# Reduced Order Modeling for Time-Dependent Optimal Control Problems with Variable Initial Values

**Matthias Heinkenschloss**

Department of Computational and Applied Mathematics
Rice University, Houston, Texas
heinken@rice.edu

August 12, 2017

London Mathematical Society EPSRC Durham Symposium
Model Order Reduction
Durham, UK, August 7 – 17, 2017

Joint work with
**Dörte Jando** (IWR Heidelberg)

Supported in part by DARPA EQUiPS, NSF DMS-1522798

# Motivation

## Linear-Quadratic Optimal Control Problem

$$\min_{\mathbf{u}\in L^2} \frac{1}{2}\int_0^T \|y(t)-y_d(t)\|^2_{L^2(\Omega_o)} dt + \frac{1}{2}\|y(T)-y_d(T)\|^2_{L^2(\Omega_o)} + \frac{\sigma}{2}\int_0^T \|\mathbf{u}(t)\|_2^2 dt,$$

where for given $\mathbf{u} : [0,T] \to I\!R^{n_u}$ the state $y \in W(0,T)$ solves

$$\begin{aligned}
\frac{\partial y(x,t)}{\partial t} - \kappa\Delta y(x,t) + \nu\cdot\nabla y(x,t) + \gamma y(x,t) &= \sum_{i=1}^{n_u} b_i(x)u_i(t), && x\in\Omega, t\in(0,T),\\
y(x,t) &= 0, && x\in\partial\Omega, t\in(0,T),\\
y(x,0) &= s(x), && x\in\Omega.
\end{aligned}$$

## After semi-discretization in space

$$\min_{\mathbf{u}\in L^2} \int_0^T \tfrac{1}{2}\mathbf{y}(t)^T\mathbf{Q}\mathbf{y}(t) + \mathbf{c}(t)^T\mathbf{y}(t) dt + \tfrac{1}{2}\mathbf{y}(T)^T\mathbf{Q}_T\mathbf{y}(T) + \mathbf{c}_T^T\mathbf{y}(T) + \int_0^T \tfrac{1}{2}\mathbf{u}(t)^T\mathbf{R}\mathbf{u}(t) dt,$$

where for given control $\mathbf{u} \in L^2((0,T); I\!R^{n_u})$ the state $\mathbf{y}$ solves

$$\mathbf{M}\frac{d}{dt}\mathbf{y}(t) + \mathbf{A}\mathbf{y}(t) = \mathbf{f}(t) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{M}\mathbf{y}(0) = \mathbf{M}\mathbf{s}.$$

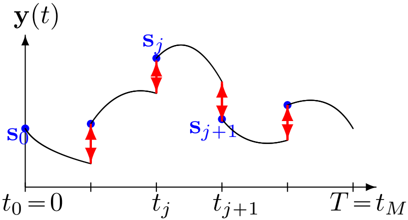
- Linear-quadratic optimal control problem (although with time dependent $\mathbf{A}, ...$) arises as subproblem in Newton-SQP-type methods for nonlinear optimal control problems.
- When $\mathbf{A}$ has negative eigenvalues ($\gamma < 0$ in model PDE), solution of PDE becomes unstable. Use multiple shooting reformulation.

## Multiple Shooting

Split time interval $[0, T]$ into $M$ subintervals $[t_j, t_{j+1}]$, $j = 0, \dots, M-1$.

Introduce auxiliary initial values $\mathbf{s}_j$; solve differential eqn. on each subinterval.

Add continuity conditions $\mathbf{M}\mathbf{y}(t_{j+1}; \mathbf{s}_j, \mathbf{u}) - \mathbf{M}\mathbf{s}_{j+1} = \mathbf{0}$ as constraints.




## Multiple Shooting Formulation

$$\min_{\mathbf{u}_j \in L^2, \mathbf{s}_j \in I\!R^n} \sum_{j=0}^{M-1} \int_{t_j}^{t_{j+1}} \frac{1}{2}\mathbf{y}_j(t)^T \mathbf{Q}\mathbf{y}_j(t) + \mathbf{c}(t)^T \mathbf{y}_j(t) dt + \int_0^T \frac{1}{2}\mathbf{u}_j(t)^T \mathbf{R}\mathbf{u}_j(t) dt,$$

$$+ \frac{1}{2}\mathbf{y}_{M-1}(T)^T \mathbf{Q}_T \mathbf{y}_{M-1}(T) + \mathbf{c}_T^T \mathbf{y}_{M-1}(T),$$

$$\text{s.t.} \quad \mathbf{M}\mathbf{y}_j(t_{j+1}; \mathbf{s}_j, \mathbf{u}) - \mathbf{M}\mathbf{s}_{j+1} = \mathbf{0},$$

where for given control $\mathbf{u}_j \in L^2((0,T); I\!R^{n_u}$ and $\mathbf{s}_j \in I\!R^n$ the state $\mathbf{y}_j$ solves

$$\mathbf{M}\frac{d}{dt}\mathbf{y}_j(t) + \mathbf{A}\mathbf{y}_j(t) = \mathbf{f}(t) + \mathbf{B}\mathbf{u}_j(t), \; t \in (t_j, t_{j+1}), \quad \mathbf{M}\mathbf{y}(t_j) = \mathbf{M}\mathbf{s}_j.$$

## Multiple Shooting Formulation

$$\min_{\mathbf{u}_j \in L^2, \mathbf{s}_j \in I\!R^n} \sum_{j=0}^{M-1} \int_{t_j}^{t_{j+1}} \tfrac{1}{2}\mathbf{y}_j(t)^T \mathbf{Q} \mathbf{y}_j(t) + \mathbf{c}(t)^T \mathbf{y}_j(t) dt + \int_0^T \tfrac{1}{2}\mathbf{u}_j(t)^T \mathbf{R} \mathbf{u}_j(t) dt,$$

$$+ \tfrac{1}{2}\mathbf{y}_{M-1}(T)^T \mathbf{Q}_T \mathbf{y}_{M-1}(T) + \mathbf{c}_T^T \mathbf{y}_{M-1}(T),$$

$$\text{s.t.} \quad \mathbf{M}\mathbf{y}_j(t_{j+1}; \mathbf{s}_j, \mathbf{u}) - \mathbf{M}\mathbf{s}_{j+1} = \mathbf{0},$$

where for given control $\mathbf{u}_j \in L^2((0,T); I\!R^{n_u}$ and $\mathbf{s}_j \in I\!R^n$ the state $\mathbf{y}_j$ solves

$$\mathbf{M}\frac{d}{dt}\mathbf{y}_j(t) + \mathbf{A}\mathbf{y}_j(t) = \mathbf{f}(t) + \mathbf{B}\mathbf{u}_j(t),\ t \in (t_j, t_{j+1}), \quad \mathbf{M}\mathbf{y}(t_j) = \mathbf{M}\mathbf{s}_j.$$

- Multiple shooting formulation can alleviate the instability, but at the price of introducing additional optimization variables $[\mathbf{s}_0, \ldots, \mathbf{s}_{M-1}] \in I\!R^{M \cdot n}$, $n \gg 1$.
- For PDE constrained problems this creates challenges for numerical solution.
- Goal: Faster solver by Reduced Order Modeling for Hessian computation

# ROM and Optimization

- Finding ROMs that well approximate the optimization problem over a range of optimization parameters can be expensive.
- Recovering from poor ROMs in the optimization can be expensive.
- Often difficult to amortize cost of ROM generation in the optimization. In his case better to use inexpensive, rough ROMs to solve subproblems, rather that to approximate original optimization problem.

## In this talk

- There is no (small) ROMs that well approximates the optimization problem (number of inputs and outputs is large).
- Therefore, we use full order model to compute objective function and gradient.
- Use ROM to generate efficient Hessian approximation for fast Newton-type algorithms.

# Outline

ROM for Initial Value Control Problems

Sequential ROMs for Initial Value Control Problems

ROMs for Problems with Initial Value and Right Hand Side Controls

# Optimal Control Problems with Initial Value Controls

$$\min_{\mathbf{s}\in \mathbb{R}^n} J(\mathbf{s}) = \int_0^T \frac{1}{2}\mathbf{y}(t)^T\mathbf{Q}\mathbf{y}(t)+\mathbf{c}(t)^T\mathbf{y}(t)dt+\tfrac{1}{2}\mathbf{y}(T)^T\mathbf{Q}_T\mathbf{y}(T)+\mathbf{c}_T^T\mathbf{y}(T)+\frac{\beta}{2}\mathbf{s}^T\mathbf{M}\mathbf{s},$$

where for given $\mathbf{s}$ the state $\mathbf{y}$ solves

$$\mathbf{M}\frac{d}{dt}\mathbf{y}(t) + \mathbf{A}\mathbf{y}(t) = \mathbf{f}(t), \quad \mathbf{M}\mathbf{y}(0) = \mathbf{M}\mathbf{s}.$$

- $\mathbf{Q}, \mathbf{Q}_T$ symmetric, pos semidefinite; $\mathbf{M}$ symmetric, pos. definite, $\beta > 0$,
- $\mathbf{A}$ possibly non-symmetric

Such problems arise

- as subproblem on multiple shooting,
- in source inversion, e.g., [Bashir et al., 2008],
- as subproblems in data assimilation, e.g., [Blum et al.,2009], [Rao&Sandu, 2016], [talk by Nancy Nichols].

ROM for variable initial conditions [H.,Reis,Antoulas 2011], [Beattie,Gugercin,Mehrmann2016], but require initial value in small dim. subspace.

# Optimality Condition and Hessian Computation

- Weighted inner product $\langle \mathbf{s}_1, \mathbf{s}_2 \rangle = \mathbf{s}_1^T \mathbf{M} \mathbf{s}_2$ .
- Gradient: Solve

$$\mathbf{M}\frac{d}{dt}\mathbf{y}(t) + \mathbf{A}\mathbf{y}(t) = \mathbf{f}(t), \qquad \mathbf{M}\mathbf{y}(0) = \mathbf{M}\mathbf{s},$$
$$-\mathbf{M}\frac{d}{dt}\mathbf{p}(t) + \mathbf{A}^T\mathbf{p}(t) = \mathbf{Q}\,\mathbf{y}(t) + \mathbf{c}(t), \qquad \mathbf{M}\mathbf{p}(T) = \mathbf{Q}_T(\mathbf{y}(T) + \mathbf{c}_T),$$

$\nabla J(\mathbf{s}) = \mathbf{p}(0) + \beta\mathbf{s}$.

- Hessian-vector-product: Solve

$$\mathbf{M}\frac{d}{dt}\mathbf{z}(t) + \mathbf{A}\mathbf{z}(t) = \mathbf{0}, \qquad \mathbf{M}\mathbf{z}(0) = \mathbf{M}\mathbf{s},$$
$$-\mathbf{M}\frac{d}{dt}\mathbf{q}(t) + \mathbf{A}^T\mathbf{q}(t) = \mathbf{Q}\,\mathbf{z}(t), \qquad \mathbf{M}\mathbf{p}(T) = \mathbf{Q}_T\mathbf{z}(T),$$

$\nabla^2 J\ \mathbf{s} = \mathbf{H}\mathbf{s} + \beta\mathbf{s}$, where $\mathbf{H}\,\mathbf{s} = \mathbf{q}(0)$.

- First order optimality condition

$$\mathbf{0} = \nabla J(\mathbf{s}_*) = \nabla^2 J\,\mathbf{s}_* + \nabla J(\mathbf{0}) = \mathbf{H}\mathbf{s}_* + \beta\mathbf{s}_* + \nabla J(\mathbf{0}).$$

# Petrov-Galerkin Projection based ROM

- Want to approximate Hessian in $(\mathbf{H} + \beta \mathbf{I})\mathbf{s}_* = -\nabla J(\mathbf{0})$, by projection based reduced order model (ROM)
- For given $\mathbf{V}, \mathbf{W} \in I\!R^{n \times k}$, $k \ll n$, with $\mathbf{W}^T\mathbf{M}\mathbf{V}$ invertible (often $= \mathbf{I}_k$ )

  $\mathbf{z} \approx \mathbf{V}\widehat{\mathbf{z}}$ and $\mathbf{q} \approx \mathbf{W}\widehat{\mathbf{q}}$ where $\widehat{\mathbf{z}}$, $\widehat{\mathbf{q}}$ solve

$$\begin{aligned} \mathbf{W}^T\mathbf{M}\mathbf{V}\frac{d}{dt}\widehat{\mathbf{z}}(t) + \mathbf{W}^T\mathbf{A}\mathbf{V}\,\widehat{\mathbf{z}}(t) &= \mathbf{0}, \\ \mathbf{W}^T\mathbf{M}\mathbf{V}\widehat{\mathbf{z}}(0) &= \mathbf{W}^T\mathbf{M}\mathbf{s} \end{aligned}$$

  and

$$\begin{aligned} -\mathbf{V}^T\mathbf{M}\mathbf{W}\frac{d}{dt}\widehat{\mathbf{q}}(t) + \mathbf{V}^T\mathbf{A}^T\mathbf{W}\,\widehat{\mathbf{q}}(t) &= \mathbf{V}^T\mathbf{Q}\,\mathbf{V}\,\widehat{\mathbf{z}}(t), \\ \mathbf{V}^T\mathbf{M}\mathbf{W}\widehat{\mathbf{q}}(T) &= \mathbf{V}^T\mathbf{Q}_T\mathbf{V}\,\widehat{\mathbf{z}}(T). \end{aligned}$$

  ROM Hessian approximation

$$\widehat{\mathbf{H}}\mathbf{s} = \mathbf{W}\widehat{\mathbf{q}}(0).$$

- Solve $\widehat{\mathbf{H}}\mathbf{s} + \beta\mathbf{s} = -\nabla J(\mathbf{0})$.
- Difficulty: $\mathbf{s}$ varies in $I\!R^n$. Can't find ROM that is good for all $\mathbf{s} \in I\!R^n$.

# Important Special Case to Illustrate Issues

- $\mathbf{A}$ is symmetric (i.e. no advection) $\quad \rightarrow \mathbf{W} = \mathbf{V}$
- $\mathbf{Q} = \mathbf{Q}_T = \mathbf{M}$
- Generalized eigenvalue decomposition of $(\mathbf{A}, \mathbf{M})$:

  $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^{n \times n}$, $\lambda_1 \leqslant \ldots \leqslant \lambda_n$

$$\mathbf{V}_n^T \mathbf{M} \mathbf{V}_n = \mathbf{I}, \quad \mathbf{A} \mathbf{V}_n = \mathbf{M} \mathbf{V}_n \boldsymbol{\Lambda}.$$

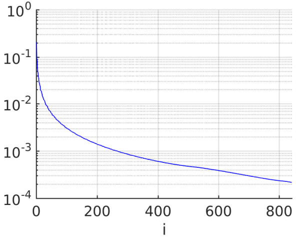
After transformation of variables $\quad \mathbf{y}(t) = \mathbf{V}_n \widetilde{\mathbf{y}}(t)$
OCP decouples into $n$ real scalar OCPs with initial data $\widetilde{\mathbf{s}} := \mathbf{V}_n^T \mathbf{M} \mathbf{s}$.

Hessian is diagonal, $\nabla^2 \widetilde{J} = \widetilde{\mathbf{H}} + \beta \mathbf{I}$, where

$$\widetilde{\mathbf{H}} = \mathrm{diag}(\widetilde{\mathbf{H}}_{11}, \ldots, \widetilde{\mathbf{H}}_{nn}),$$

and

$$\widetilde{\mathbf{H}}_{11} \geqslant \widetilde{\mathbf{H}}_{22} \geqslant \ldots \widetilde{\mathbf{H}}_{nn} > 0.$$



# Basic ROM = best rank $k$ approximation

- Approximate $\widetilde{\mathbf{H}} = \text{diag}(\widetilde{\mathbf{H}}_{11}, \ldots, \widetilde{\mathbf{H}}_{nn})$ by best rank $k$-approximation

$$\widetilde{\mathbf{H}}^{\text{bsc}} = \text{diag}(\widetilde{\mathbf{H}}_{11}, \ldots, \widetilde{\mathbf{H}}_{kk}, 0, \ldots, 0).$$

  - This is basic ROM
  - $\widetilde{\mathbf{H}}^{\text{bsc}}\, \widetilde{\mathbf{s}}$ obtained by projection with $\widetilde{\mathbf{V}} = \widetilde{\mathbf{W}} = [\mathbf{e}_1, \ldots, \mathbf{e}_k] \in \mathbb{R}^{n \times k}$,
  - Only two differential equations of size $k$ need to be solved.
  - Need to compute $k$ smallest generalized eigenvalues $(\mathbf{A}, \mathbf{M})$.

- $\widetilde{\mathbf{H}}\widetilde{\mathbf{s}} + \beta\widetilde{\mathbf{s}} = -\widetilde{\mathbf{g}} \stackrel{\text{def}}{=} -\nabla\widetilde{J}(\mathbf{0})$ is approximated by

$$\begin{aligned} \widetilde{\mathbf{H}}_{ii}\widetilde{\mathbf{s}}_i + \beta\widetilde{\mathbf{s}}_i &= -\widetilde{\mathbf{g}}_i, \quad i = 1, \ldots, k && \text{exact} \\ \beta\widetilde{\mathbf{s}}_i &= -\widetilde{\mathbf{g}}_i, \quad i = k+1, \ldots, n && \text{potentially large error} \end{aligned}$$

- Error large if $\widetilde{\mathbf{H}}_{ii} > \beta$ for some $i \geqslant k+1$, and there are relatively large components $\widetilde{\mathbf{g}}_{k+1}, \ldots, \widetilde{\mathbf{g}}_n$

# Augmentation of Basic ROM by right hand side $\widetilde{\mathbf{g}}$

Add rhs $\widetilde{\mathbf{g}}$ to reduced basis $\widetilde{\mathbf{V}} = \widetilde{\mathbf{W}} = [\mathbf{e}_1, \ldots, \mathbf{e}_k]$ to generate $\hat{\widetilde{\mathbf{V}}} = \widehat{\widetilde{\mathbf{W}}}$ with

$$\text{range}(\hat{\widetilde{\mathbf{V}}}) = \text{range}([\widetilde{\mathbf{V}}, \widetilde{\mathbf{g}}]), \quad \hat{\widetilde{\mathbf{V}}}^T \hat{\widetilde{\mathbf{V}}} = \mathbf{I}.$$

If $\widetilde{\mathbf{g}} \notin \widetilde{\mathbf{V}}$, then new basis vector is

$$\widetilde{\mathbf{v}} = \Big(0, \ldots, 0, \widetilde{\mathbf{g}}_{k+1}, \ldots, \widetilde{\mathbf{g}}_n\Big)^T \Big/ \Big(\sum_{i=k+1}^{n} \widetilde{\mathbf{g}}_i^2\Big)^{1/2}$$

Hessian approximation

$$\widetilde{\mathbf{H}}^{\text{aug}}\, \widetilde{\mathbf{s}} = \sum_{i=1}^{k} \widetilde{\mathbf{H}}_{ii}\, \widetilde{\mathbf{s}}_i\, \mathbf{e}_i + \widetilde{\boldsymbol{\rho}}(0)\, \widetilde{\mathbf{v}}$$

required solution of two additional scalar ODEs

$$\begin{aligned}
\frac{d}{dt}\widetilde{\boldsymbol{\zeta}}(t) + \lambda(\widetilde{\mathbf{g}})\, \widetilde{\boldsymbol{\zeta}}(t) &= 0, & \widetilde{\boldsymbol{\zeta}}(0) &= \widetilde{\mathbf{v}}^T \widetilde{\mathbf{s}}, \\
-\frac{d}{dt}\widetilde{\boldsymbol{\rho}}(t) + \lambda(\widetilde{\mathbf{g}})\, \widetilde{\boldsymbol{\rho}}(t) &= \widetilde{\boldsymbol{\zeta}}(t), & \widetilde{\boldsymbol{\rho}}(T) &= \widetilde{\boldsymbol{\zeta}}(T).
\end{aligned}$$

# Error Estimate

$$(\widetilde{\mathbf{H}} + \beta\mathbf{I})\widetilde{\mathbf{s}} = -\widetilde{\mathbf{g}} \text{ with full Hessian}$$

$$(\widetilde{\mathbf{H}}^{\rm bsc} + \beta\mathbf{I})\widetilde{\mathbf{s}}^{\rm bsc} = -\widetilde{\mathbf{g}} \text{ with basic ROM Hessian}$$

$$(\widetilde{\mathbf{H}}^{\rm aug} + \beta\mathbf{I})\widetilde{\mathbf{s}}^{\rm aug} = -\widetilde{\mathbf{g}} \text{ with augm. ROM Hessian}$$

Error estimate

$$\|\widetilde{\mathbf{s}} - \widetilde{\mathbf{s}}^{\rm aug}\|^2 \leqslant \|\widetilde{\mathbf{s}} - \widetilde{\mathbf{s}}^{\rm bsc}\|^2 - \left(\frac{h(\lambda(\widetilde{\mathbf{g}}))}{\beta\ (h(\lambda_{k+1}) + \beta)}\right)^2 \sum_{i=k+1}^{n} \widetilde{\mathbf{g}}_i^2 < \|\widetilde{\mathbf{s}} - \widetilde{\mathbf{s}}^{\rm bsc}\|^2.$$

where

$$h(\lambda) = e^{-2\lambda T} + \left(1 - e^{-2\lambda T}\right)/(2\lambda), \qquad (\geqslant 0, \text{ mon } \searrow, \text{ convex})$$

$$\widetilde{\mathbf{H}}_{ii} = h(\lambda_i),$$

$$\lambda(\widetilde{\mathbf{g}}) = \sum_{i=k+1}^{n} \lambda_i \widetilde{\mathbf{g}}_i^2 \Big/ \sum_{i=k+1}^{n} \widetilde{\mathbf{g}}_i^2. \qquad \text{(weighted arithmetic mean of } \lambda_{k+1}, \ldots)$$

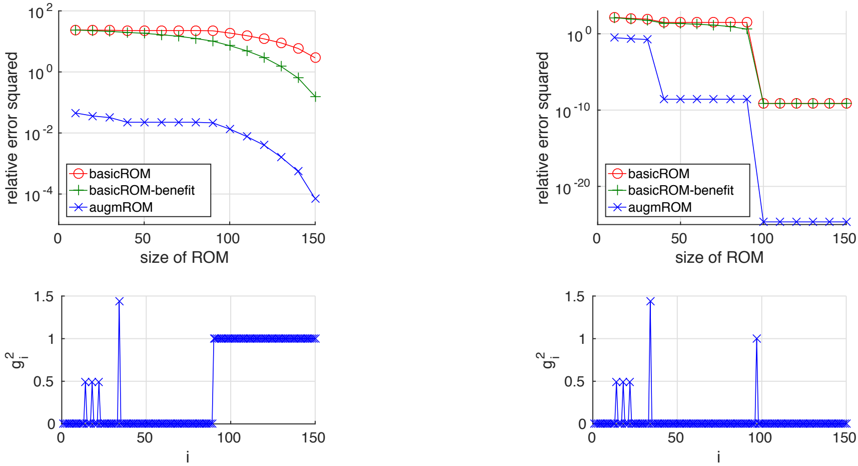
Roughly speaking, augmentation works the better the fewer important right hand side components $\widetilde{\mathbf{g}}_{k+1}, \ldots, \widetilde{\mathbf{g}}_n$ exist

# Illustration

FEM disc. of 2D model problem with $\kappa = 0.1$, $\nu = (0,0)^T$, $\gamma = 0.5$, $\beta = 5 \cdot 10^{-4}$.

$$\widetilde{g}_{14} = \widetilde{g}_{18} = \widetilde{g}_{22} = .7,\ \widetilde{g}_{34} = 1.2,\ \widetilde{g}_{840} = \widetilde{g}_{841} = 10^{-5}$$

$\widetilde{g}_{90} = \cdots = \widetilde{g}_{160} = 1$ $\qquad\qquad$ $\widetilde{g}_{97} = 1$



# General ROM Augmentation & Use of ROM Hessians

- In practice, apply ROM augmentation to original OCP.
- $\mathbf{A}$ symmetric:

  Compute $k$ smallest generalized eigenvalues $\lambda_1 \leqslant \ldots \leqslant \lambda_k$ of $(\mathbf{A}, \mathbf{M})$ and matrix of corresponding eigenvectors $\mathbf{V} \in I\!R^{n \times k}$, $\mathbf{V}^T \mathbf{M} \mathbf{V} = \mathbf{I}_{k \times k}$, $\mathbf{A}\mathbf{V} = \mathbf{M}\mathbf{V}\text{diag}(\lambda_1, \ldots, \lambda_k)$.

  Augment: $\text{range}(\widehat{\mathbf{V}}) = \text{range}([\mathbf{V}, \mathbf{g}])$, $\widehat{\mathbf{V}}^T \mathbf{M} \widehat{\mathbf{V}} = \mathbf{I}$.
- $\mathbf{A}$ non-symmetric:

  Compute $k$ generalized eigenvalues of $(\mathbf{A}, \mathbf{M})$ with smallest real part $\text{Re}(\lambda_1) \leqslant \ldots \leqslant \text{Re}(\lambda_k)$ and corresponding left and right eigenvectors $\mathbf{V}, \mathbf{W}$ (each complex vector gives two real vectors).

  Augment: $\text{range}(\widehat{\mathbf{V}}) = \text{range}([\mathbf{V}, \mathbf{g}])$, $\text{range}(\widehat{\mathbf{W}}) = \text{range}([\mathbf{W}, \mathbf{g}])$, $\widehat{\mathbf{W}}^T \mathbf{M} \widehat{\mathbf{V}} = \mathbf{I}$.
- Use conjugate gradient method to solve

$$(\mathbf{H} + \beta \mathbf{I})\mathbf{s} = -\mathbf{g} \quad \text{where } \mathbf{g} = \nabla J(\mathbf{0})$$

# CG with Fixed ROM Hessian Approximation

**Algorithm 1:**

1. Generate reduced order model $\mathbf{W}, \mathbf{V} \in \mathbb{R}^{n \times k}$.
   Compute $(\widehat{\mathbf{H}} + \beta \mathbf{I})\mathbf{s}_0$ and set $\mathbf{r}_0 = -\mathbf{g} - (\widehat{\mathbf{H}} + \beta \mathbf{I})\mathbf{s}_0$, $\mathbf{d}_0 = \mathbf{r}_0$.
2. For $i = 0, 1, 2, ..., i_{\max}$
   2.1 If $\|\mathbf{r}_i\| \leqslant tol^{\text{cg}}$, return $\mathbf{s}_i$.
   2.2 Compute $\mathbf{h}_i = (\widehat{\mathbf{H}} + \beta \mathbf{I})\mathbf{d}_i$.
   2.3 $\alpha_i = \langle \mathbf{r}_i, \mathbf{r}_i \rangle / \langle \mathbf{d}_i, \mathbf{h}_i \rangle$
   2.4 $\mathbf{s}_{i+1} = \mathbf{s}_i + \alpha_i \mathbf{d}_i$
   2.5 $\mathbf{r}_{i+1} = \mathbf{r}_i - \alpha_i \mathbf{h}_i$
   2.6 $\mathbf{d}_{i+1} = \mathbf{r}_{i+1} + \frac{\langle \mathbf{r}_{i+1}, \mathbf{r}_{i+1} \rangle}{\langle \mathbf{r}_i, \mathbf{r}_i \rangle} \mathbf{d}_i$

If eigenvalues $\widehat{\mu}_1 \geqslant \ldots \geqslant \widehat{\mu}_k \geqslant \widehat{\mu}_{k+1} = \ldots = 0$ of $\widehat{\mathbf{H}}$, then Algorithm 1 converges in $k+1$ iterations.

If $\mathbf{r}_0 \in R(\mathbf{W})$, then $\mathbf{d}_i \in R(\mathbf{W})$ for all $i$.

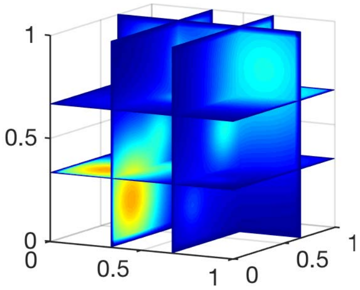
# Numerical Results

Model problem with $\beta = 10^{-4}$, $\gamma = 0.5$, $\Omega_o = \Omega = (0,1)^3$ discretized by P1 elements on a $30 \times 30 \times 30$ spatial grid
$\rightarrow n = 29,791$, $\mathbf{Q} = \mathbf{Q}_T = \mathbf{M}$, $\quad \mathbf{R} = \beta\mathbf{M}$

Desired state computed using initial state

$$\begin{aligned} s(x) =& 2e^{-10\|x-x_1\|_2^2} + e^{-5\|x-x_2\|_2^2} \\ & + 2e^{-50\|x-x_3\|_2^2} + 2e^{-40\|x-x_4\|_2^2} \\ & + e^{-100\|x-x_5\|_2^2}, \end{aligned}$$

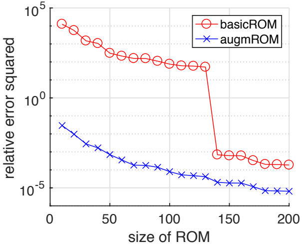
where $x_1 = (0.2, 0.2, 0.2)$,
$x_2 = (0.8, 0.8, 0.8)$, $x_3 = (0.5, 0.5, 0.5)$,
$x_4 = (0.2, 0.8, 0.8)$, $x_5 = (0.8, 0.2, 0.2)$.



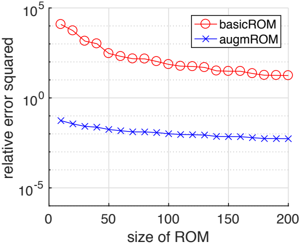

Implicit Euler method in time with $50$ steps.

# Model Problem with Symmetric $\mathbf{A}$ ($\kappa = 0.1$, $\nu = (0,0)^T$)

Solve linear system using CG with ROM $k = 20$ Hessians and two CG tolerances.




(a) $tol^{\mathrm{cg}} = 10^{-2}\|\mathbf{g}\|_{\mathbf{M}}$




(b) $tol^{\mathrm{cg}} = 10^{-4}\|\mathbf{g}\|_{\mathbf{M}}$

- Augmented ROM Hessians produce better results
- Coarse CG tolerance can negatively impact improvement achieved by ROM augmentation

# Model Problem with Symmetric $\mathbf{A}$

Solve model problem with ROM size $k = 20$ and $tol^{cg} = 10^{-4}\|\mathbf{g}\|_{\mathbf{M}}$.

| | $J$ | $\|\nabla J\|$ | CG iters | ROM size | PDE solves full / ROM | rel err |
|---|---|---|---|---|---|---|
| at $\mathbf{s}_0 = \mathbf{0}$ | $1.46e-2$ | $5.91e-2$ | | | | |
| Full Hessian | $1.75e-5$ | $5.61e-6$ | 24 | | 50 / 0 | |
| $\mathbf{H}^{\text{bsc}}$ | $1.82e+1$ | $8.47e-1$ | 11 | 20 | 2 / 22 | $5.44e+3$ |
| $\mathbf{H}^{\text{aug}}$, $\mathbf{a} = \mathbf{g}$ | $5.27e-5$ | $8.68e-4$ | 9 | 21 | 2 / 18 | $3.59e-2$ |
| $\mathbf{H}^{\text{aug}}$, $\mathbf{a} = \mathbf{d}_i$ | $5.27e-5$ | $8.68e-4$ | 9 | 21 | 2 / 18 | $3.59e-2$ |

- Solution by augmented ROM is much better than by basic ROM (failed)
- Computation using augmented ROM much cheaper than full Hessian

## Experiments with Nonsymmetric $\mathbf{A}$ ($\nu \neq \mathbf{0}$)

- Augmented ROM $\widehat{\mathbf{W}}, \widehat{\mathbf{V}}$ gives better results than basic ROM
- Benefit of ROM augmentation increases for more advection-dominated prob.

# Outline

ROM for Initial Value Control Problems

Sequential ROMs for Initial Value Control Problems

ROMs for Problems with Initial Value and Right Hand Side Controls

# Sequential Reduced Order Modeling

ROM augmentation provides substantial improvement over basic ROM, but resulting $\mathbf{s}$ approximation may still not be sufficiently accurate

$\rightarrow$ Apply ROM approach sequentially

- Each iteration uses a small sized ROM that works on a different subspace
- Isolate subspaces by projections

For an approximation $\mathbf{s}^{(\ell)}$ of the OCP solution $\mathbf{s}_*$,
ROM results in Hessian approximation $\widehat{\mathbf{H}}^{(\ell)}$,
correction $\Delta\mathbf{s}$ solves

$$(\widehat{\mathbf{H}}^{(\ell)} + \mathbf{M}^{-1}\mathbf{R})\Delta\mathbf{s} = -\nabla J(\mathbf{s}^{(\ell)})$$

and yields new approximation $\mathbf{s}^{(\ell+1)} = \mathbf{s}^{(\ell)} + \Delta\mathbf{s}$.

# Sequential ROM for Decoupled Problem

Consider diagonalizable case:

- $\mathbf{A}$ is symmetric (i.e. no advection) $\quad \rightarrow \mathbf{W} = \mathbf{V}$
- $\mathbf{Q} = \mathbf{Q}_T = \mathbf{M}, \quad \mathbf{R} = \beta \mathbf{M}, \quad \beta > 0$
- Generalized eigenvectors $\mathbf{V}_n$ of $(\mathbf{A}, \mathbf{M})$

$\rightarrow$ OCP decouples in $n$ scalar OCPs with initial data $\widetilde{\mathbf{s}} := \mathbf{V}_n^T \mathbf{M} \mathbf{s}$ (by $\mathbf{y} = \mathbf{V}_n \widetilde{\mathbf{y}}$).

$$
\begin{aligned}
(\widehat{\widetilde{\mathbf{H}}} + \beta \mathbf{I}) \Delta \widetilde{\mathbf{s}} &= -\nabla \widetilde{J}(\widetilde{\mathbf{s}}^{(\ell)}) \\
\widetilde{\mathbf{s}}^{(\ell+1)} &= \widetilde{\mathbf{s}}^{(\ell)} + \Delta \widetilde{\mathbf{s}}
\end{aligned}
$$

- After initial ($\ell = 0$) solution with basic ROM:
  first $k$ components of $\widetilde{\mathbf{s}}^{(1)}$ are exact, first $k$ components of gradient are zero.
- In step $\ell$: Repeat procedure on subsystem with indices $\ell k + 1$ to $n$.
- Use projection $\mathbf{P}$ to essentially remove first $\ell k$ equations and unknowns (needed to compute augmentation).
  Use projection $\mathbf{Q}$ to essentially solve equations $\ell k + 1, \ldots, (\ell + 1)k$.

# Optimization with Sequential ROM Hessians

**Algorithm 2:**

0. Given $tol \in (0,1)$, $\mathbf{s}^{(0)}$, $k$, $\mathbf{V}^{\text{old}} = []$, $\mathbf{P} = \mathbf{I}$, $\mathbf{Q} = \mathbf{I}$.
1. Generate $\mathbf{V}_n$, i.e. eigenvectors of $(\mathbf{A}, \mathbf{M})$.
2. For $\ell = 0, 1, 2, \ldots, \ell_{max}$
   - 2.1 If $\|\nabla J(\mathbf{s}^{(\ell)})\|_{\mathbf{M}} < tol$, return $\mathbf{s}^{(\ell)}$.
   - 2.2 Choose ROM $\mathbf{V}^c$:
     - 2.2.1 Either basic ROM $\mathbf{V}^c = [(\mathbf{V}_n)_{\ell k+1}, \cdots, (\mathbf{V}_n)_{(\ell+1)k}] \in \mathbb{R}^{n \times k}$
     - 2.2.2 or this basic ROM augmented by $\mathbf{P}^* \nabla J(\mathbf{s}^{(\ell)})$ which gives

$$\mathbf{V}^c = [(\mathbf{V}_n)_{\ell k+1}, \cdots, (\mathbf{V}_n)_{(\ell+1)k}, \mathbf{v}] \in \mathbb{R}^{n \times (k+1)}$$

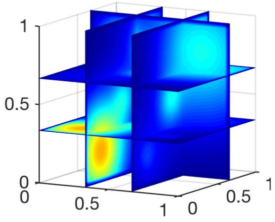
   - 2.3 Compute orthogonal projection on subspace $\mathbf{V}^c$: $\mathbf{Q} = \mathbf{V}^c (\mathbf{V}^c)^T \mathbf{M}$.
   - 2.4 For the Hessian approximation $\widehat{\mathbf{H}}^c$ obtained by $\mathbf{V}^c$ solve

$$\mathbf{Q}^* (\widehat{\mathbf{H}}^c + \mathbf{M}^{-1}\mathbf{R}) \mathbf{Q} \Delta \mathbf{s} = -\mathbf{Q}^* \nabla J(\mathbf{s}^{(\ell)})$$

   - 2.5 Update $\mathbf{s}^{(\ell+1)} = \mathbf{s}^{(\ell)} + \Delta \mathbf{s}$
   - 2.6 Set $\mathbf{V}^{\text{old}} = [\mathbf{V}^{\text{old}}\ \mathbf{V}_1^c \ldots \mathbf{V}_k^c] \in \mathbb{R}^{n \times (\ell+1)k}$
   - 2.7 Compute orthogonal projection on $(\mathbf{V}^{\text{old}})^{\perp}$: $\mathbf{P} = \mathbf{I} - \mathbf{V}^{\text{old}} (\mathbf{V}^{\text{old}})^T \mathbf{M}$.

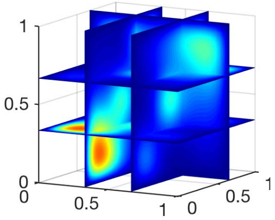
| | iter | $J$ | $\|\nabla J\|$ | CG iters | full / ROM solves |
|---|---|---|---|---|---|
| full Hessian | 0 | 1.4578e-02 | 5.9128e-02 | | 2 / 0 |
| (size $n = 29,791$) | 1 | 1.7529e-05 | 5.6103e-06 | 24 | 48 / 0 |
| basic ROM | 0 | 1.4578e-02 | 5.9128e-02 | | 2 / 0 |
| (size $20$) | 1 | 5.5269e-04 | 4.3797e-03 | 11 | 2 / 22 |
| | 2 | 1.7462e-04 | 1.9109e-03 | 7 | 2 / 14 |
| | 3 | 6.9784e-05 | 8.5360e-04 | 5 | 2 / 10 |
| | 4 | 6.0665e-05 | 7.2887e-04 | 5 | 2 / 10 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 14 | 2.3966e-05 | 1.7130e-04 | 3 | 2 / 6 |
| | 15 | 2.1713e-05 | 1.2443e-04 | 4 | 2 / 8 |
| augmented ROM | 0 | 1.4578e-02 | 5.9128e-02 | | 2 / 0 |
| (size $21$) | 1 | 5.2692e-05 | 8.6795e-04 | 12 | 2 / 24 |
| | 2 | 2.4754e-05 | 3.3800e-04 | 7 | 2 / 14 |
| | 3 | 1.9291e-05 | 1.5485e-04 | 6 | 2 / 12 |
| | 4 | 1.8066e-05 | 6.7495e-05 | 5 | 2 / 10 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 8 | 1.7542e-05 | 6.0700e-06 | 4 | 2 / 8 |
| | 9 | 1.7527e-05 | 2.4308e-06 | 4 | 2 / 8 |

# Numerical Results

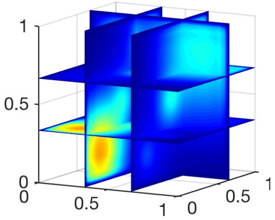



exact

50 full PDE solves




one step of augmented ROM
4 full PDE solves




sequential application of augmented ROM
20 full PDE solves

# Outline

ROM for Initial Value Control Problems

Sequential ROMs for Initial Value Control Problems

ROMs for Problems with Initial Value and Right Hand Side Controls

# Problem with Controls

Recall: $n_u$ is small $\rightarrow$ No model order reduction wrt $\mathbf{u}$
$n$ is large $\rightarrow$ Model order reduction wrt $\mathbf{y}$

$$\min_{\mathbf{s},\mathbf{u}} J(\mathbf{s},\mathbf{u}) = \int_0^T \tfrac{1}{2}\mathbf{y}(t)^T\mathbf{Q}\mathbf{y}(t) + \mathbf{c}(t)^T\mathbf{y}(t)dt + \tfrac{1}{2}\mathbf{y}(T)^T\mathbf{Q}_T\mathbf{y}(T) + \mathbf{c}_T^T\mathbf{y}(T)$$
$$+ \tfrac{1}{2}\mathbf{s}^T\mathbf{R}\mathbf{s} + \frac{\sigma}{2}\int_0^T \mathbf{u}(t)^T\mathbf{u}(t)dt,$$

where for given $\mathbf{s},\mathbf{u}$ the state $\mathbf{y} \in H^1(0,T;\mathbb{R}^n)$ solves

$$\mathbf{M}\frac{d}{dt}\mathbf{y}(t) + \mathbf{A}\mathbf{y}(t) = \mathbf{f}(t) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{M}\mathbf{y}(0) = \mathbf{M}\mathbf{s}.$$

Weighted inner product in $\mathbb{R}^{n+n_u}$ with matrix $\begin{pmatrix} \mathbf{M} & \\ & \mathbf{I} \end{pmatrix}$

# Hessian Computation & Reduced Order Modeling

$$\nabla^2 J \begin{pmatrix} \mathbf{v} \\ \mathbf{d} \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{v} \\ \mathbf{d} \end{pmatrix}$$

First and second row are

$$\mathbf{q}(0) + \mathbf{M}^{-1}\mathbf{R}\mathbf{v} \quad \text{and} \quad -\int_0^T \mathbf{B}(t)^T \mathbf{q}(t) dt + \sigma \int_0^T \mathbf{d}(t) dt$$

where $\mathbf{z}$ and $\mathbf{q}$ solve

$$\begin{aligned} \mathbf{M}\frac{d}{dt}\mathbf{z}(t) + \mathbf{A}\mathbf{z}(t) &= \mathbf{B}\mathbf{d}(t), & \mathbf{M}\mathbf{z}(0) &= \mathbf{M}\mathbf{v}, \\ -\mathbf{M}\frac{d}{dt}\mathbf{q}(t) + \mathbf{A}^T\mathbf{q}(t) &= \mathbf{Q}\,\mathbf{z}(t), & \mathbf{M}\mathbf{q}(T) &= \mathbf{Q}_T\mathbf{z}(T). \end{aligned}$$

Modified basic ROM: Add columns of $\mathbf{B}$ to basic ROM.
Now in setting as before.
Control inputs included since range($\mathbf{B}$) $\subset$ range($\mathbf{W}$).
But, outputs not fully captured - are working on better 'control ROM' (rigorous incorporation of final time observation).

# Numerical Results

- Model problem with symmetric $\mathbf{A}$
- Control regularization $\sigma = 10^{-3}$ and controls are acting on

$$p_1 = (0.6, 0.3), \quad p_2 = (0.2, 0.5), \quad p_3 = (0.70.6).$$

$\rightarrow$ $n = 841$ and 150 discrete control variables

$k = 30$, $tol^{cg} = 10^{-4}\|\mathbf{g}\|$

| | $J$ | $\|\nabla J\|$ | CG iters | ROM size | full/ROM |
|---|---|---|---|---|---|
| at $(\mathbf{0}, \mathbf{0})$ | 1.14e-1 | 2.12e-1 | | | |
| full Hess | 2.67e-4 | 1.67e-5 | 33 | | 68 / 0 |
| $\widehat{\mathbf{H}}^{\rm bsc}$ | 3.78e+0 | 4.31e-1 | 26 | 33 | 2 / 52 |
| $\widehat{\mathbf{H}}$,$\mathbf{a} = \mathbf{g}$ | 3.00e-4 | 1.42e-3 | 21 | 33+1 | 2 / 42 |

# Summary

- Optimal control problems with controls in initial conditions arise often
  – large number of control variables makes these problems expensive to solve
- Use ROMs to reduce computational expense
  – initial data vary in $\mathbb{R}^n$, $n \gg 1$
  – no single ROM provides good approximation
- Introduced ROM augmentation for inexpensive, but good Hessian approx.
  – substantially improves ROM quality for variable initial data / right hand side,
  – computationally inexpensive,
  – complete analysis for important diagonalizable case.
- Sequential application of augmented ROMs gives approximate optimizer with same accuracy as the full model approach but at a faction of cost.
- Introduced augmented ROMs for OCPs with controls in initial conditions and right hand side.
- Extensions still needed:
  – to time varying problems $\mathbf{A} \to \mathbf{A}(t)$,
  – ...