
- Applied to aerodynamic design and topology optimization
  - Order of magnitude speedup speedup observed
  - Competitive warm-start method

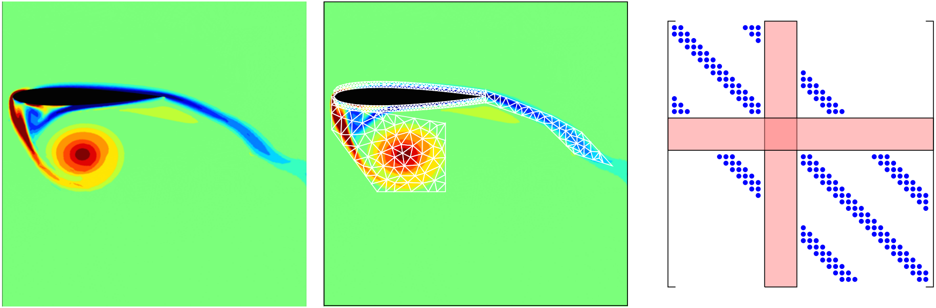








Faster Computational Physics: Adaptive Data-Driven Discretization



(a) Vorticity around heaving airfoil (b) Potential Ω^l, Ω^g decomposition (c) Idealized sparsity structure

- Methods to *transform* features in global basis functions - minimize reliance on local shape functions
- Linear algebra for sparse operators with a few dense rows and columns
- Elements of: **high-order methods**, **adaptive mesh refinement**, **numerical linear algebra**

Fewer Queries: Second-Order Methods for Accelerated Convergence

Hessian information highly desired in optimization and UQ, but expensive due to $\mathcal{O}(N_{\boldsymbol{\mu}})$ required linear system solves

Sensitivity/Adjoint Method for Computing Hessian

$$\frac{\mathrm{d}^2\mathcal{J}}{\mathrm{d}\boldsymbol{\mu}_j\mathrm{d}\boldsymbol{\mu}_k} = \frac{\partial^2\mathcal{J}}{\partial\boldsymbol{\mu}_j\partial\boldsymbol{\mu}_k} + \frac{\partial^2\mathcal{J}}{\partial\boldsymbol{\mu}_j\partial\boldsymbol{u}}\frac{\partial\boldsymbol{u}}{\partial\boldsymbol{\mu}_k} + \frac{\partial\boldsymbol{u}}{\partial\boldsymbol{\mu}_j}^T\frac{\partial^2\mathcal{J}}{\partial\boldsymbol{u}\partial\boldsymbol{\mu}_k} + \frac{\partial\boldsymbol{u}}{\partial\boldsymbol{\mu}_j}^T\frac{\partial^2\mathcal{J}}{\partial\boldsymbol{u}\partial\boldsymbol{u}}\frac{\partial\boldsymbol{u}}{\partial\boldsymbol{\mu}_k}$$
$$-\frac{\partial\mathcal{J}}{\partial\boldsymbol{u}}\frac{\partial\boldsymbol{r}}{\partial\boldsymbol{u}}^{-1}\left[\frac{\partial^2\boldsymbol{r}}{\partial\boldsymbol{\mu}_j\partial\boldsymbol{\mu}_k} + \frac{\partial^2\boldsymbol{r}}{\partial\boldsymbol{\mu}_j\partial\boldsymbol{u}}\frac{\partial\boldsymbol{u}}{\partial\boldsymbol{\mu}_k} + \frac{\partial^2\boldsymbol{r}}{\partial\boldsymbol{\mu}_k\partial\boldsymbol{u}}\frac{\partial\boldsymbol{u}}{\partial\boldsymbol{\mu}_j} + \frac{\partial^2\boldsymbol{r}}{\partial\boldsymbol{u}\partial\boldsymbol{u}} : \frac{\partial\boldsymbol{u}}{\partial\boldsymbol{\mu}_j}\otimes\frac{\partial\boldsymbol{u}}{\partial\boldsymbol{\mu}_k}\right]$$

where

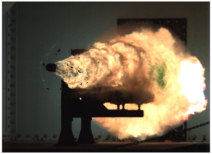
$$\frac{\partial\boldsymbol{u}}{\partial\boldsymbol{\mu}_j} = \frac{\partial\boldsymbol{r}}{\partial\boldsymbol{u}}^{-1}\frac{\partial\boldsymbol{r}}{\partial\boldsymbol{\mu}_j}$$

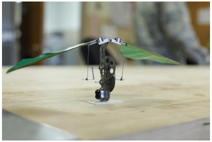
- Fast, *multiple right-hand side* linear solver by building data-driven subspace for image of $\frac{\partial\boldsymbol{r}}{\partial\boldsymbol{u}}^{-1}, \frac{\partial\boldsymbol{r}}{\partial\boldsymbol{u}}^{-T}$
- MOR concepts in context of **numerical linear algebra**

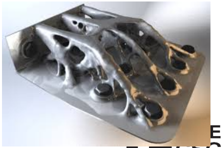
Approaching Many-Query, Extreme-Scale Computational Physics

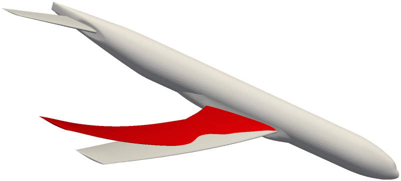
Leveraging Inexactness For Acceleration of Many-Query Multiphysics Problems

- Framework introduced for accelerating PDE-constrained optimization problems with **side constraints** and **large-dimensional parameter space**
  - Adaptive reduction of state and parameter spaces
- Applied to aerodynamic design and topology optimization
  - Order of magnitude speedup speedup observed
  - Competitive warm-start method
- **Future work:** combine advantages of MOR/AMR for drastic computational savings with *in-situ* training; second-order methods for rapidly converging many-query algorithms; new (multiphysics) applications









Acknowledgement

References I

Barbič, J. and James, D. (2007).
Time-critical distributed contact for 6-dof haptic rendering of adaptively sampled reduced deformable models.
In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 171–180. Eurographics Association.

Barrault, M., Maday, Y., Nguyen, N. C., and Patera, A. T. (2004).
An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations.
Comptes Rendus Mathematique, 339(9):667–672.

Belytschko, T., Liu, W., Moran, B., et al. (2000).
Nonlinear finite elements for continua and structures, volume 26.
Wiley New York.

Bonet, J. and Wood, R. (1997).
Nonlinear continuum mechanics for finite element analysis.
Cambridge university press.

Bui-Thanh, T., Willcox, K., and Ghattas, O. (2008).
Model reduction for large-scale systems with high-dimensional parametric input space.
SIAM Journal on Scientific Computing, 30(6):3270–3288.

References II

Carlberg, K., Bou-Mosleh, C., and Farhat, C. (2011).
Efficient non-linear model reduction via a least-squares petrov–galerkin projection and compressive tensor approximations.
International Journal for Numerical Methods in Engineering, 86(2):155–181.

Chapman, T., Collins, P., Avery, P., and Farhat, C. (2015).
Accelerated mesh sampling for model hyper reduction.
International Journal for Numerical Methods in Engineering.

Chaturantabut, S. and Sorensen, D. C. (2010).
Nonlinear model reduction via discrete empirical interpolation.
SIAM Journal on Scientific Computing, 32(5):2737–2764.

Chen, Y., Davis, T. A., Hager, W. W., and Rajamanickam, S. (2008).
Algorithm 887: Cholmod, supernodal sparse cholesky factorization and update/downdate.
ACM Transactions on Mathematical Software (TOMS), 35(3):22.

Constantine, P. G., Dow, E., and Wang, Q. (2014).
Active subspace methods in theory and practice: Applications to kriging surfaces.
SIAM Journal on Scientific Computing, 36(4):A1500–A1524.

References III

Fahl, M. (2001).
Trust-region methods for flow control based on reduced order modelling.
PhD thesis, Universitätsbibliothek.

Gill, P. E., Murray, W., and Saunders, M. A. (2002).
Snopt: An sqp algorithm for large-scale constrained optimization.
SIAM journal on optimization, 12(4):979–1006.

Lawson, C. L. and Hanson, R. J. (1974).
Solving least squares problems, volume 161.
SIAM.

Lieberman, C., Willcox, K., and Ghattas, O. (2010).
Parameter and state model reduction for large-scale statistical inverse problems.
SIAM Journal on Scientific Computing, 32(5):2523–2542.

Maute, K. and Ramm, E. (1995).
Adaptive topology optimization.
Structural optimization, 10(2):100–112.

References IV

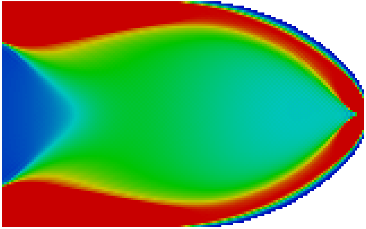
Nguyen, N. and Peraire, J. (2008).
An efficient reduced-order modeling approach for non-linear parametrized partial differential equations.
International journal for numerical methods in engineering, 76(1):27–55.

Nocedal, J. and Wright, S. (2006).
Numerical optimization, series in operations research and financial engineering.
Springer.

Rewienski, M. J. (2003).
A trajectory piecewise-linear approach to model order reduction of nonlinear dynamical systems.
PhD thesis, Citeseer.

Zahr, M. J. and Farhat, C. (2014).
Progressive construction of a parametric reduced-order model for pde-constrained optimization.
International Journal for Numerical Methods in Engineering.

Standard Difficulty: Binary Solutions



(a) Without penalization

Standard Difficulty: Binary Solutions

Relaxed, Penalized Problem Setup

$$\underset{\boldsymbol{u}\in\mathbb{R}^{n_{\boldsymbol{u}}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad \boldsymbol{f}_{\text{ext}}{}^T\boldsymbol{u}$$

$$\text{subject to} \quad V(\boldsymbol{\mu}) \leq \frac{1}{2}V_0$$

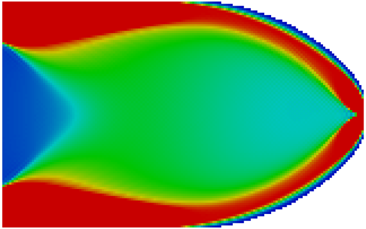
$$\mathbf{r}(\boldsymbol{u},\ \boldsymbol{\mu}^p) = 0$$

$$\boldsymbol{\mu} \in [0,1]^{k_{\boldsymbol{\mu}}}$$

Effect of Penalization

$$\mathbf{K}^e \leftarrow (\boldsymbol{\mu}^e)^p\mathbf{K}^e$$

- $\mathbf{K}^e$: eth element stiffness matrix



(a) Without penalization

Standard Difficulty: Binary Solutions

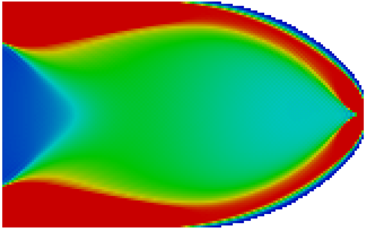
Relaxed, Penalized Problem Setup

$$\begin{aligned} \underset{\boldsymbol{u}\in\mathbb{R}^{n_{\boldsymbol{u}}},\ \boldsymbol{\mu}\in\mathbb{R}^{n_{\boldsymbol{\mu}}}}{\text{minimize}} \quad & \boldsymbol{f}_{\text{ext}}{}^{T}\boldsymbol{u} \\ \text{subject to} \quad & V(\boldsymbol{\mu}) \leq \frac{1}{2}V_0 \\ & \mathbf{r}(\boldsymbol{u},\ \boldsymbol{\mu}^p) = 0 \\ & \boldsymbol{\mu} \in [0,1]^{k_{\boldsymbol{\mu}}} \end{aligned}$$

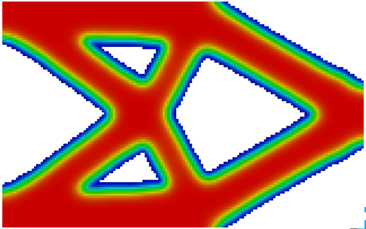
Effect of Penalization

$$\mathbf{K}^e \leftarrow (\boldsymbol{\mu}^e)^p \mathbf{K}^e$$

- $\mathbf{K}^e$: eth element stiffness matrix



(a) Without penalization

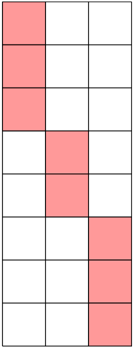


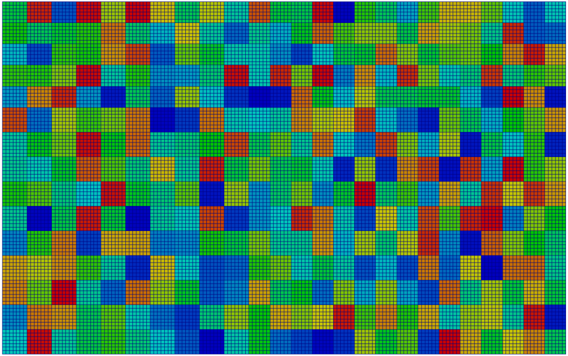
(b) With penalization

Standard Difficulty: Binary Solutions

Implication for ROM

- From parameter restriction, $\boldsymbol{\mu}^p = (\boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r)^p$
- Precomputation relies on separability of $\boldsymbol{\Phi}_{\boldsymbol{\mu}}$ and $\boldsymbol{\mu}_r$
- Separability maintained if $(\boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r)^p = \boldsymbol{\Phi}_{\boldsymbol{\mu}} \boldsymbol{\mu}_r^p$
- Sufficient condition: *columns of* $\boldsymbol{\Phi}_{\boldsymbol{\mu}}$ *have non-overlapping non-zeros*





Efficient Evaluation of Nonlinear Terms

- Due to the mixing of high-dimensional and low-dimensional terms in the ROM expression, only limited speedups available

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

- To enable *pre-computation* of all large-dimensional quantities into low-dimensional ones, leverage *Taylor series expansion*

$$\begin{aligned}[\mathbf{r}_r(\boldsymbol{u}_r,\ \boldsymbol{\mu}_r)]_i = \mathbf{D}^0_{im}(\boldsymbol{\mu}_r)_m + \mathbf{D}^1_{ijm}(\boldsymbol{u}_r \times \boldsymbol{\mu}_r)_{jm} + \mathbf{D}^2_{ijkm}(\boldsymbol{u}_r \times \boldsymbol{u}_r \times \boldsymbol{\mu}_r)_{jkm} \\ + \mathbf{D}^3_{ijklm}(\boldsymbol{u}_r \times \boldsymbol{u}_r \times \boldsymbol{u}_r \times \boldsymbol{\mu}_r)_{jklm} = 0\end{aligned}$$

where

$$\mathbf{D}^3_{ijklm} = \frac{\partial^3 \mathbf{r}_t}{\partial \boldsymbol{u}_p \partial \boldsymbol{u}_q \partial \boldsymbol{u}_s}(\hat{\boldsymbol{u}},\ \phi^m_{\boldsymbol{\mu}})(\phi^i_{\boldsymbol{u}} \times \phi^j_{\boldsymbol{u}} \times \phi^k_{\boldsymbol{u}} \times \phi^l_{\boldsymbol{u}})_{tpqs}$$

- Related work: [Rewienski, 2003, Barrault et al., 2004, Barbič and James, 2007, Nguyen and Peraire, 2008, Chaturantabut and Sorensen, 2010, Carlberg et al., 2011]

Lagrange Multiplier Estimate

Lagrange Multiplier, Constraint Pairs

$\boldsymbol{\lambda}$	$\boldsymbol{\lambda}_r$	$\boldsymbol{\tau}$	$\boldsymbol{\tau}_r$
$\mathbf{c}(\boldsymbol{u},\ \boldsymbol{\mu}) \geq 0$	$\mathbf{c}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}) \geq 0$	$\mathbf{A}\boldsymbol{\mu} \geq \mathbf{b}$	$\mathbf{A}_r\boldsymbol{\mu}_r \geq \mathbf{b}_r$

Goal: Given $\boldsymbol{u}_r,\ \boldsymbol{\mu}_r,\ \boldsymbol{\tau}_r \geq 0,\ \boldsymbol{\lambda}_r \geq 0$, estimate $\tilde{\boldsymbol{\tau}} \geq 0,\ \tilde{\boldsymbol{\lambda}} \geq 0$ to compute

$$\nabla_{\boldsymbol{\mu}}\mathcal{L}(\boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r,\ \tilde{\boldsymbol{\lambda}},\ \tilde{\boldsymbol{\tau}}) = \frac{\partial \mathcal{J}}{\partial \boldsymbol{\mu}}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) - \frac{\partial \mathbf{c}}{\partial \boldsymbol{\mu}}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r)^T\tilde{\boldsymbol{\lambda}} - \mathbf{A}^T\tilde{\boldsymbol{\tau}}$$

Lagrange Multiplier Estimates

$$\tilde{\boldsymbol{\lambda}} = \boldsymbol{\lambda}_r$$

$$\tilde{\boldsymbol{\tau}} = \underset{\boldsymbol{\tau}\geq 0}{\arg\min} \left|\left|\mathbf{A}^T\boldsymbol{\tau} - \left(\frac{\partial \mathcal{J}}{\partial \boldsymbol{\mu}}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r) - \frac{\partial \mathbf{c}}{\partial \boldsymbol{\mu}}(\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{u}_r,\ \boldsymbol{\Phi}_{\boldsymbol{\mu}}\boldsymbol{\mu}_r)^T\tilde{\boldsymbol{\lambda}}\right)\right|\right|$$

Non-negative least squares: [Lawson and Hanson, 1974, Chapman et al., 2015]

Standard Difficulty: Checkerboarding

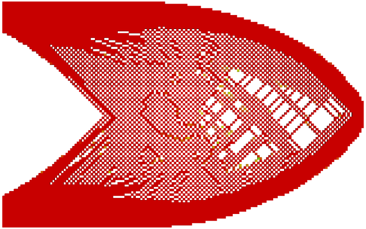
Gradient Filtering, Nodal Projection

- Minimum length scale, $r_{\min}$
- Gradient Filtering[11]

$$\widehat{\frac{\partial \mathcal{J}}{\partial \boldsymbol{\mu}_k}} = \frac{\sum_{j \in S_k} H_{kj} \boldsymbol{\mu}_i \frac{\partial \mathcal{J}}{\partial \boldsymbol{\mu}_i}}{\boldsymbol{\mu}_k \sum_{j \in S_k} H_{kj}}$$

- Nodal Projection

$$\boldsymbol{\mu}_k = \frac{\sum_{j \in \mathcal{S}_k} \boldsymbol{\tau}_j H_{jk}}{\sum_{j \in \mathcal{S}_k} H_{jk}}$$



(a) Without projection/filtering

[11] $H_{ki} = r_{\min} - \mathrm{dist}(k,\ i)$

Standard Difficulty: Checkerboarding

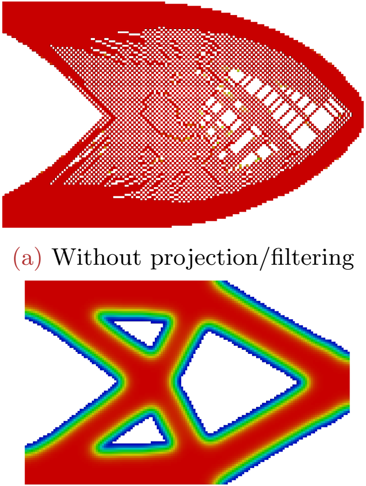
Gradient Filtering, Nodal Projection

- Minimum length scale, $r_{\min}$
- Gradient Filtering[11]

$$\widehat{\frac{\partial \mathcal{J}}{\partial \boldsymbol{\mu}_k}} = \frac{\sum_{j \in S_k} H_{kj} \boldsymbol{\mu}_i \frac{\partial \mathcal{J}}{\partial \boldsymbol{\mu}_i}}{\boldsymbol{\mu}_k \sum_{j \in S_k} H_{kj}}$$

- Nodal Projection

$$\boldsymbol{\mu}_k = \frac{\sum_{j \in \mathcal{S}_k} \boldsymbol{\tau}_j H_{jk}}{\sum_{j \in \mathcal{S}_k} H_{jk}}$$

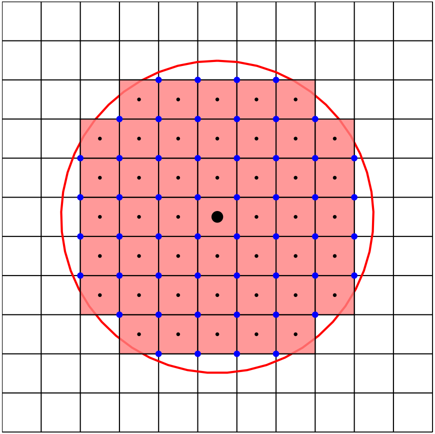


(a) Without projection/filtering

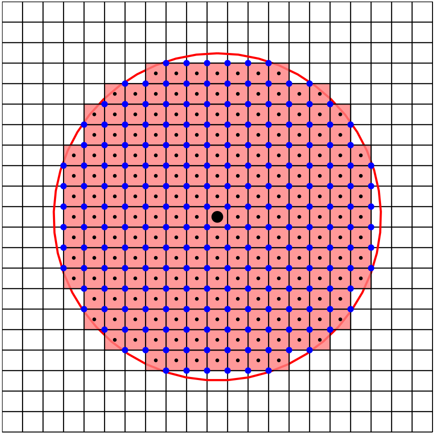
(b) With projection

[11] $H_{ki} = r_{\min} - \text{dist}(k,\ i)$

Standard Difficulty: Checkerboarding



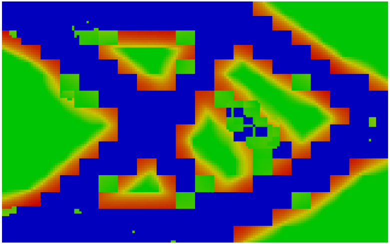
Standard Difficulty: Checkerboarding



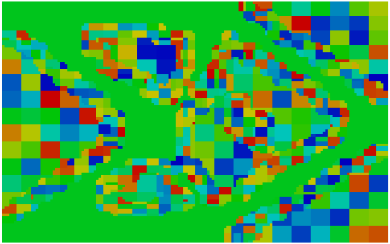
Standard Difficulty: Checkerboarding

Implication for ROM

- Nonlocality introduced through projection/filtering
- μ_e influences volume fraction of all elements within $r_{\min}$ of element/node e
- Clashes with requirement on Φ_μ of columns with non-overlapping non-zeros
- Handled heuristically by performing parameter basis adaptation to eliminate "checkerboard" regions of parameter space, uses concept of $r_{\min}$
- *Next: Helmholtz filtering*



Gradient of Lagrangian



Updated Macroelements

Standard Difficulty: Checkerboarding

Implication for ROM

- Nonlocality introduced through projection/filtering
- μ_e influences volume fraction of all elements within $r_{\min}$ of element/node e
- Clashes with requirement on Φ_μ of columns with non-overlapping non-zeros
- Handled heuristically by performing parameter basis adaptation to eliminate "checkerboard" regions of parameter space, uses concept of $r_{\min}$
- *Next: Helmholtz filtering*