An Introduction to System-theoretic Methods for Model Reduction - Part III Preserving System Structure

Christopher Beattie

Department of Mathematics, Virginia Tech
Division of Computational Modeling and Data Analytics, Virginia Tech

ICERM Semester Program - Spring 2020
Model and dimension reduction in uncertain and dynamic systems
Providence, RI

February 4, 2020

Outline of First Lecture

Explicitly Structured Dynamical Systems

- Reminder: Projective reduction for linear dynamical systems presented in standard first order form.
- Examples of Explicitly Structured Dynamical Systems
  - partitioning state variables (don't mix apples and oranges)
  - second-order structure (vibrating systems)
  - propagation delays and system memory (viscoelastic models)
  - parametrized systems (inverse problems and optimization)
- Unifying framework: General coprime realizations
  - Interpolatory projections that retain structure
  - Backward stability for interpolatory methods with inexact solves

Classic problem setting

Standard "state space" description:

$$\mathbf{u}(t) \longrightarrow \boxed{\begin{array}{c} \mathbf{E}\dot{\mathbf{x}} = \mathbf{A}\,\mathbf{x} + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\,\mathbf{x}(t) \end{array}} \longrightarrow \mathbf{y}(t)$$

- $\mathbf{E}$, $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times p}$ and $\mathbf{C} \in \mathbb{R}^{m\times n}$ with n (state space dimension) very large: $n \gg m, p$.
- "Internal state" $\mathbf{x}(t)$ is assumed to be unimportant.
- Usual goal: Reduce the state space dimension without degrading the input-output map "$u \mapsto y$"

Find a "smaller" dynamical system
with nearly the same input/output map.

Model Reduction Heuristics

$$\begin{aligned} \mathbf{E}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) \end{aligned} \quad \overset{?}{\approx} \quad \begin{aligned} \mathbf{E}_r\dot{\mathbf{x}}_r &= \mathbf{A}_r\mathbf{x}_r + \mathbf{B}_r\,\mathbf{u}(t) \\ \mathbf{y}_r(t) &= \mathbf{C}_r\mathbf{x}_r(t) \end{aligned}$$

(Full system) (Reduced system)

Eliminate low value subspaces

- Original $\mathbf{x}(t)$ may linger close to low dimensional subspaces that are relatively insensitive to variations in input $\mathbf{u}(t)$.
- Original $\mathbf{x}(t)$ may have components of motion that have little influence on $\mathbf{y}(t)$ - low-visibility components.
- We may eliminate attractive components with low visibility and high visibility components that are not attractive but do not eliminate attractive components with high visibility.

Model Reduction Heuristics

$$\begin{aligned} \mathbf{E}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) \end{aligned} \quad \overset{?}{\approx} \quad \begin{aligned} \mathbf{E}_r\dot{\mathbf{x}}_r &= \mathbf{A}_r\mathbf{x}_r + \mathbf{B}_r\,\mathbf{u}(t) \\ \mathbf{y}_r(t) &= \mathbf{C}_r\mathbf{x}_r(t) \end{aligned}$$

(Full system) (Reduced system)

Eliminate low value subspaces

- Original $\mathbf{x}(t)$ may linger close to low dimensional subspaces that are relatively insensitive to variations in input $\mathbf{u}(t)$.
  Project dynamics *onto* "attractive" r-dimensional subspaces.
- Original $\mathbf{x}(t)$ may have components of motion that have little influence on $\mathbf{y}(t)$ - low-visibility components.
  Project dynamics *along* "low-visibility" codimension-r subspaces.
- We may eliminate attractive components with low visibility and high visibility components that are not attractive but do not eliminate attractive components with high visibility.
  Balancing addresses this tradeoff rigorously.

Projection Framework

$$\begin{array}{c}\mathbf{E}\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)\end{array} \approx \begin{array}{c}\mathbf{E}_r\dot{\mathbf{x}}_r = \mathbf{A}_r\mathbf{x}_r + \mathbf{B}_r\,\mathbf{u}(t) \\ \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t)\end{array}$$

(Original system) (Reduced system)

- Suppose $\mathcal{W}_r = \mathrm{Ran}(\mathbf{W}_r)$ and $\mathcal{V}_r = \mathrm{Ran}(\mathbf{V}_r)$ are r-dimensional subspaces such that $\mathcal{V}_r \cap \mathcal{W}_r^\perp = \{0\}$. Choose bases so that $\mathbf{W}_r^T\mathbf{V}_r = \mathbf{I}$.
The (skew) projection $\mathbf{P}_r = \mathbf{V}_r\mathbf{W}_r^T$ projects onto $\mathcal{V}_r$ along $\mathcal{W}_r^\perp$.
- $\mathcal{V}_r$ should represent an “attractive” r-dimensional subspace
$\mathcal{W}_r^\perp$ should represent a “low-visibility” codimension-r subspace.
- “Project dynamics” by approximating $\mathbf{x}(t) \approx \mathbf{V}_r\mathbf{x}_r(t)$
and constraining the reduced trajectory $\mathbf{x}_r(t)$ to satisfy

$$\mathbf{W}_r^T\left(\mathbf{E}\mathbf{V}_r\dot{\mathbf{x}}_r(t) - \mathbf{A}\mathbf{V}_r\mathbf{x}_r(t) - \mathbf{B}\,\mathbf{u}(t)\right) = 0 \quad \text{(Petrov-Galerkin)}$$

- Leads to a reduced model: $\mathbf{E}_r = \mathbf{W}_r^T\mathbf{E}\mathbf{V}_r \in \mathbb{R}^{r\times r}$,

$$\mathbf{A}_r = \mathbf{W}_r^T\mathbf{A}\mathbf{V}_r \in \mathbb{R}^{r\times r}, \quad \mathbf{C}_r = \mathbf{C}\mathbf{V}_r \in \mathbb{R}^{m\times r}, \quad \mathbf{B}_r = \mathbf{W}_r^T\mathbf{B} \in \mathbb{R}^{r\times p}.$$

Projection Framework

$$\begin{aligned}\mathbf{E}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\,\mathbf{u}(t)\\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t)\end{aligned} \quad \approx \quad \begin{aligned}\mathbf{E}_r\dot{\mathbf{x}}_r &= \mathbf{A}_r\mathbf{x}_r + \mathbf{B}_r\,\mathbf{u}(t)\\ \mathbf{y}_r(t) &= \mathbf{C}_r\mathbf{x}_r(t)\end{aligned}$$

(Original system) (Reduced system)

- Suppose $\mathcal{W}_r = \mathrm{Ran}(\mathbf{W}_r)$ and $\mathcal{V}_r = \mathrm{Ran}(\mathbf{V}_r)$ are r-dimensional subspaces such that $\mathcal{V}_r \cap \mathcal{W}_r^{\perp} = \{0\}$. Choose bases so that $\mathbf{W}_r^T\mathbf{V}_r = \mathbf{I}$.
  The (skew) projection $\mathbf{P}_r = \mathbf{V}_r\mathbf{W}_r^T$ projects onto $\mathcal{V}_r$ along $\mathcal{W}_r^{\perp}$.
- $\mathcal{V}_r$ should represent an "attractive" r-dimensional subspace
  $\mathcal{W}_r^{\perp}$ should represent a "low-visibility" codimension-r subspace.
- "Project dynamics" by approximating $\mathbf{x}(t) \approx \mathbf{V}_r\mathbf{x}_r(t)$ and constraining the reduced trajectory $\mathbf{x}_r(t)$ to satisfy

$$\mathbf{W}_r^T\left(\mathbf{E}\mathbf{V}_r\dot{\mathbf{x}}_r(t) - \mathbf{A}\mathbf{V}_r\mathbf{x}_r(t) - \mathbf{B}\,\mathbf{u}(t)\right) = 0 \quad \text{(Petrov-Galerkin)}$$

- Leads to a reduced model: $\mathbf{E}_r = \mathbf{W}_r^T\mathbf{E}\mathbf{V}_r \in \mathbb{R}^{r\times r}$,

$$\mathbf{A}_r = \mathbf{W}_r^T\mathbf{A}\mathbf{V}_r \in \mathbb{R}^{r\times r}, \quad \mathbf{C}_r = \mathbf{C}\mathbf{V}_r \in \mathbb{R}^{m\times r}, \quad \mathbf{B}_r = \mathbf{W}_r^T\mathbf{B} \in \mathbb{R}^{r\times p}.$$

$$\begin{aligned}\mathbf{E}_r\dot{\mathbf{x}}_r &= \mathbf{A}_r\mathbf{x}_r + \mathbf{B}_r\,\mathbf{u}(t)\\ \mathbf{y}_r(t) &= \mathbf{C}_r\mathbf{x}_r(t)\end{aligned}$$

Rational Approximation

$$\begin{array}{c}\mathbf{E}\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)\end{array} \approx \begin{array}{c}\mathbf{E}_r\dot{\mathbf{x}}_r = \mathbf{A}_r\mathbf{x}_r + \mathbf{B}_r\,\mathbf{u}(t) \\ \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t)\end{array}$$

(Original system) (Reduced system)

Want outputs to be close, $\mathbf{y}_r \approx \mathbf{y}$, over a large class of inputs $\mathbf{u}$.

- Fourier Transforms: $\mathbf{u}(t) \mapsto \hat{\mathbf{u}}(\omega), \quad \mathbf{y}(t) \mapsto \hat{\mathbf{y}}(\omega)$

  Original response: $\hat{\mathbf{y}}(\omega) = \mathcal{H}(\imath\omega)\hat{\mathbf{u}}(\omega)$

  Reduced response: $\hat{\mathbf{y}}_r(\omega) = \mathcal{H}_r(\imath\omega)\hat{\mathbf{u}}(\omega)$

  with transfer functions:

  $\mathcal{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$ and $\mathcal{H}_r(s) = \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r$

- $$\hat{\mathbf{y}}(\omega) - \hat{\mathbf{y}}_r(\omega) = \Big(\mathcal{H}(\imath\omega) - \mathcal{H}_r(\imath\omega)\Big)\,\hat{\mathbf{u}}(\omega)$$

  Want $\mathcal{H}_r(\imath\omega) \approx \mathcal{H}(\imath\omega)$ for $\omega \in \mathbb{R}$.

Interpolation Framework

$$\begin{array}{c} \mathbf{E}\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \end{array} \approx \begin{array}{c} \mathbf{E}_r\dot{\mathbf{x}}_r = \mathbf{A}_r\mathbf{x}_r + \mathbf{B}_r\,\mathbf{u}(t) \\ \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t) \end{array}$$

(Original system) (Reduced system)

$$\mathcal{H}(s) \approx \mathcal{H}_r(s)\ ?$$

- Performance measures:

$$\|\mathcal{H} - \mathcal{H}_r\|_{\mathcal{H}_2} = \left(\int_{-\infty}^{\infty} |\mathcal{H}(\imath\omega) - \mathcal{H}_r(\imath\omega)|^2 d\omega\right)^{1/2} \quad \text{"}\mathcal{H}_2 \text{ error"}$$

(try to make $\|y - y_r\|_{L_\infty} / \|u\|_{L_2}$ small)

$$\|\mathcal{H} - \mathcal{H}_r\|_{\mathcal{H}_\infty} = \sup_\omega |\mathcal{H}(\imath\omega) - \mathcal{H}_r(\imath\omega)| \quad \text{"}\mathcal{H}_\infty \text{ error"}$$

(try to make $\|y - y_r\|_{L_2} / \|u\|_{L_2}$ small)

- *Interpolation* is a necessary condition for a best rational approximation, $\mathcal{H}_r \approx \mathcal{H}$ in each case.

Interpolation Framework

$$\begin{aligned} \mathbf{E}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) \end{aligned} \quad \approx \quad \begin{aligned} \mathbf{E}_r\dot{\mathbf{x}}_r &= \mathbf{A}_r\mathbf{x}_r + \mathbf{B}_r\,\mathbf{u}(t) \\ \mathbf{y}_r(t) &= \mathbf{C}_r\mathbf{x}_r(t) \end{aligned}$$

(Original system) (Reduced system)

$$\mathcal{H}(s) \quad \approx \quad \mathcal{H}_r(s) \text{ ?}$$

- Performance measures:

$$\|\mathcal{H} - \mathcal{H}_r\|_{\mathcal{H}_2} = \left(\int_{-\infty}^{\infty} |\mathcal{H}(\imath\omega) - \mathcal{H}_r(\imath\omega)|^2 d\omega\right)^{1/2} \qquad \text{"}\mathcal{H}_2 \text{ error"}$$

(try to make $\|y - y_r\|_{L_\infty} / \|u\|_{L_2}$ small)

$$\|\mathcal{H} - \mathcal{H}_r\|_{\mathcal{H}_\infty} = \sup_{\omega} |\mathcal{H}(\imath\omega) - \mathcal{H}_r(\imath\omega)| \qquad \text{"}\mathcal{H}_\infty \text{ error"}$$

(try to make $\|y - y_r\|_{L_2} / \|u\|_{L_2}$ small)

- *Interpolation* is a necessary condition for a best rational approximation, $\mathcal{H}_r \approx \mathcal{H}$ in each case.

Find reduced models, $\mathcal{H}_r(s)$, that interpolate $\mathcal{H}(s)$ at selected points $\sigma_1, \sigma_2, \ldots \subset \mathbb{C}$.

Interpolatory projections

The key fact that ties interpolation together with projection methods for standard first-order state-space realizations:

Theorem

Suppose $\mathbb{b} \in \mathbb{R}^p$ and $\mathbb{c} \in \mathbb{R}^m$ are arbitrary vectors.

a. *If* $(\sigma \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathbb{b} \in \text{Ran}(\mathbf{V}_r)$ *then* $\mathcal{H}(\sigma)\mathbb{b} = \mathcal{H}_r(\sigma)\mathbb{b}$

b. *If* $(\mu \mathbf{E}^T - \mathbf{A}^T)^{-1}\mathbf{C}^T\mathbb{c} \in \text{Ran}(\mathbf{W}_r)$ *then* $\mathbb{c}^T\mathcal{H}(\mu) = \mathbb{c}^T\mathcal{H}_r(\mu)$.

c. *If both (a) and (b) hold with* $\sigma = \mu$ *then* $\mathbb{c}^T\mathcal{H}'(\sigma)\mathbb{b} = \mathbb{c}^T\mathcal{H}_r'(\sigma)\mathbb{b}$.

Thus, given r distinct interpolation points: $\{\sigma_i\}_{i=1}^r$ and directions $\{\mathbb{b}_i\}_{i=1}^r$, $\{\mathbb{c}_i\}_{i=1}^r$, if

$$\mathbf{V}_r = \left[(\sigma_1 \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathbb{b}_1, \cdots, (\sigma_r \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathbb{b}_r\right], \quad \mathbf{W}_r^T = \begin{bmatrix} \mathbb{c}_1^T\mathbf{C}(\sigma_1 \mathbf{E} - \mathbf{A})^{-1} \\ \vdots \\ \mathbb{c}_r^T\mathbf{C}(\sigma_r \mathbf{E} - \mathbf{A})^{-1} \end{bmatrix},$$

then $\mathcal{H}(\sigma_i)\mathbb{b}_i = \mathcal{H}_r(\sigma_i)\mathbb{b}_i$, $\mathbb{c}_i^T\mathcal{H}(\sigma_i) = \mathbb{c}_i^T\mathcal{H}_r(\sigma_i)$, and $\mathbb{c}_i^T\mathcal{H}'(\sigma_i)\mathbb{b}_i = \mathbb{c}_i^T\mathcal{H}_r'(\sigma_i)\mathbb{b}_i$ for $i = 1, \ldots, r$

Structure-preserving model reduction

$$\mathbf{u}(t) \longrightarrow \boxed{\begin{array}{c} \mathbf{A}_0 \frac{d^\ell \mathbf{x}}{dt^\ell} + \mathbf{A}_1 \frac{d^{\ell-1} \mathbf{x}}{dt^{\ell-1}} + \ldots + \mathbf{A}_\ell \mathbf{x} = \mathbf{B}_0 \frac{d^k \mathbf{u}}{dt^k} + \ldots + \mathbf{B}_k \mathbf{u} \\ \mathbf{y}(t) = \mathbf{C}_0 \frac{d^q \mathbf{x}}{dt^q} + \ldots + \mathbf{C}_q \mathbf{x}(t) \end{array}} \longrightarrow \mathbf{y}(t)$$

- "Every linear ODE is equivalent to a first-order ODE system"
  Might not be the best approach ...
- The "state space" is an aggregate of dynamic variables some of which may be internal and "locked" to other variables.
- *Refined goal*: Want to develop model reduction methods that can reduce selected state variables (i.e., on selected subspaces) while leaving other state variables untouched; maintaining the previous structural relationships among the variables.

  Note *order* reduction is distinguished from *dimension* reduction

Structure-preserving model reduction

$$\mathbf{u}(t) \longrightarrow \boxed{\begin{array}{c} \mathbf{A}_0 \frac{d^\ell \mathbf{x}}{dt^\ell} + \mathbf{A}_1 \frac{d^{\ell-1} \mathbf{x}}{dt^{\ell-1}} + \ldots + \mathbf{A}_\ell \mathbf{x} = \mathbf{B}_0 \frac{d^k \mathbf{u}}{dt^k} + \ldots + \mathbf{B}_k \mathbf{u} \\ \mathbf{y}(t) = \mathbf{C}_0 \frac{d^q \mathbf{x}}{dt^q} + \ldots + \mathbf{C}_q \mathbf{x}(t) \end{array}} \longrightarrow \mathbf{y}(t)$$

- "Every linear ODE is equivalent to a first-order ODE system"
  Might not be the best approach ...
- The "state space" is an aggregate of dynamic variables some of which may be internal and "locked" to other variables.
- *Refined goal*: Want to develop model reduction methods that can reduce selected state variables (i.e., on selected subspaces) while leaving other state variables untouched; maintaining the previous structural relationships among the variables.

  Note *order* reduction is distinguished from *dimension* reduction

Example 1: Incompressible viscoelastic vibration

$$\partial_{tt}\mathbf{w}(x,t) - \eta\,\Delta\mathbf{w}(x,t) - \int_0^t \rho(t-\tau)\,\Delta\mathbf{w}(x,\tau)\,d\tau + \nabla\varpi(x,t) = \mathbf{b}(x)\cdot\mathbf{u}(t),$$

$\nabla\cdot\mathbf{w}(x,t) = 0$ which determines $\mathbf{y}(t) = [\varpi(x_1,t),\,\ldots,\,\varpi(x_p,t)]^T$

- $\mathbf{w}(x,t)$ is the displacement field; $\varpi(x,t)$ is the pressure field; $\rho(\tau)$ is a “relaxation function”

Example 1: Incompressible viscoelastic vibration

$$\partial_{tt}\mathbf{w}(x,t) - \eta\,\Delta\mathbf{w}(x,t) - \int_0^t \rho(t-\tau)\,\Delta\mathbf{w}(x,\tau)\,d\tau + \nabla\varpi(x,t) = \mathbf{b}(x)\cdot\mathbf{u}(t),$$

$$\nabla\cdot\mathbf{w}(x,t) = 0 \quad \text{which determines} \quad \mathbf{y}(t) = [\varpi(x_1,t),\,\ldots,\,\varpi(x_p,t)]^T$$

- $\mathbf{w}(x,t)$ is the displacement field; $\varpi(x,t)$ is the pressure field; $\rho(\tau)$ is a “relaxation function”

$$\mathbf{M}\,\ddot{\mathbf{x}}(t) + \eta\,\mathbf{K}\,\mathbf{x}(t) + \int_0^t \rho(t-\tau)\,\mathbf{K}\,\mathbf{x}(\tau)\,d\tau + \mathbf{D}\,\boldsymbol{\varpi}(t) = \mathbf{B}\,\mathbf{u}(t),$$

$$\mathbf{D}^T\,\mathbf{x}(t) = \mathbf{0}, \qquad \text{which determines} \quad \mathbf{y}(t) = \mathbf{C}\,\boldsymbol{\varpi}(t)$$

- $\mathbf{x}\in\mathbb{R}^{n_1}$ discretization of $\mathbf{w}$; $\boldsymbol{\varpi}\in\mathbb{R}^{n_2}$ discretization of ϖ.
- $\mathbf{M}$ and $\mathbf{K}$ are real, symmetric, positive-definite matrices,
  $\mathbf{B}\in\mathbb{R}^{n_1\times m}$, $\mathbf{C}\in\mathbb{R}^{p\times n_2}$, and $\mathbf{D}\in\mathbb{R}^{n_1\times n_2}$.

Example 1: Incompressible viscoelastic vibration

Transfer function (need not be a rational function !):

$$\mathcal{H}(s) = \begin{bmatrix} \mathbf{0} & \mathbf{C} \end{bmatrix} \begin{bmatrix} s^2\mathbf{M} + (\widehat{\rho}(s) + \eta)\,\mathbf{K} & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix}$$

- Want a reduced order model that replicates input-output response with high fideliety yet retains “viscoelasticity”:

$$\mathbf{M}_r\,\ddot{\mathbf{x}}_r(t) + \eta\,\mathbf{K}_r\,\mathbf{x}_r(t) + \int_0^t \rho(t-\tau)\,\mathbf{K}_r\,\mathbf{x}_r(\tau)\,d\tau + \mathbf{D}_r\,\varpi_r(t) = \mathbf{B}_r\,\mathbf{u}(t),$$

$$\mathbf{D}_r^T\,\mathbf{x}_r(t) = \mathbf{0}, \qquad \text{which determines} \quad \mathbf{y}_r(t) = \mathbf{C}_r\,\varpi_r(t)$$

  with symmetric positive semidefinite $\mathbf{M}_r$, $\mathbf{K}_r \in \mathbb{R}^{r\times r}$, with $\mathbf{B}_r \in \mathbb{R}^{r\times m}$, $\mathbf{C}_r \in \mathbb{R}^{p\times r}$, and $\mathbf{D}_r \in \mathbb{R}^{r\times r}$.

- Because of the memory term, both reduced and original systems have *infinite-order.*

Example 1: Incompressible viscoelastic vibration

Transfer function (need not be a rational function !):

$$\mathcal{H}(s) = [\,\mathbf{0} \quad \mathbf{C}\,] \begin{bmatrix} s^2\mathbf{M} + (\widehat{\rho}(s) + \eta)\,\mathbf{K} & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix}$$

- Want a reduced order model that replicates input-output response with high fideliety yet retains "viscoelasticity":

$$\mathbf{M}_r\,\ddot{\mathbf{x}}_r(t) \;+\; \eta\,\mathbf{K}_r\,\mathbf{x}_r(t) \;+\; \int_0^t \rho(t-\tau)\,\mathbf{K}_r\,\mathbf{x}_r(\tau)\,d\tau + \mathbf{D}_r\,\boldsymbol{\varpi}_r(t) = \mathbf{B}_r\,\mathbf{u}(t),$$

$$\mathbf{D}_r^T\,\mathbf{x}_r(t) = \mathbf{0}, \qquad \text{which determines} \quad \mathbf{y}_r(t) = \mathbf{C}_r\,\boldsymbol{\varpi}_r(t)$$

with symmetric positive semidefinite $\mathbf{M}_r$, $\mathbf{K}_r \in \mathbb{R}^{r\times r}$, with $\mathbf{B}_r \in \mathbb{R}^{r\times m}$, $\mathbf{C}_r \in \mathbb{R}^{p\times r}$, and $\mathbf{D}_r \in \mathbb{R}^{r\times r}$.

- Because of the memory term, both reduced and original systems have *infinite-order.*

Example 1: Incompressible viscoelastic vibration

Transfer function (need not be a rational function !):

$$\mathcal{H}(s) = [\,\mathbf{0} \quad \mathbf{C}\,] \begin{bmatrix} s^2\mathbf{M} + (\widehat{\rho}(s) + \eta)\,\mathbf{K} & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix}$$

- Want a reduced order model that replicates input-output response with high fideliety yet retains “viscoelasticity”:

$$\mathbf{M}_r\,\ddot{\mathbf{x}}_r(t) \;+\; \eta\,\mathbf{K}_r\,\mathbf{x}_r(t) \;+\; \int_0^t \rho(t-\tau)\,\mathbf{K}_r\,\mathbf{x}_r(\tau)\,d\tau + \mathbf{D}_r\,\boldsymbol{\varpi}_r(t) = \mathbf{B}_r\,\mathbf{u}(t),$$

$$\mathbf{D}_r^T\,\mathbf{x}_r(t) = \mathbf{0}, \qquad \text{which determines} \quad \mathbf{y}_r(t) = \mathbf{C}_r\,\boldsymbol{\varpi}_r(t)$$

with symmetric positive semidefinite $\mathbf{M}_r,\ \mathbf{K}_r \in \mathbb{R}^{r\times r}$, with $\mathbf{B}_r \in \mathbb{R}^{r\times m}$, $\mathbf{C}_r \in \mathbb{R}^{p\times r}$, and $\mathbf{D}_r \in \mathbb{R}^{r\times r}$.

- Because of the memory term, both reduced and original systems have *infinite-order.*

Example 2: Delay Differential System

- Many complex processes exhibit some sort of delayed response in their input, output, or internal dynamics.

Often related to ancillary unmodeled dynamics that create a time lag due to communication, material transport, or inertial effects occuring at a finer scale than captured in the model.

$$\begin{aligned}\mathbf{E}\dot{\mathbf{x}}(t) &= \mathbf{A}_1\mathbf{x}(t) + \mathbf{A}_2\mathbf{x}(t-\tau) + \mathbf{F}\,\mathbf{u}(t),\\ \mathbf{y}(t) &= \mathbf{D}\mathbf{x}(t)\end{aligned}$$

- Delay systems are also infinite-order. The dynamic effects of even a small delay can be profound.
- Find a reduced order model retaining the same delay structure:

$$\begin{aligned}\mathbf{E}_r\dot{\mathbf{x}}_r(t) &= \mathbf{A}_{1r}\mathbf{x}_r(t) + \mathbf{A}_{2r}\mathbf{x}_r(t-\tau) + \mathbf{F}_r\,\mathbf{u}(t),\\ \mathbf{y}_r(t) &= \mathbf{D}_r\mathbf{x}_r(t)\end{aligned}$$

$$\mathcal{H}(s) = \mathbf{D}\left(s\mathbf{E} - \mathbf{A}_1 - \mathbf{A}_2\, e^{-s\tau}\right)^{-1}\mathbf{F} \;\rightarrow\; \mathcal{H}_r(s) = \mathbf{D}_r\left(s\mathbf{E}_r - \mathbf{A}_{1r} - \mathbf{A}_{2r}\, e^{-s\tau}\right)^{-1}\mathbf{F}_r$$

Projective reduction for coprime realizations

Suppose a transfer function $\mathcal{H}(s)$ has a known decomposition:

$$\mathcal{H}(s) = \mathcal{C}(s)\mathcal{K}(s)^{-1}\mathcal{B}(s)$$

("general coprime realization") where the factors

- $\mathcal{C}(s) \in \mathbb{C}^{p\times n}$ and $\mathcal{B}(s) \in \mathbb{C}^{n\times m}$ are analytic for s in the right half plane, and
- $\mathcal{K}(s) \in \mathbb{C}^{n\times n}$ is both analytic and full rank for s in the right half plane.

This realization should reflect the system "structure" that is valued.

Reduced models can be constructed via projection as before:

- Pick full rank constant matrices $\mathbf{V}_r \in \mathbb{C}^{n\times r}$ and $\mathbf{W}_r \in \mathbb{C}^{n\times r}$

Reduced model $\mathcal{H}_r(s) = \mathcal{C}_r(s)\mathcal{K}_r(s)^{-1}\mathcal{B}_r(s)$ is obtained by defining

$$\mathcal{K}_r(s) = \mathbf{W}_r^T\mathcal{K}(s)\mathbf{V}_r, \quad \mathcal{B}_r(s) = \mathbf{W}_r^T\mathcal{B}(s), \quad \text{and} \quad \mathcal{C}_r(s) = \mathcal{C}(s)\mathbf{V}_r.$$

$\mathbf{V}_r$ and $\mathbf{W}_r$ can often be chosen so that structure is preserved.

Example 1 again

Framework: $\mathcal{H}(s) = \mathcal{C}(s)\mathcal{K}(s)^{-1}\mathcal{B}(s)$ and $\mathcal{H}_r(s) = \mathcal{C}_r(s)\mathcal{K}_r(s)^{-1}\mathcal{B}_r(s)$

$$\mathcal{K}_r(s) = \mathbf{W}_r^T\mathcal{K}(s)\mathbf{V}_r, \quad \mathcal{B}_r(s) = \mathbf{W}_r^T\mathcal{B}(s), \quad \text{and} \quad \mathcal{C}_r(s) = \mathcal{C}(s)\mathbf{V}_r.$$

- $\mathcal{H}(s) = [\,\mathbf{0}\ \mathbf{C}\,] \begin{bmatrix} s^2\mathbf{M} + (\widehat{\rho}(s) + \eta)\,\mathbf{K} & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix}$
- $$\mathcal{K}(s) = \begin{bmatrix} s^2\mathbf{M} + (\widehat{\rho}(s) + \eta)\,\mathbf{K} & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0} \end{bmatrix};$$
  $$\mathcal{B}(s) = \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix}; \quad \mathcal{C}(s) = [\,\mathbf{0} \quad \mathbf{C}\,].$$

To maintain symmetry and positive definiteness, $\mathbf{W}_r = \mathbf{V}_r = \begin{bmatrix} \mathbf{U}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_r \end{bmatrix}$:

$$\mathcal{K}_r(s) = \mathbf{V}_r^T\mathcal{K}(s)\mathbf{V}_r = \begin{bmatrix} s^2\mathbf{M}_r + (\widehat{\rho}(s) + \eta)\,\mathbf{K}_r & \mathbf{D}_r \\ \mathbf{D}_r^T & \mathbf{0} \end{bmatrix}$$

with $\mathbf{M}_r = \mathbf{U}_r^T\mathbf{M}\mathbf{U}_r$; $\mathbf{K}_r = \mathbf{V}_r^T\mathbf{K}\mathbf{V}_r$ and $\mathbf{D}_r = \mathbf{U}_r^T\mathbf{D}\mathbf{Z}_r$.

Example 1 again

Framework: $\mathcal{H}(s) = \mathcal{C}(s)\mathcal{K}(s)^{-1}\mathcal{B}(s)$ and $\mathcal{H}_r(s) = \mathcal{C}_r(s)\mathcal{K}_r(s)^{-1}\mathcal{B}_r(s)$

$$\mathcal{K}_r(s) = \mathbf{W}_r^T \mathcal{K}(s)\mathbf{V}_r, \quad \mathcal{B}_r(s) = \mathbf{W}_r^T \mathcal{B}(s), \quad \text{and} \quad \mathcal{C}_r(s) = \mathcal{C}(s)\mathbf{V}_r.$$

- $\mathcal{H}(s) = [\,\mathbf{0}\ \mathbf{C}\,] \begin{bmatrix} s^2\mathbf{M} + (\widehat{\rho}(s)+\eta)\,\mathbf{K} & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix}$
- $$\mathcal{K}(s) = \begin{bmatrix} s^2\mathbf{M} + (\widehat{\rho}(s)+\eta)\,\mathbf{K} & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0} \end{bmatrix};$$
  $$\mathcal{B}(s) = \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix}; \quad \mathcal{C}(s) = [\,\mathbf{0} \quad \mathbf{C}\,].$$

To maintain symmetry and positive definiteness, $\mathbf{W}_r = \mathbf{V}_r = \begin{bmatrix} \mathbf{U}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_r \end{bmatrix}$:

$$\mathcal{K}_r(s) = \mathbf{V}_r^T \mathcal{K}(s)\mathbf{V}_r = \begin{bmatrix} s^2\mathbf{M}_r + (\widehat{\rho}(s)+\eta)\,\mathbf{K}_r & \mathbf{D}_r \\ \mathbf{D}_r^T & \mathbf{0} \end{bmatrix}$$

with $\mathbf{M}_r = \mathbf{U}_r^T\mathbf{M}\mathbf{U}_r$; $\mathbf{K}_r = \mathbf{V}_r^T\mathbf{K}\mathbf{V}_r$ and $\mathbf{D}_r = \mathbf{U}_r^T\mathbf{D}\mathbf{Z}_r$.

Interpolatory projections for structured systems

Theorem

Suppose that $\mathcal{B}(s)$, $\mathcal{C}(s)$, and $\mathcal{K}(s)$ are analytic at a point $\sigma \in \mathbb{C}$ and both $\mathcal{K}(\sigma)$ and $\mathcal{K}_r(\sigma) = \mathbf{W}_r^T \mathcal{K}(\sigma) \mathbf{V}_r$ have full rank. Suppose $\mathbb{b} \in \mathbb{C}^p$ and $\mathbb{c} \in \mathbb{C}^m$ are arbitrary nontrivial vectors.

- *If* $\mathcal{K}(\sigma)^{-1}\mathcal{B}(\sigma)\mathbb{b} \in \mathrm{Ran}(\mathbf{V}_r)$ *then* $\mathcal{H}(\sigma)\mathbb{b} = \mathcal{H}_r(\sigma)\mathbb{b}$.
- *If* $\left(\mathbb{c}^T \mathcal{C}(\sigma)\mathcal{K}(\sigma)^{-1}\right)^T \in \mathrm{Ran}(\mathbf{W}_r)$ *then* $\mathbb{c}^T\mathcal{H}(\sigma) = \mathbb{c}^T\mathcal{H}_r(\sigma)$
- *If* $\mathcal{K}(\sigma)^{-1}\mathcal{B}(\sigma)\mathbb{b} \in \mathrm{Ran}(\mathbf{V}_r)$ *and* $\left(\mathbb{c}^T \mathcal{C}(\sigma)\mathcal{K}(\sigma)^{-1}\right)^T \in \mathrm{Ran}(\mathbf{W}_r)$
  then $\mathbb{c}^T\mathcal{H}'(\sigma)\mathbb{b} = \mathbb{c}^T\mathcal{H}_r'(\sigma)\mathbb{b}$

Can build up projecting subspaces based
on interpolation data as in the standard case.

Optimal interpolation points are difficult to characterise;
(but good ones are often not hard to obtain)

Proof Outline

For ε small, $\mathfrak{K}(\sigma+\varepsilon)^{-1}\mathfrak{B}(\sigma+\varepsilon)\mathbb{b} = \mathfrak{K}(\sigma)^{-1}\mathfrak{B}(\sigma)\mathbb{b} + \mathcal{O}(\varepsilon)$

and $\mathbb{c}^T\mathfrak{C}(\sigma+\varepsilon)\mathfrak{K}(\sigma+\varepsilon)^{-1} = \mathbb{c}^T\mathfrak{C}(\sigma)\mathfrak{K}(\sigma)^{-1} + \mathcal{O}(\varepsilon)$.

Define

$$\Pi_{\mathcal{V}} = \mathbf{V}_r\mathfrak{K}_r(\sigma+\varepsilon)^{-1}\mathbf{W}_r^T\mathfrak{K}(\sigma+\varepsilon) \quad \text{and}$$
$$\Pi_{\mathcal{W}} = \mathfrak{K}(\sigma+\varepsilon)\mathbf{V}_r\mathfrak{K}_r(\sigma+\varepsilon)^{-1}\mathbf{W}_r^T$$

First key point:

- $\Pi_{\mathcal{V}}$ is a skew projection onto $\text{Ran}(\mathbf{V}_r)$ independent of ε, and
- $\Pi_{\mathcal{W}}$ is a skew projection with $\text{Ker}(\Pi_{\mathcal{W}}) = \text{Ran}(\mathbf{W}_r)^{\perp}$ independent of ε.

Proof Outline

For ε small, $\mathcal{K}(\sigma+\varepsilon)^{-1}\mathcal{B}(\sigma+\varepsilon)\mathbb{b} = \mathcal{K}(\sigma)^{-1}\mathcal{B}(\sigma)\mathbb{b} + \mathcal{O}(\varepsilon)$

and $\mathbb{c}^T\mathcal{C}(\sigma+\varepsilon)\mathcal{K}(\sigma+\varepsilon)^{-1} = \mathbb{c}^T\mathcal{C}(\sigma)\mathcal{K}(\sigma)^{-1} + \mathcal{O}(\varepsilon)$.

Define

$$\Pi_{\mathcal{V}} = \mathbf{V}_r\mathcal{K}_r(\sigma+\varepsilon)^{-1}\mathbf{W}_r^T\mathcal{K}(\sigma+\varepsilon) \quad \text{and}$$

$$\Pi_{\mathcal{W}} = \mathcal{K}(\sigma+\varepsilon)\mathbf{V}_r\mathcal{K}_r(\sigma+\varepsilon)^{-1}\mathbf{W}_r^T$$

First key point:

- $\Pi_{\mathcal{V}}$ is a skew projection onto Ran$(\mathbf{V}_r)$ independent of ε, and
- $\Pi_{\mathcal{W}}$ is a skew projection with Ker$(\Pi_{\mathcal{W}}) =$ Ran$(\mathbf{W}_r)^{\perp}$ independent of ε.

Examine the pointwise error: e.g., to show $\mathbb{c}^T\mathcal{H}'(\sigma)\mathbb{b} = \mathbb{c}^T\mathcal{H}_r'(\sigma)\mathbb{b}$

$$\mathbb{c}^T\mathcal{H}(\sigma+\varepsilon)\mathbb{b} - \mathbb{c}^T\mathcal{H}_r(\sigma+\varepsilon)\mathbb{b} = \mathbb{c}^T\mathcal{C}(\sigma+\varepsilon)(\mathcal{K}(\sigma+\varepsilon)^{-1} - \mathcal{K}_r(\sigma+\varepsilon)^{-1})\mathcal{B}(\sigma+\varepsilon)\mathbb{b}$$

$$= \mathbb{c}^T\mathcal{C}(\sigma+\varepsilon)\mathcal{K}(\sigma+\varepsilon)^{-1}\left(\mathbf{I}-\Pi_{\mathcal{W}}\right)\mathcal{K}(\sigma+\varepsilon)\left(\mathbf{I}-\Pi_{\mathcal{V}}\right)\mathcal{K}(\sigma+\varepsilon)^{-1}\mathcal{B}(\sigma+\varepsilon)\mathbb{b}$$

Proof Outline

For ε small, $\mathcal{K}(\sigma+\varepsilon)^{-1}\mathcal{B}(\sigma+\varepsilon)\mathbb{b} = \mathcal{K}(\sigma)^{-1}\mathcal{B}(\sigma)\mathbb{b} + \mathcal{O}(\varepsilon)$
and $\mathbb{c}^T\mathcal{C}(\sigma+\varepsilon)\mathcal{K}(\sigma+\varepsilon)^{-1} = \mathbb{c}^T\mathcal{C}(\sigma)\mathcal{K}(\sigma)^{-1} + \mathcal{O}(\varepsilon)$.

Define

$$\Pi_{\mathcal{V}} = \mathbf{V}_r\mathcal{K}_r(\sigma+\varepsilon)^{-1}\mathbf{W}_r^T\mathcal{K}(\sigma+\varepsilon) \quad \text{and}$$

$$\Pi_{\mathcal{W}} = \mathcal{K}(\sigma+\varepsilon)\mathbf{V}_r\mathcal{K}_r(\sigma+\varepsilon)^{-1}\mathbf{W}_r^T$$

First key point:

- $\Pi_{\mathcal{V}}$ is a skew projection onto $\mathrm{Ran}(\mathbf{V}_r)$ independent of ε, and
- $\Pi_{\mathcal{W}}$ is a skew projection with $\mathrm{Ker}(\Pi_{\mathcal{W}}) = \mathrm{Ran}(\mathbf{W}_r)^{\perp}$ independent of ε.

Examine the pointwise error: e.g., to show $\mathbb{c}^T\mathcal{H}'(\sigma)\mathbb{b} = \mathbb{c}^T\mathcal{H}_r'(\sigma)\mathbb{b}$

$$\mathbb{c}^T\mathcal{H}(\sigma+\varepsilon)\mathbb{b} - \mathbb{c}^T\mathcal{H}_r(\sigma+\varepsilon)\mathbb{b} = \mathbb{c}^T\mathcal{C}(\sigma+\varepsilon)(\mathcal{K}(\sigma+\varepsilon)^{-1} - \mathcal{K}_r(\sigma+\varepsilon)^{-1})\mathcal{B}(\sigma+\varepsilon)\mathbb{b}$$

$$= \mathbb{c}^T\mathcal{C}(\sigma+\varepsilon)\mathcal{K}(\sigma+\varepsilon)^{-1}\Big(\mathbf{I}-\Pi_{\mathcal{W}}\Big)\mathcal{K}(\sigma+\varepsilon)\Big(\mathbf{I}-\Pi_{\mathcal{V}}\Big)\mathcal{K}(\sigma+\varepsilon)^{-1}\mathcal{B}(\sigma+\varepsilon)\mathbb{b}$$

$$= \mathcal{O}(\varepsilon^2)$$

Interpolatory projections in model reduction

- Given distinct (complex) frequencies $\{\sigma_1, \sigma_2, \ldots, \sigma_r\} \subset \mathbb{C}$, left tangent directions $\{\mathbb{c}_1, \ldots, \mathbb{c}_r\}$, and right tangent directions $\{\mathbb{b}_1, \ldots, \mathbb{b}_r\}$:

$$\mathbf{V}_r = \left[\mathcal{K}(\sigma_1)^{-1}\mathcal{B}(\sigma_1)\mathbb{b}_1, \cdots, \mathcal{K}(\sigma_r)^{-1}\mathcal{B}(\sigma_r)\mathbb{b}_r\right]$$

$$\mathbf{W}_r^T = \begin{bmatrix} \mathbb{c}_1^T\mathcal{C}(\sigma_1)\mathcal{K}(\sigma_1)^{-1} \\ \vdots \\ \mathbb{c}_r^T\mathcal{C}(\sigma_r)\mathcal{K}(\sigma_r)^{-1} \end{bmatrix}$$

- Guarantees that $\mathcal{H}(\sigma_j)\mathbb{b}_j = \mathcal{H}_r(\sigma_j)\mathbb{b}_j$,

$\mathbb{c}_j^T\mathcal{H}(\sigma_j) = \mathbb{c}_j^T\mathcal{H}_r(\sigma_j)$, $\quad \mathbb{c}_j^T\mathcal{H}'(\sigma_j)\mathbb{b}_j = \mathbb{c}_j^T\mathcal{H}_r'(\sigma_j)\mathbb{b}_j$

for $j = 1, 2, \ldots, r$.

Interpolatory projections in model reduction

- Given distinct (complex) frequencies $\{\sigma_1, \sigma_2, \ldots, \sigma_r\} \subset \mathbb{C}$, left tangent directions $\{\mathbb{c}_1, \ldots, \mathbb{c}_r\}$, and right tangent directions $\{\mathbb{b}_1, \ldots, \mathbb{b}_r\}$:

$$\mathbf{V}_r = \left[\mathcal{K}(\sigma_1)^{-1}\mathcal{B}(\sigma_1)\mathbb{b}_1, \cdots, \mathcal{K}(\sigma_r)^{-1}\mathcal{B}(\sigma_r)\mathbb{b}_r\right]$$

$$\mathbf{W}_r^T = \begin{bmatrix} \mathbb{c}_1^T\mathcal{C}(\sigma_1)\mathcal{K}(\sigma_1)^{-1} \\ \vdots \\ \mathbb{c}_r^T\mathcal{C}(\sigma_r)\mathcal{K}(\sigma_r)^{-1} \end{bmatrix}$$

- Guarantees that $\mathcal{H}(\sigma_j)\mathbb{b}_j = \mathcal{H}_r(\sigma_j)\mathbb{b}_j$,

  $\mathbb{c}_j^T\mathcal{H}(\sigma_j) = \mathbb{c}_j^T\mathcal{H}_r(\sigma_j), \quad \mathbb{c}_j^T\mathcal{H}'(\sigma_j)\mathbb{b}_j = \mathbb{c}_j^T\mathcal{H}_r'(\sigma_j)\mathbb{b}_j$

  for $j = 1, 2, \ldots, r$.

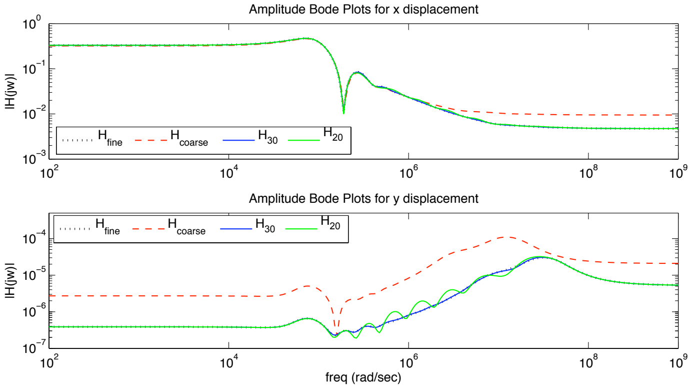
Example 1 (last time)

- A simple variation of the previous model:
- $\Omega = [0,1] \times [0,1]$: a volume filled with a viscoelastic material with boundary separated into a top edge ("lid"), $\partial\Omega_1$, and the complement, $\partial\Omega_0$ (bottom, left, and right edges).
- Excitation through shearing forces caused by transverse displacement of the lid, $u(t)$.
- Output: displacement $\mathbf{w}(\hat{x},t)$, at a fixed point $\hat{x} = (0.5, 0.5)$.

$$\partial_{tt}\mathbf{w}(x,t) - \eta_0\,\Delta\mathbf{w}(x,t) \;-\; \eta_1\partial_t \int_0^t \frac{\Delta\mathbf{w}(x,\tau)}{(t-\tau)^\alpha}\,d\tau \;+\; \nabla\varpi(x,t) = 0 \ \text{ for } x \in \Omega$$

$$\nabla \cdot \mathbf{w}(x,t) = 0 \text{ for } x \in \Omega,$$

$$\mathbf{w}(x,t) = 0 \text{ for } x \in \partial\Omega_0, \qquad \mathbf{w}(x,t) = u(t) \text{ for } x \in \partial\Omega_1$$



$\mathcal{H}_{\rm fine}$: $n_x = 51,842$ and $n_p = 6,651$ $\mathcal{H}_{30}$: $n_x = n_p = 30$

$\mathcal{H}_{\rm coarse}$: $n_x = 13,122$ $n_p = 1,681$ $\mathcal{H}_{20}$: $n_x = n_p = 20$

- $\mathcal{H}_{30}, \mathcal{H}_{20}$: reduced interpolatory viscoelastic models.
- $\mathcal{H}_{30}$ almost exactly replicates $\mathcal{H}_{\rm fine}$ and outperforms $\mathcal{H}_{\rm coarse}$
- Since input is a boundary *displacement* (as opposed to a boundary *force*), $\mathcal{B}(s) = s^2\,\mathbf{m} + \rho(s)\mathbf{k}$,

Example 2: Reduction of a Delay System

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}_0\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-\tau) + \mathbf{e}\,u(t) \quad \text{with} \quad \mathbf{y}(t) = \mathbf{e}^T\mathbf{x}(t)$$

with
$$\mathbf{E} = \kappa\,\mathbf{I} + \mathbf{T},\quad \mathbf{A}_0 = \frac{3}{\tau}(\mathbf{T} - \kappa\,\mathbf{I}),\quad \mathbf{A}_1 = \frac{1}{\tau}(\mathbf{T} - \kappa\,\mathbf{I})$$

$$\mathbf{T} = \text{diag}\begin{pmatrix} & 1 & & 1 & \cdots & & 1 & \\ 1 & & 0 & & & 0 & & 1 \\ & 1 & & 1 & \cdots & & 1 & \end{pmatrix}$$

Compare approaches:

- Direct (generalized) interpolation:
  $\mathcal{H}_r(s) = \mathbf{e}^T\mathbf{V}_r\left(s\mathbf{W}_r^T\mathbf{E}\mathbf{V}_r - \mathbf{W}_r^T\mathbf{A}_0\mathbf{V}_r + \mathbf{W}_r^T\mathbf{A}_1\mathbf{V}_r\,e^{-s\tau}\right)^{-1}\mathbf{W}_r^T\mathbf{e}.$
- Approximate delay term with rational function:

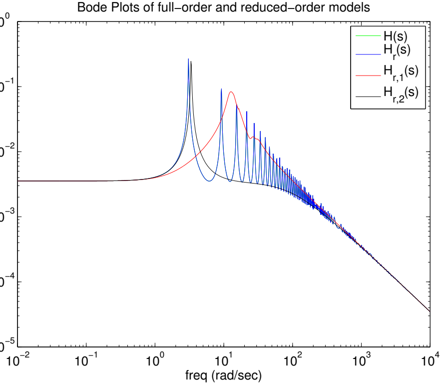
$$e^{-\tau s} \approx \frac{p_\ell(-\tau s)}{p_\ell(\tau s)}$$

  - Pass to $(\ell+1)^{st}$ order ODE system: $\mathbf{D}(s)\,\widehat{x}(s) = p_\ell(\tau s)\,\mathbf{e}\,\widehat{u}(s)$ with $\mathbf{D}(s) = (s\mathbf{E} - \mathbf{A}_0)\,p_\ell(\tau s) - \mathbf{A}_1 p_\ell(-\tau s)$.
  - Model reduction on linearization: first order system of dimension $(\ell+1)*n$. ($\rightarrow$Loss of structure!)

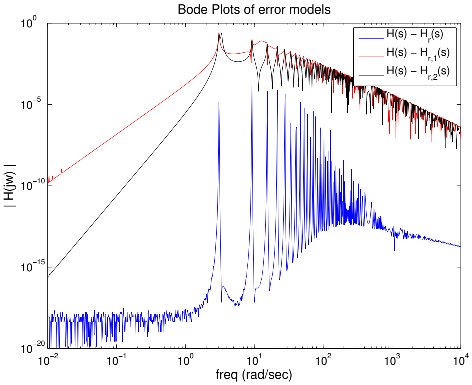
Example 2: Reduction of a Delay System

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}_1\mathbf{x}(t) + \mathbf{A}_2\mathbf{x}(t-\tau) + \mathbf{e}\,u(t) \quad \text{with} \quad \mathbf{y}(t) = \mathbf{e}^T\mathbf{x}(t)$$

$\mathcal{H}_r(s)$ - Generalized interpolation; $\mathcal{H}_{r,1}(s)$ - First-order Padé;



P



Original system dim: $n = 500$. Reduced system dim: $r = 10$.

Interpolation points: $\pm 1.0\text{E-}3\,\imath$, $\pm 3.16\text{E-}1\,\imath$, $\pm 5.0\,\imath$, $3.16\text{E+}1\,\imath$, $\pm 1.0\text{E+}3\imath$

Parametrized Dynamical Systems

Systems often depend on parameters...

- A designer/engineer/forecaster searches for optimal parameter values: to reduce cost, to improve efficiency, to minimize disturbance, predict trouble, etc.
- This results in complex large-scale optimization problems.
- Goal: Give the designer/engineer/forecaster a reduced parametric model with the same knobs to turn and optimize !!!
- Surrogate Optimization: Instead of solving

$$\min_{\mathsf{p}} \mathcal{J}(\mathbf{y};\, \mathbf{u},\, \mathsf{p}) \qquad \text{such that}$$

$$\mathbf{E}(\mathsf{p})\,\dot{\mathbf{x}}(t) = \mathbf{A}(\mathsf{p})\,\mathbf{x}(t) + \mathbf{B}(\mathsf{p})\,\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}(\mathsf{p})\,\mathbf{x}(t)$$

  solve

$$\min_{\mathsf{p}} \mathcal{J}(\mathbf{y}_r;\, \mathbf{u},\, \mathsf{p}) \qquad \text{such that}$$

$$\mathbf{E}_r(\mathsf{p})\,\dot{\mathbf{x}}_r(t) = \mathbf{A}_r(\mathsf{p})\,\mathbf{x}_r(t) + \mathbf{B}_r(\mathsf{p})\,\mathbf{u}(t), \quad \mathbf{y}_r(t) = \mathbf{C}_r(\mathsf{p})\,\mathbf{x}_r(t)$$

Parametrized Dynamical Systems

Systems often depend on parameters...

- A designer/engineer/forecaster searches for optimal parameter values: to reduce cost, to improve efficiency, to minimize disturbance, predict trouble, etc.
- This results in complex large-scale optimization problems.
- Goal: Give the designer/engineer/forecaster a reduced parametric model with the same knobs to turn and optimize !!!
- Surrogate Optimization: Instead of solving

$$\min_{\mathsf{p}} \mathcal{J}(\mathbf{y};\, \mathbf{u},\, \mathsf{p}) \qquad \text{such that}$$

$$\mathbf{E}(\mathsf{p})\, \dot{\mathbf{x}}(t) = \mathbf{A}(\mathsf{p})\, \mathbf{x}(t) + \mathbf{B}(\mathsf{p})\, \mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}(\mathsf{p})\, \mathbf{x}(t)$$

solve

$$\min_{\mathsf{p}} \mathcal{J}(\mathbf{y}_r;\, \mathbf{u},\, \mathsf{p}) \qquad \text{such that}$$

$$\mathbf{E}_r(\mathsf{p})\, \dot{\mathbf{x}}_r(t) = \mathbf{A}_r(\mathsf{p})\, \mathbf{x}_r(t) + \mathbf{B}_r(\mathsf{p})\, \mathbf{u}(t), \quad \mathbf{y}_r(t) = \mathbf{C}_r(\mathsf{p})\, \mathbf{x}_r(t)$$

Parametrized Dynamical Systems

Systems often depend on parameters...

- A designer/engineer/forecaster searches for optimal parameter values: to reduce cost, to improve efficiency, to minimize disturbance, predict trouble, etc.
- This results in complex large-scale optimization problems.
- Goal: Give the designer/engineer/forecaster a reduced parametric model with the same knobs to turn and optimize !!!
- Surrogate Optimization: Instead of solving

$$\min_{\mathsf{p}} \mathcal{J}(\mathbf{y};\, \mathbf{u},\, \mathsf{p}) \qquad \text{such that}$$

$$\mathbf{E}(\mathsf{p})\, \dot{\mathbf{x}}(t) = \mathbf{A}(\mathsf{p})\, \mathbf{x}(t) + \mathbf{B}(\mathsf{p})\, \mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}(\mathsf{p})\, \mathbf{x}(t)$$

solve

$$\min_{\mathsf{p}} \mathcal{J}(\mathbf{y}_r;\, \mathbf{u},\, \mathsf{p}) \qquad \text{such that}$$

$$\mathbf{E}_r(\mathsf{p})\, \dot{\mathbf{x}}_r(t) = \mathbf{A}_r(\mathsf{p})\, \mathbf{x}_r(t) + \mathbf{B}_r(\mathsf{p})\, \mathbf{u}(t), \quad \mathbf{y}_r(t) = \mathbf{C}_r(\mathsf{p})\, \mathbf{x}_r(t)$$

Parametrized Dynamical Systems

$\mathcal{H}(\mathsf{p}, s) = \mathbf{C}(\mathsf{p})\,(s\mathbf{I} - \mathbf{A}(\mathsf{p}))^{-1}\mathbf{B}(\mathsf{p})$ with $\mathsf{p} = \{\mathsf{p}_1, \mathsf{p}_1, \ldots, \mathsf{p}_{\mathrm{par}}\}$.

- Assume

$$\mathbf{A}(\mathsf{p}) = \mathbf{A}_0 + f_1(\mathsf{p})\mathbf{A}_1 + \ldots + f_M(\mathsf{p})\mathbf{A}_M$$
$$\mathbf{B}(\mathsf{p}) = \mathbf{B}_0 + g_1(\mathsf{p})\mathbf{B}_1 + \ldots + g_M(\mathsf{p})\mathbf{B}_M$$
$$\mathbf{C}(\mathsf{p}) = \mathbf{C}_0 + h_1(\mathsf{p})\mathbf{C}_1 + \ldots + g_M(\mathsf{p})\mathbf{C}_M$$

  with $M \ll n$.

- Want to preserve the parametric dependence in the reduced model in a way the maintains effective reduction: $\mathcal{H}_r(\mathsf{p}, s) = \mathbf{C}_r(\mathsf{p})\,(s\mathbf{I} - \mathbf{A}_r(\mathsf{p}))^{-1}\mathbf{B}_r(\mathsf{p})$ with

$$\mathbf{A}_r(\mathsf{p}) = \mathbf{W}_r^T\mathbf{A}(\mathsf{p})\mathbf{V}_r = \mathbf{W}_r^T\mathbf{A}_0\mathbf{V}_r + f_1(\mathsf{p})\mathbf{W}_r^T\mathbf{A}_1\mathbf{V}_r + \ldots + f_M(\mathsf{p})\mathbf{W}_r^T\mathbf{A}_M\mathbf{V}_r$$
$$\mathbf{B}_r(\mathsf{p}) = \mathbf{W}_r^T\mathbf{B}(\mathsf{p}) = \mathbf{W}_r^T\mathbf{B}_0 + g_1(\mathsf{p})\mathbf{W}_r^T\mathbf{B}_1 + \ldots + g_M(\mathsf{p})\mathbf{W}_r^T\mathbf{B}_M$$
$$\mathbf{C}_r(\mathsf{p}) = \mathbf{C}(\mathsf{p})\mathbf{V}_r = \mathbf{C}_0\mathbf{V}_r + h_1(\mathsf{p})\mathbf{C}_1\mathbf{V}_r + \ldots + h_M(\mathsf{p})\mathbf{C}_M\mathbf{V}_r$$

The parametric structure of $\mathcal{H}(\mathsf{p}, s)$ is retained in $\mathcal{H}_r(\mathsf{p}, s)$.

Parametrized Dynamical Systems

$\mathcal{H}(\mathsf{p}, s) = \mathbf{C}(\mathsf{p})\,(s\mathbf{I} - \mathbf{A}(\mathsf{p}))^{-1}\mathbf{B}(\mathsf{p})$ with $\mathsf{p} = \{\mathsf{p}_1, \mathsf{p}_1, \ldots, \mathsf{p}_{\mathrm{par}}\}$.

- Assume

$$\begin{aligned}
\mathbf{A}(\mathsf{p}) &= \mathbf{A}_0 + f_1(\mathsf{p})\mathbf{A}_1 + \ldots + f_M(\mathsf{p})\mathbf{A}_M \\
\mathbf{B}(\mathsf{p}) &= \mathbf{B}_0 + g_1(\mathsf{p})\mathbf{B}_1 + \ldots + g_M(\mathsf{p})\mathbf{B}_M \\
\mathbf{C}(\mathsf{p}) &= \mathbf{C}_0 + h_1(\mathsf{p})\mathbf{C}_1 + \ldots + g_M(\mathsf{p})\mathbf{C}_M
\end{aligned}$$

  with $M \ll n$.

- Want to preserve the parametric dependence in the reduced model in a way the maintains effective reduction: $\mathcal{H}_r(\mathsf{p}, s) = \mathbf{C}_r(\mathsf{p})\,(s\mathbf{I} - \mathbf{A}_r(\mathsf{p}))^{-1}\mathbf{B}_r(\mathsf{p})$ with

$$\begin{aligned}
\mathbf{A}_r(\mathsf{p}) &= \mathbf{W}_r^T\mathbf{A}(\mathsf{p})\mathbf{V}_r = \mathbf{W}_r^T\mathbf{A}_0\mathbf{V}_r + f_1(\mathsf{p})\mathbf{W}_r^T\mathbf{A}_1\mathbf{V}_r + \ldots + f_M(\mathsf{p})\mathbf{W}_r^T\mathbf{A}_M\mathbf{V}_r \\
\mathbf{B}_r(\mathsf{p}) &= \mathbf{W}_r^T\mathbf{B}(\mathsf{p}) = \mathbf{W}_r^T\mathbf{B}_0 + g_1(\mathsf{p})\mathbf{W}_r^T\mathbf{B}_1 + \ldots + g_M(\mathsf{p})\mathbf{W}_r^T\mathbf{B}_M \\
\mathbf{C}_r(\mathsf{p}) &= \mathbf{C}(\mathsf{p})\mathbf{V}_r = \mathbf{C}_0\mathbf{V}_r + h_1(\mathsf{p})\mathbf{C}_1\mathbf{V}_r + \ldots + h_M(\mathsf{p})\mathbf{C}_M\mathbf{V}_r
\end{aligned}$$

The parametric structure of $\mathcal{H}(\mathsf{p}, s)$ is retained in $\mathcal{H}_r(\mathsf{p}, s)$.

Parametrized Dynamical Systems

$\mathcal{H}(\mathsf{p}, s) = \mathbf{C}(\mathsf{p})\,(s\mathbf{I} - \mathbf{A}(\mathsf{p}))^{-1}\mathbf{B}(\mathsf{p})$ with $\mathsf{p} = \{\mathsf{p}_1, \mathsf{p}_1, \ldots, \mathsf{p}_{\text{par}}\}$.

- Assume

$$\begin{aligned}
\mathbf{A}(\mathsf{p}) &= \mathbf{A}_0 + f_1(\mathsf{p})\mathbf{A}_1 + \ldots + f_M(\mathsf{p})\mathbf{A}_M \\
\mathbf{B}(\mathsf{p}) &= \mathbf{B}_0 + g_1(\mathsf{p})\mathbf{B}_1 + \ldots + g_M(\mathsf{p})\mathbf{B}_M \\
\mathbf{C}(\mathsf{p}) &= \mathbf{C}_0 + h_1(\mathsf{p})\mathbf{C}_1 + \ldots + g_M(\mathsf{p})\mathbf{C}_M
\end{aligned}$$

  with $M \ll n$.

- Want to preserve the parametric dependence in the reduced model in a way the maintains effective reduction: $\mathcal{H}_r(\mathsf{p}, s) = \mathbf{C}_r(\mathsf{p})\,(s\mathbf{I} - \mathbf{A}_r(\mathsf{p}))^{-1}\mathbf{B}_r(\mathsf{p})$ with

$$\mathbf{A}_r(\mathsf{p}) = \mathbf{W}_r^T\mathbf{A}(\mathsf{p})\mathbf{V}_r = \mathbf{W}_r^T\mathbf{A}_0\mathbf{V}_r + f_1(\mathsf{p})\mathbf{W}_r^T\mathbf{A}_1\mathbf{V}_r + \ldots + f_M(\mathsf{p})\mathbf{W}_r^T\mathbf{A}_M\mathbf{V}_r$$

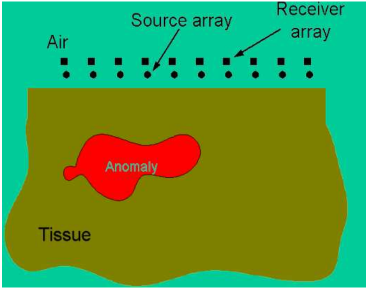
$$\mathbf{B}_r(\mathsf{p}) = \mathbf{W}_r^T\mathbf{B}(\mathsf{p}) = \mathbf{W}_r^T\mathbf{B}_0 + g_1(\mathsf{p})\mathbf{W}_r^T\mathbf{B}_1 + \ldots + g_M(\mathsf{p})\mathbf{W}_r^T\mathbf{B}_M$$

$$\mathbf{C}_r(\mathsf{p}) = \mathbf{C}(\mathsf{p})\mathbf{V}_r = \mathbf{C}_0\mathbf{V}_r + h_1(\mathsf{p})\mathbf{C}_1\mathbf{V}_r + \ldots + h_M(\mathsf{p})\mathbf{C}_M\mathbf{V}_r$$

The parametric structure of $\mathcal{H}(\mathsf{p}, s)$ is retained in $\mathcal{H}_r(\mathsf{p}, s)$.

Example 3: Diffuse Optical Tomography

- Tissue illuminated by near-infrared, frequency modulated light
- Light detected in array(s)
- Tumors have different optical properties than surrounding tissue
- Recover images of optical properties (diffusion and absorption fields) from data
- Problem is ill-posed and underdetermined, and data is noisy



Example 3: Diffuse Optical Tomography

- DOT forward problem given by the 3D PDE [Arridge 1999]

$$\frac{1}{\nu}\frac{\partial}{\partial t}\eta(x,t) = \nabla\cdot D(x)\nabla\eta(x,t) - \mu(x)\eta(x,t) + b_j(x)u_j(t), \quad \text{for } x\in\Omega$$

$$0 = \eta(x,t) + 2\,\mathcal{A}\,D(x)\,\frac{\partial}{\partial\xi}\eta(x,t), \quad \text{for } x\in\partial\Omega_{\pm}$$

$$y_i(t) = \int_{\partial\Omega} c_i(x)\eta(x,t)\,dx \quad \text{for } i=1,\,\ldots,\,n_d$$

- Utilize observations, $\mathbf{y}(t)$, to determine $D(x)$ and $\mu(x)$.
- For simplicity, assume $D(x)$ is well specified.
- $\mu(x) = \mu(\cdot,\mathrm{p})$ for $\mathrm{p} = [p_1,\ldots,p_\ell]^T$.
- Given n_s sources and n_f frequency modulations:
  a measurement for p: solution of $n_s\cdot n_f$ discretized 3D PDEs!

Discretized Problem

$$\mathbf{E}\,\dot{\mathbf{x}}(t;\mathsf{p}) = \mathbf{A}(\mathsf{p})\,\mathbf{x}(t;\mathsf{p}) + \mathbf{B}\,\mathbf{u}(t) \quad \text{with} \quad \mathbf{y}(t;\mathsf{p}) = \mathbf{C}\,\mathbf{x}(t;\mathsf{p})$$

- $\mathbf{E},\ \mathbf{A}(\mathbf{p}) \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times n_s}$, and $\mathbf{C} \in \mathbb{R}^{n_d\times n}$.
- $\mathbf{x} \in \mathbb{R}^n$ is the discretized photon flux, $\mathbf{A}(\mathsf{p}) = \mathbf{A}^{[0]} + \mathbf{A}^{[1]}(\mathsf{p})$
- $\mathbf{y} = [y_1, \ldots, y_{n_d}]^T$: vector of outputs.
- $\mathbf{Y}(\omega;\mathsf{p}) = \mathcal{F}(\mathbf{y}(t;\mathsf{p})),\ \mathbf{U}(\omega) = \mathcal{F}(\mathbf{u}(t))$

  $\mathbf{Y}(\omega;\mathsf{p}) = \mathcal{H}(\imath\omega;\mathsf{p})\ \mathbf{U}(\omega)$ where $\mathcal{H}(s;\mathsf{p}) = \mathbf{C}\,(s\,\mathbf{E} - \mathbf{A}(\mathsf{p}))^{-1}\,\mathbf{B}$;
- $\mathcal{H}(s,\mathsf{p})$: mapping from sources (inputs) to measurements (outputs) in the frequency domain.

Inverse Problem for Parameterized Tomography

- $\mathbf{Y}_i(\omega_j; \mathsf{p}) \in \mathbb{C}^{n_d}$: for input source $\mathbf{U}_i$ at frequency ω_j

$$\mathcal{Y}(\mathsf{p}) = [\mathbf{Y}_1(\omega_1; \mathsf{p})^T, \ldots, \mathbf{Y}_1(\omega_{n_\omega}; \mathsf{p})^T, \mathbf{Y}_2(\omega_1; \mathsf{p})^T, \ldots, \mathbf{Y}_{n_s}(\omega_{n_\omega}; \mathsf{p})^T]^T \in \mathbb{C}^{n_d \cdot n_s \cdot n_\omega}$$

- The nonlinear least squares problem:

$$\min_{\mathsf{p} \in \mathbb{R}^\ell} \|\mathcal{Y}(\mathsf{p}) - \mathbb{D}\|_2 \quad \text{s.t.}$$

$$\mathbf{E}\,\dot{\mathbf{x}}(t; \mathsf{p}) = \mathbf{A}(\mathsf{p})\,\mathbf{x}(t; \mathsf{p}) + \mathbf{B}\,\mathbf{u}(t) \quad \text{with} \quad \mathcal{Y}(\mathsf{p}) = \mathcal{F}\left(\mathbf{C}\,\mathbf{x}(t; \mathsf{p})\right)$$

- Parameterization by compactly supported radial basis functions (CSRBF) [Aghassi, Kilmer, Miller 2011]
- Solve using trust region method with regularized Gauss-Newton search directions [deSturler,Kilmer 2011]
- What are the objective function and Jacobian evaluations?

Forward Problem: Function and Jacobian Evaluations

- $\mathcal{Y}(\mathsf{p}) - \mathbb{D}$ eval. requires for $i = 1, \ldots, n_s$ and $j = 1, \ldots, n_w$

$$\mathcal{H}(\imath\omega_j; \mathsf{p}) = \mathbf{C}(\imath\omega_j \mathbf{E} - \mathbf{A}(\mathsf{p}))^{-1} \mathbf{B},$$

- Jacobian evaluation $\dfrac{\partial}{\partial p_k}\mathbf{Y}_i(\omega_j; \mathsf{p})$ requires

$$\frac{\partial}{\partial p_k}\left[\mathcal{H}(\imath\omega_j; \mathsf{p})\right] = -\mathbf{C}\left(\imath\omega_j \mathbf{E} - \mathbf{A}(\mathsf{p})\right)^{-1} \frac{\partial}{\partial p_k}\mathbf{A}(\mathsf{p}) \left(\imath\omega_j \mathbf{E} - \mathbf{A}(\mathsf{p})\right)^{-1} \mathbf{B}$$

for $i = 1, \ldots, n_d$ and $j = 1, \ldots, n_w$

- Use interpolatory model reduction to replace
  - $\mathcal{H}(\imath\omega_j; \mathsf{p}) = \mathbf{C}\left(\imath\omega_j \mathbf{E} - \mathbf{A}(\mathsf{p})\right)^{-1} \mathbf{B}$ with $\mathcal{H}_r(\imath\omega_j; \mathsf{p}) = \mathbf{C}_r(\imath\omega_j \mathbf{E}_r - \mathbf{A}_r(\mathsf{p}))^{-1} \mathbf{B}_r$
  - $\dfrac{\partial}{\partial p_k}\left[\mathcal{H}(\imath\omega_j; \mathsf{p})\right]$ with $\dfrac{\partial}{\partial p_k}\left[\mathcal{H}_r(\imath\omega_j; \mathsf{p})\right]$

Parametric Model Order Reduction

- Given

$$\mathcal{H}(s, \mathsf{p}) = \mathbf{C}(\mathsf{p}) \left(s\mathbf{E}(\mathsf{p}) - \mathbf{A}(\mathsf{p})\right)^{-1}\mathbf{B}(\mathsf{p})$$

- Construct

$$\mathcal{H}_r(\mathsf{p}, s) = \mathbf{C}_r(\mathsf{p}) \left(s\mathbf{E}(\mathsf{p}) - \mathbf{A}_r(\mathsf{p})\right)^{-1}\mathbf{B}_r(\mathsf{p})$$

via projection

$$\mathbf{E}_r(\mathsf{p}) = \mathbf{W}_r^T\mathbf{E}(\mathsf{p})\mathbf{V}_r$$

$$\mathbf{A}_r(\mathsf{p}) = \mathbf{W}_r^T\mathbf{A}(\mathsf{p})\mathbf{V}_r$$

$$\mathbf{B}_r(\mathsf{p}) = \mathbf{W}_r^T\mathbf{B}(\mathsf{p})$$

$$\mathbf{C}_r(\mathsf{p}) = \mathbf{C}(\mathsf{p})\mathbf{V}_r$$

Parameter interpolation

Theorem ([Baur/Beattie/Benner/G.,09])

Suppose $\sigma\mathbf{E}(\mathsf{p}) - \mathbf{A}(\mathsf{p})$, $\mathbf{B}(\mathsf{p})$, *and* $\mathbf{C}(\mathsf{p})$ *are continuously differentiable with respect to* p *in a neighborhood of* $\pi \in \mathbb{R}^{\ell}$, *where* $\sigma \in \mathbb{C}$.

- *if* $[\sigma\mathbf{E}(\pi) - \mathbf{A}(\pi)]^{-1}\mathbf{B}(\pi) \in$ *Range*$(\mathbf{V}_r)$ *and*

  $\left[\mathbf{C}(\pi)\,(\sigma\mathbf{E}(\pi) - \mathbf{A}(\pi))^{-1}\right]^T \in$ *Range*$(\mathbf{W}_r)$ *then*

$$\mathcal{H}(\sigma,\pi) = \mathcal{H}_r(\sigma,\pi), \quad \frac{\partial}{\partial s}\mathcal{H}(\sigma,\pi) = \frac{\partial}{\partial s}\mathcal{H}_r(\sigma,\pi), \text{ and}$$

$$\nabla_{\mathsf{p}}\mathcal{H}(\sigma,\pi) = \nabla_{\mathsf{p}}\mathcal{H}_r(\sigma,\pi)$$

- Two-sided interpolatory projection automatically matches parameter gradients.
- [Daniel *et al.*, 2004], [Gunupudi *et al.*, 2004], [Weile *et al.*, 1999], [Feng/Benner, 2009],....

Interpolatory Parametric Model Reduction in DOT

- Recall:
  - $\mathcal{H}(s,\mathbf{p}) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A}(\mathbf{p})\right)^{-1}\mathbf{B}$
  - $\mathcal{H}_r(s,\mathbf{p}) = \mathbf{C}_r\left(s\mathbf{E}_r - \mathbf{A}_r(\mathbf{p})\right)^{-1}\mathbf{B}_r$
- Choose frequency interpolation points $\sigma_1, \sigma_2, \ldots, \sigma_K \in \mathbb{C}$ and the parameter interpolation points $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_J \in \mathbb{R}^{\ell}$ to enforce

$$\begin{aligned} \mathcal{H}(\sigma_k, \boldsymbol{\pi}_j) &= \mathcal{H}_r(\sigma_k, \boldsymbol{\pi}_j) \\ \mathcal{H}'(\sigma_k, \boldsymbol{\pi}_j) &= \mathcal{H}_r'(\sigma_k, \boldsymbol{\pi}_j) \\ \nabla_{\mathsf{p}}\mathcal{H}(\sigma_k, \boldsymbol{\pi}_j) &= \nabla_{\mathsf{p}}\mathcal{H}_r(\sigma_k, \boldsymbol{\pi}_j) \end{aligned}$$

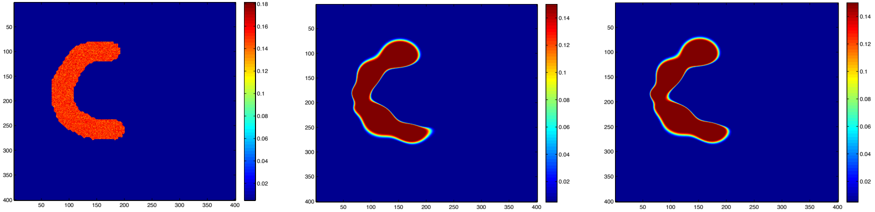
  for $k = 1, \ldots, K$ and $j = 1, \ldots, J$.
- In DOT application: $\sigma_k = \imath\omega_k$.

- In DOT, function evaluations amount to evaluating of $\mathcal{H}(s, \mathbf{p})$ for chosen $\sigma_k = \imath\omega_k$.
- The Jacobian evaluations are $\nabla_{\mathrm{p}}\mathcal{H}(s, \mathbf{p})$
- Perfect application for interpolatory model reduction.
- Replace $\mathcal{H}(s, \mathbf{p})$ with $\mathcal{H}_r(s, \mathbf{p})$ and $\nabla_{\mathrm{p}}\mathcal{H}(s, \mathbf{p})$ with $\nabla_{\mathrm{p}}\mathcal{H}_r(s, \mathbf{p})$
- Solving $r \times r$ linear systems as opposed to $n \times n$
- *For the values of $\mathbf{p}$ that are sampled, the minimization algorithm does not see the difference.*
- [Arian/Fahl/Sachs, 2002], [Fahl/Sachs, 2003], [Willcox *et al.*, 2010], [Druskin *et al.*, 2011], [Meerbergen,Yue 2011], [Benner/Sachs/Volkwein, 2014],..

Example 3a: $n = 160801$

- 5cm by 5cm uniformly spaced grid
- Discretization leading to $n = 160801$ degrees of freedom.
- There are 32 sources and detectors.
- 25 CSRBF leading to $\ell = 100$ parameters.
- Five sampling points $\boldsymbol{\pi}_j \in \mathbb{R}^{100}$
- Use same noise level and initialization for the full-order parametric model, $n = 160801$, and the surrogate parametric model, $r = 250$.

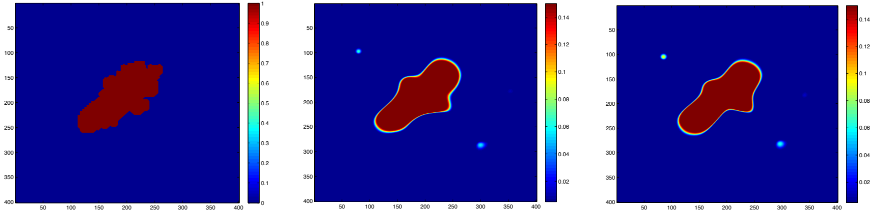
Example 3a - cont



- Full inversion problem: 1120 linear systems of size 160801×160801
- Reduced-inversion problem: 1216 linear systems of *size* 250×250
- Initial cost: 160 linear systems of size 160801×160801

Example 3b

- Use the same basis from the previous reconstruction



- Full inversion problem: 896 linear systems of size 160801×160801
- Reduced-inversion problem: 992 linear systems of *size* 250×250
- 0 linear systems of size 160801×160801

Inexact solves in interpolatory projections

- The (exact) primitive interpolating bases are

$$\mathbf{V}_r = \left[\mathcal{K}(\sigma_1)^{-1}\mathcal{B}(\sigma_1)\mathbf{b}_1,\ \mathcal{K}(\sigma_2)^{-1}\mathcal{B}(\sigma_2)\mathbf{b}_2,\ \cdots,\ \mathcal{K}(\sigma_r)^{-1}\mathcal{B}(\sigma_r)\mathbf{b}_r\ \right]$$

$$\mathbf{W}_r^T = \begin{bmatrix} \mathbf{c}_1^T\mathcal{C}(\sigma_1)\mathcal{K}(\sigma_1)^{-1} \\ \vdots \\ \mathbf{c}_r^T\mathcal{C}(\sigma_r)\mathcal{K}(\sigma_r)^{-1} \end{bmatrix}$$

- Persistent need for more detail and accuracy in the modeling stage makes n big: $\mathcal{O}(10^6)$ or more
- $\mathcal{K}(\sigma)\,\mathbf{v} = \mathcal{B}(\sigma)\mathbf{b}$ and $\mathcal{K}(\sigma)^T\,\mathbf{w} = \mathcal{C}(\sigma)^T\mathbf{c}$ cannot be solved directly.
- Inexact solves need to be used in constructing $\mathbf{V}_r$ and $\mathbf{W}_r$
- Inexact solves create new issues.

Reduced order models no longer interpolate $\mathcal{H}(s)$

Inexact solves in interpolatory projections

- Let $\widetilde{\mathbf{v}}_j$ be an inexact solution for $\mathcal{K}(\sigma_j)\mathbf{v} = \mathcal{B}(\sigma_j)\mathbf{b}_j$ and $\widetilde{\mathbf{w}}_j$ be an inexact solution for $\mathcal{K}(\sigma_j)^T\mathbf{w} = \mathcal{C}(\sigma_j)^T\mathbf{c}_j$.
- Inexact solutions are associated with residuals:

$$\delta\mathbf{b}_j = \mathcal{K}(\sigma_j)\widetilde{\mathbf{v}}_j - \mathcal{B}(\sigma_j)\mathbf{b}_j \qquad \delta\mathbf{c}_j = \mathcal{K}(\sigma_j)^T\widetilde{\mathbf{w}}_j - \mathcal{C}(\sigma_j)^T\mathbf{c}_j$$

- Define resulting "inexact bases"

$$\widetilde{\mathbf{V}}_r = [\widetilde{\mathbf{v}}_1, \widetilde{\mathbf{v}}_2, \ldots, \widetilde{\mathbf{v}}_r] \qquad \widetilde{\mathbf{W}}_r = [\widetilde{\mathbf{w}}_1, \widetilde{\mathbf{w}}_2, \ldots, \widetilde{\mathbf{w}}_r]$$

Inexact solves in interpolatory projections

- Let $\widetilde{\mathbf{v}}_j$ be an inexact solution for $\mathcal{K}(\sigma_j)\mathbf{v} = \mathcal{B}(\sigma_j)\mathbf{b}_j$ and $\widetilde{\mathbf{w}}_j$ be an inexact solution for $\mathcal{K}(\sigma_j)^T\mathbf{w} = \mathcal{C}(\sigma_j)^T\mathbf{c}_j$.
- Inexact solutions are associated with residuals:

$$\delta\mathbf{b}_j = \mathcal{K}(\sigma_j)\widetilde{\mathbf{v}}_j - \mathcal{B}(\sigma_j)\mathbf{b}_j \qquad \delta\mathbf{c}_j = \mathcal{K}(\sigma_j)^T\widetilde{\mathbf{w}}_j - \mathcal{C}(\sigma_j)^T\mathbf{c}_j$$

- Define resulting "inexact bases"

$$\widetilde{\mathbf{V}}_r = [\widetilde{\mathbf{v}}_1, \widetilde{\mathbf{v}}_2, \ldots, \widetilde{\mathbf{v}}_r] \qquad \widetilde{\mathbf{W}}_r = [\widetilde{\mathbf{w}}_1, \widetilde{\mathbf{w}}_2, \ldots, \widetilde{\mathbf{w}}_r]$$

The "inexact" model, $\tilde{\mathcal{H}}_r(s) = \tilde{\mathcal{C}}_r(s)\tilde{\mathcal{K}}_r(s)^{-1}\tilde{\mathcal{B}}_r(s)$, is defined by

$$\tilde{\mathcal{K}}_r(s) = \widetilde{\mathbf{W}}_r^T\mathcal{K}(s)\widetilde{\mathbf{V}}_r, \quad \tilde{\mathcal{B}}_r(s) = \widetilde{\mathbf{W}}_r^T\mathcal{B}(s), \quad \text{and} \quad \tilde{\mathcal{C}}_r(s) = \mathcal{C}(s)\widetilde{\mathbf{V}}_r.$$

Inexact solves with Petrov-Galerkin

- No unified backward error if approximate solution of each system $\mathcal{K}(\sigma_j)\mathbf{v} = \mathcal{B}(\sigma_j)\mathbf{b}_j$ and $\mathcal{K}(\sigma_j)^T\mathbf{w} = \mathcal{C}(\sigma_j)^T\mathbf{c}_j$ occurs independently. *Stronger conclusions possible if there is more structure.*
- Assume that the linear systems $\mathcal{K}(\sigma_j)\mathbf{v} = \mathcal{B}(\sigma_j)\mathbf{b}_j$ and $\mathcal{K}(\sigma_j)^T\mathbf{w} = \mathcal{C}(\sigma_j)^T\mathbf{c}_j$ are solved approximately with a

  Petrov-Galerkin process:
  Let $\mathcal{P}_m$ and $\mathcal{Q}_m$ be subspaces of $\mathbb{C}^n$ with $\mathcal{P}_m^\perp \cap \mathcal{Q}_m = \{0\}$.
  Let $\widetilde{\mathbf{v}}_j$ and $\widetilde{\mathbf{w}}_j$ be solutions of

$$\widetilde{\mathbf{v}}_j \in \mathcal{P}_m \quad \text{such that} \quad \mathcal{K}(\sigma_j)\mathbf{v} - \mathcal{B}(\sigma_j)\mathbf{b}_j \in \mathcal{Q}_m^\perp$$

  and

$$\widetilde{\mathbf{w}}_j \in \mathcal{Q}_m \quad \text{such that} \quad \mathcal{K}(\sigma_j)^T\mathbf{w} - \mathcal{C}(\sigma_j)^T\mathbf{c}_j \in \mathcal{P}_m^\perp$$

Inexact solves with Petrov-Galerkin

- No unified backward error if approximate solution of each system $\mathcal{K}(\sigma_j)\mathbf{v} = \mathcal{B}(\sigma_j)\mathbf{b}_j$ and $\mathcal{K}(\sigma_j)^T\mathbf{w} = \mathcal{C}(\sigma_j)^T\mathbf{c}_j$ occurs independently. *Stronger conclusions possible if there is more structure.*
- Assume that the linear systems $\mathcal{K}(\sigma_j)\mathbf{v} = \mathcal{B}(\sigma_j)\mathbf{b}_j$ and $\mathcal{K}(\sigma_j)^T\mathbf{w} = \mathcal{C}(\sigma_j)^T\mathbf{c}_j$ are solved approximately with a

  Petrov-Galerkin process:
  Let $\mathcal{P}_m$ and $\mathcal{Q}_m$ be subspaces of $\mathbb{C}^n$ with $\mathcal{P}_m^\perp \cap \mathcal{Q}_m = \{0\}$.
  Let $\widetilde{\mathbf{v}}_j$ and $\widetilde{\mathbf{w}}_j$ be solutions of

$$\widetilde{\mathbf{v}}_j \in \mathcal{P}_m \quad \text{such that} \quad \mathcal{K}(\sigma_j)\mathbf{v} - \mathcal{B}(\sigma_j)\mathbf{b}_j \in \mathcal{Q}_m^\perp$$

  and

$$\widetilde{\mathbf{w}}_j \in \mathcal{Q}_m \quad \text{such that} \quad \mathcal{K}(\sigma_j)^T\mathbf{w} - \mathcal{C}(\sigma_j)^T\mathbf{c}_j \in \mathcal{P}_m^\perp$$

Backward error with Petrov-Galerkin

- Define residual matrices

$$\mathbf{R_b} = [\delta\mathbf{b}_1,\ \delta\mathbf{b}_2,\ \ldots\delta\mathbf{b}_r] \qquad \mathbf{R_c} = [\delta\mathbf{c}_1,\ \delta\mathbf{c}_2,\ \ldots\delta\mathbf{c}_r]$$

Backward error with Petrov-Galerkin

- Define residual matrices

$$\mathbf{R_b} = [\delta \mathbf{b}_1,\ \delta \mathbf{b}_2,\ \ldots \delta \mathbf{b}_r] \qquad \mathbf{R_c} = [\delta \mathbf{c}_1,\ \delta \mathbf{c}_2,\ \ldots \delta \mathbf{c}_r]$$

and backward error

$$\mathbf{E}_{2r} = \mathbf{R_b}(\widetilde{\mathbf{W}}_r^T \widetilde{\mathbf{V}}_r)^{-1} \widetilde{\mathbf{W}}_r^T + \widetilde{\mathbf{V}}_r (\widetilde{\mathbf{W}}_r^T \widetilde{\mathbf{V}}_r)^{-1} \mathbf{R_c}^T$$

Backward error with Petrov-Galerkin

- Define residual matrices

$$\mathbf{R_b} = [\delta\mathbf{b}_1,\, \delta\mathbf{b}_2,\, \ldots \delta\mathbf{b}_r] \qquad \mathbf{R_c} = [\delta\mathbf{c}_1,\, \delta\mathbf{c}_2,\, \ldots \delta\mathbf{c}_r]$$

and backward error

$$\mathbf{E}_{2r} = \mathbf{R_b}(\widetilde{\mathbf{W}}_r^T\widetilde{\mathbf{V}}_r)^{-1}\widetilde{\mathbf{W}}_r^T + \widetilde{\mathbf{V}}_r(\widetilde{\mathbf{W}}_r^T\widetilde{\mathbf{V}}_r)^{-1}\mathbf{R_c}^T$$

then $\tilde{\mathcal{H}}_r(s)$ interpolates a perturbed dynamical system,

$\tilde{\mathcal{H}}(s) = \mathcal{C}(s)^T(\mathcal{K}(s) + \mathbf{E}_{2r})^{-1}\mathcal{B}(s)$ at $s = \sigma_1, \ldots \sigma_r$.

Backward error with Petrov-Galerkin

- Define residual matrices

$$\mathbf{R_b} = [\delta\mathbf{b}_1,\ \delta\mathbf{b}_2,\ \ldots\delta\mathbf{b}_r] \qquad \mathbf{R_c} = [\delta\mathbf{c}_1,\ \delta\mathbf{c}_2,\ \ldots\delta\mathbf{c}_r]$$

and backward error

$$\mathbf{E}_{2r} = \mathbf{R_b}(\widetilde{\mathbf{W}}_r^T\widetilde{\mathbf{V}}_r)^{-1}\widetilde{\mathbf{W}}_r^T + \widetilde{\mathbf{V}}_r(\widetilde{\mathbf{W}}_r^T\widetilde{\mathbf{V}}_r)^{-1}\mathbf{R_c}^T$$

then $\tilde{\mathcal{H}}_r(s)$ interpolates a perturbed dynamical system,

$$\tilde{\mathcal{H}}(s) = \mathcal{C}(s)^T(\mathcal{K}(s) + \mathbf{E}_{2r})^{-1}\mathcal{B}(s) \text{ at } s = \sigma_1, \ldots \sigma_r.$$

- The *computed* $\tilde{\mathcal{H}}_r(s)$ is an *exact* reduced order model of a perturbed system $\tilde{\mathcal{H}}(s)$ obtained by projection using "inexact" bases:

$$\tilde{\mathcal{K}}_r(s) = \widetilde{\mathbf{W}}_r^T\mathcal{K}(s)\widetilde{\mathbf{V}}_r = \widetilde{\mathbf{W}}_r^T(\mathcal{K}(s) + \mathbf{E}_{2r})\widetilde{\mathbf{V}}_r$$

Conclusions

- Useful distinction between model order and state space dimension.
- Interpolatory methods allow for straightforward extension to general system structures that reflect important underlying model features.
  - Optimal choices for interpolation points are no longer straightforward, but good choices are usually easy to obtain (for nonparametric problems).
  - For parameterized problems, effective strategies for choosing interpolation points rely on greedy selection (similar to best practices for RB methods).
- Example from tomographic image reconstruction
- As for standard interpolatory methods, the principal off-line cost is tied to solving large (generally sparse) linear algebraic systems.
  - For truncated iterative methods are used, backward stability is guaranteed within a Petrov-Galerkin framework.
  - Necessary step for well-grounded, rigorous termination criteria.

Valentine's Day is just around the corner !!

Great Gift Idea !!