Model Reduction via Interpolation

Thanos Antoulas
ECE Department, Rice University & MPI-Magdeburg

Christopher Beattie
Department of Mathematics, Virginia Tech.

Serkan Gugercin
Department of Mathematics, Virginia Tech.

LMS Durham Symposium on Model Order Reduction,
7 - 17 August 2017, Durham, UK

Outline

Lecture 1: (Beattie)

a. Linear (time-invariant, nonparametric) case: $\begin{cases} \mathbf{E}\,\dot{\mathbf{x}}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t), \\ \mathbf{y}(t) = \mathbf{C}\,\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{cases}$

- Rational Krylov subspaces
- Tangential interpolation

b. The Loewner Framework: Nonintrusive model reduction directly from observations of system response without access to $\mathbf{E}, \mathbf{A}, \mathbf{B}, \mathbf{C}$.

c. Reducing structured dynamical systems

Lecture 2: (Beattie)

- More on structure-preserving model reduction
- *Optimal* model reduction by interpolation and IRKA

Lecture 3: (Antoulas)

- Data-driven interpolatory methods for nonlinear systems
- Chef's surprize

Linear Dynamical Systems

$$\mathcal{S}: \qquad \mathbf{u}(t) \longrightarrow \boxed{\begin{aligned} \mathbf{E}\,\dot{\mathbf{x}}(t) &= \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\,\mathbf{x}(t) + \mathbf{D}\,\mathbf{u}(t) \end{aligned}} \longrightarrow \mathbf{y}(t)$$

- $\mathbf{A}, \mathbf{E} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\mathbf{C} \in \mathbb{R}^{q\times n}$ and $\mathbf{D} \in \mathbb{R}^{q\times m}$
- $\mathbf{x}(t) \in \mathbb{R}^n$: states, $\mathbf{u}(t) \in \mathbb{R}^m$: Input, $\mathbf{y}(t) \in \mathbb{R}^q$: Output
- We will assume $\lambda_i(\mathbf{A}, \mathbf{E}) \in \mathbb{C}_-$ for $i = 1, 2, \ldots, n$
- State-space dimension, n, is quite large, $n \approx \mathcal{O}(10^4, 10^7)$ or higher
- What is important is the mapping "$u \mapsto y$",
  NOT full information on state evolution: $\mathbf{x}(t)$
  $\Longrightarrow$ Remove unimportant states having small impact on $\mathbf{y}(t)$

- Produce a smaller dynamical system

$$\mathcal{S}_r : \mathbf{u}(t) \longrightarrow \boxed{\begin{aligned} \mathbf{E}_r\,\dot{\mathbf{x}}_r(t) &= \mathbf{A}_r\,\mathbf{x}_r(t) + \mathbf{B}_r\,\mathbf{u}(t) \\ \mathbf{y}_r(t) &= \mathbf{C}_r\,\mathbf{x}_r(t) + \mathbf{D}_r\,\mathbf{u}(t) \end{aligned}} \longrightarrow \mathbf{y}_r(t) \approx \mathbf{y}(t)$$

  where $\mathbf{A}_r, \mathbf{E}_r \in \mathbb{R}^{r\times r}$, $\mathbf{B}_r \in \mathbb{R}^{r\times m}$, $\mathbf{C}_r \in \mathbb{R}^{q\times r}$, and $\mathbf{D}_r \in \mathbb{R}^{q\times m}$ such that

  - r-dimensional state space with $r \ll n$;
  - $\|\mathbf{y} - \mathbf{y}_r\|$ is *small* wrt an appropriate norm;
  - important structural properties of $\mathcal{S}$ are preserved;
  - the procedure is *computationally efficient*.

- "Project dynamics" onto an r-dimensional subspace;
- Eliminate states that:
  - are insensitive to variations in $\mathbf{u}(t)$: "Hard to reach"
  - have little influence on $\mathbf{y}(t)$: "Hard to observe"

- $\mathcal{S}_r$ then used as a surrogate for the original model.

- Produce a smaller dynamical system

$$\mathcal{S}_r : \mathbf{u}(t) \longrightarrow \boxed{\begin{aligned} \mathbf{E}_r\, \dot{\mathbf{x}}_r(t) &= \mathbf{A}_r\, \mathbf{x}_r(t) + \mathbf{B}_r\, \mathbf{u}(t) \\ \mathbf{y}_r(t) &= \mathbf{C}_r\, \mathbf{x}_r(t) + \mathbf{D}_r\, \mathbf{u}(t) \end{aligned}} \longrightarrow \mathbf{y}_r(t) \approx \mathbf{y}(t)$$

  where $\mathbf{A}_r, \mathbf{E}_r \in \mathbb{R}^{r\times r}$, $\mathbf{B}_r \in \mathbb{R}^{r\times m}$, $\mathbf{C}_r \in \mathbb{R}^{q\times r}$, and $\mathbf{D}_r \in \mathbb{R}^{q\times m}$ such that

  - r-dimensional state space with $r \ll n$;
  - $\|\mathbf{y} - \mathbf{y}_r\|$ is *small* wrt an appropriate norm;
  - important structural properties of $\mathcal{S}$ are preserved;
  - the procedure is *computationally efficient*.

- "Project dynamics" onto an r-dimensional subspace;
- Eliminate states that:
  - are insensitive to variations in $\mathbf{u}(t)$: "Hard to reach"
  - have little influence on $\mathbf{y}(t)$: "Hard to observe"

- $\mathcal{S}_r$ then used as a surrogate for the original model.

Model Reduction via Projection

Choose

- $\mathcal{V}_r = \text{Range}(\mathbf{V}_r)$: the r-dimensional *right modeling subspace* (trial subspace) where $\mathbf{V}_r \in \mathbb{R}^{n\times r}$, and
- $\mathcal{W}_r = \text{Range}(\mathbf{W}_r)$, the r-dimensional *left modeling subspace* (test subspace) where $\mathbf{W}_r \in \mathbb{R}^{n\times r}$
- Approximate $\underbrace{\mathbf{x}(t)}_{n\times 1} \approx \underbrace{\mathbf{V}_r}_{n\times r}\underbrace{\mathbf{x}_r(t)}_{r\times 1}$ by forcing $\mathbf{x}_r(t)$ to satisfy

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

- Leads to a reduced order model:

$$\mathbf{E}_r = \underbrace{\mathbf{W}_r{}^T\mathbf{E}\mathbf{V}_r}_{r\times r},\quad \mathbf{A}_r = \underbrace{\mathbf{W}_r{}^T\mathbf{A}\mathbf{V}_r}_{r\times r},\quad \mathbf{B}_r = \underbrace{\mathbf{W}_r{}^T\mathbf{B}}_{r\times m},\quad \mathbf{C}_r = \underbrace{\mathbf{C}\mathbf{V}_r}_{q\times r},\quad \mathbf{D}_r = \underbrace{\mathbf{D}}_{q\times m}$$

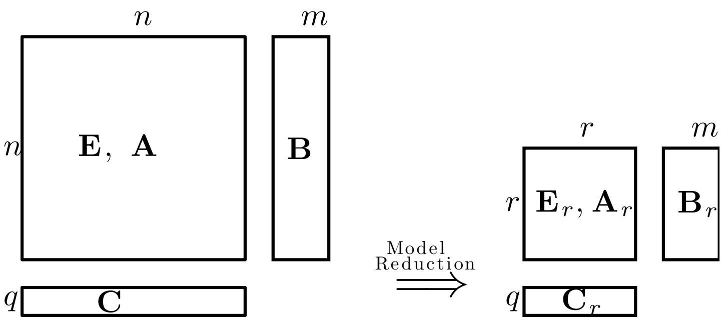


Figure: Projection-based Model Reduction

- Basis independence - Only $\mathcal{V}_r = \mathrm{Ran}(\mathbf{V}_r)$ and $\mathcal{W}_r = \mathrm{Ran}(\mathbf{W}_r)$ matters.
- Once $\mathcal{V}_r$ and $\mathcal{W}_r$ are selected, $\mathcal{S}_r$ is fully determined.

Transfer Functions and the Frequency Domain

- $\mathcal{S}: \quad \mathbf{u}(t) \mapsto \mathbf{y}(t) = (\mathcal{S}\mathbf{u})(t) = \int_{-\infty}^{t} h(t-\tau)\mathbf{u}(\tau)d\tau.$
- $\mathbf{H}(s) = (\mathcal{L}h)(s) = \int_{0}^{\infty} h(\tau)e^{-s\tau}d\tau = \mathbf{C}(s\mathbf{E}-\mathbf{A})^{-1}\mathbf{B} + \mathbf{D}.$
- $\mathbf{H}(s)$ is called the transfer function of $\mathcal{S}$.
- $\mathbf{H}(s)$: matrix-valued $(q \times p)$ rational function in $s \in \mathbb{C}$.
- Consider the simple $n = m = q = 2$ example with $\mathbf{D} = \mathbf{0}$,

$$\mathbf{E} = \mathbf{I}_2, \quad \mathbf{A} = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix},$$

- $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E}-\mathbf{A})^{-1}\mathbf{B} = \dfrac{1}{s^2+3s+2}\begin{bmatrix} 1 & s+4 \\ s-1 & -6 \end{bmatrix}$

- Let $\hat{\mathbf{z}}(\omega) = \mathcal{F}(\mathbf{z}(t))$

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

with transfer functions:

$$\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \quad \text{and} \quad \mathbf{H}_r(s) = \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r + \mathbf{D}_r$$

- $\mathbf{H}(s) = \dfrac{\alpha_0 s^n + \alpha_1 s^{n-1} + \alpha_2 s^{n-2} + \cdots + \alpha_n}{s^n + \beta_1 s^{n-1} + \beta_2 s^{n-2} + \cdots + \beta_n}$ (Assuming SISO)
- $\mathbf{H}_r(s) = \dfrac{\gamma_0 s^r + \gamma_1 s^{r-1} + \gamma_2 s^{r-2} + \cdots + \gamma_r}{s^r + \eta_1 s^{r-1} + \eta_2 s^{r-2} + \cdots + \eta_r}$ (Assuming SISO)
- Model Reduction = Rational Approximation

Error measures: $\mathcal{H}_\infty$ Norm

- $\mathcal{L}^2$ - $\mathcal{L}^2$ induced norm associated with $\mathcal{S} : \mathbf{u} \to \mathbf{y}$

$$\|\mathcal{S}\|_{\mathcal{H}_\infty} = \sup_{\mathbf{u}\neq 0} \frac{\|\mathbf{y}\|_2}{\|\mathbf{u}\|_2} = \sup_{\mathbf{u}\neq 0} \frac{\|\mathcal{S}\mathbf{u}\|_2}{\|\mathbf{u}\|_2} = \sup_{w\in\mathbb{R}} \|\mathbf{H}(\imath w)\|_2$$

- $\|\mathcal{S} - \mathcal{S}_r\|_{\mathcal{H}_\infty}$ is worst-case output error $\|\mathbf{y}(t) - \mathbf{y}_r(t)\|_2$ with $\|\mathbf{u}\|_2 = 1$.

$$\|\mathbf{y} - \mathbf{y}_r\|_2 \leq \|\mathcal{S} - \mathcal{S}_r\|_{\mathcal{H}_\infty} \|\mathbf{u}\|_2, \quad t \geq 0.$$

Suppose $\|\mathbf{u}\|_2 = 1$,

$$\begin{aligned}
\int_0^\infty \|\mathbf{y}(t) - \mathbf{y}_r(t)\|_2^2\, dt &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \|\,\widehat{\mathbf{y}}(\imath\omega) - \widehat{\mathbf{y}}_r(\imath\omega)\|_2^2 \; d\omega \\
&\leq \frac{1}{2\pi}\int_{-\infty}^{\infty} \|\mathbf{H}(\imath\omega) - \mathbf{H}_r(\imath\omega)\|_2^2 \; \|\widehat{\mathbf{u}}(\imath\omega)\|_2^2\, d\omega \\
&\leq \sup_\omega \|\mathbf{H}(\imath\omega) - \mathbf{H}_r(\imath\omega)\|_2^2 \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \|\widehat{\mathbf{u}}(\imath\omega)\|_2^2\, d\omega\right)^{1/2} \\
&\leq \sup_\omega \|\mathbf{H}(\imath\omega) - \mathbf{H}_r(\imath\omega)\|_2^2 \overset{\text{def}}{=} \|\mathcal{S} - \mathcal{S}_r\|_{\mathcal{H}_\infty}^2
\end{aligned}$$

Error measures: $\mathcal{H}_2$ Norm

- $\mathcal{L}_2$ norm of $\mathbf{h}(t)$ in time domain.

$$\|\mathcal{S}\|_{\mathcal{H}_2} = \left(\int_0^{\infty} \|h(t)\|_2^2 \, dt \right)^{\frac{1}{2}} = \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} \|\mathbf{H}(\imath\omega)\|_F^2 \, d\omega \right)^{\frac{1}{2}}$$

- $\mathcal{L}_2$-$\mathcal{L}_\infty$ induced norm of $\mathcal{S}$ for MISO and SIMO systems:

$$\|\mathcal{S}\|_{\mathcal{H}_2} = \sup_{\mathbf{u} \neq 0} \frac{\|\mathbf{y}\|_\infty}{\|\mathbf{u}\|_2} \quad \text{for MISO and SIMO systems}$$

- In the general case of MIMO systems:

$$\|\mathbf{y} - \mathbf{y}_r\|_{L_\infty} \leq \|\mathcal{S} - \mathcal{S}_r\|_{\mathcal{H}_2} \, \|\mathbf{u}\|_{L_2}$$

Computing the $\mathcal{H}_2$ norm:

- In order for $\|\mathcal{S}\|_{\mathcal{H}_2} < \infty$, it's necessary that $\mathbf{D} = \mathbf{0}$.
- Given $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$, let $\mathbf{P}$ be the unique solution to

$$\mathbf{A}\mathbf{P}\mathbf{E}^T + \mathbf{E}\mathbf{P}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T = \mathbf{0}.$$

  Then,

$$\|\mathcal{S}\|_{\mathcal{H}_2} = \sqrt{\mathrm{Tr}(\mathbf{C}\,\mathbf{P}\,\mathbf{C}^T)}$$

- Directly follows from definition of $\mathcal{H}_2$ norm $+$ residue thm.
- Matlab commands: `norm(`$\mathcal{S}$`,2)`, `normh2(`$\mathcal{S}$`)`, `h2norm(`$\mathcal{S}$`)`,

Frequency Domain Plots

- System response described graphically in the frequency domain.
- Amplitude Bode Plot: Plot $\|\mathbf{H}(\imath\omega)\|_2$ vs $\omega \in \mathbb{R}$.
- For the dynamical system on Slide 8:

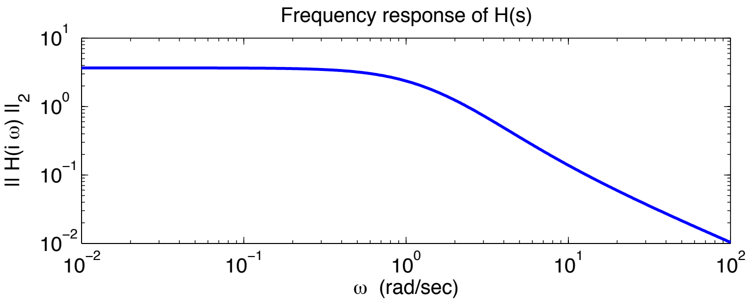


Figure: Frequency Response of $\mathbf{H}(s)$

Interpolatory Model Reduction

- Seek a reduced model $\mathcal{S}_r$ whose transfer function $\mathbf{H}_r(s)$ is a rational interpolant to $\mathbf{H}(s)$ in selected directions.

Tangential Interpolation Problem:

left interpolation points:
$\{\mu_i\}_{i=1}^r \subset \mathbb{C}$,
with corresponding
left tangent directions:
$\{\tilde{\mathsf{c}}_i\}_{i=1}^r \subset \mathbb{C}^q$,

and

right interpolation points:
$\{\sigma_i\}_{i=1}^r \subset \mathbb{C}$
with corresponding
right tangent directions:
$\{\tilde{\mathsf{b}}_i\}_{i=1}^r \subset \mathbb{C}^m$.

Find $\mathbf{E}_r$, $\mathbf{A}_r$, $\mathbf{B}_r$, $\mathbf{C}_r$, *and* $\mathbf{D}_r$ *(hence* $\mathbf{H}_r(s)$*) such that*

$$\tilde{\mathsf{c}}_i^T \mathbf{H}_r(\mu_i) = \tilde{\mathsf{c}}_i^T \mathbf{H}(\mu_i)$$
for $i = 1, \cdots, r$,

and

$$\mathbf{H}_r(\sigma_j)\tilde{\mathsf{b}}_j = \mathbf{H}(\sigma_j)\tilde{\mathsf{b}}_j,$$
for $j = 1, \cdots, r$,

- We are *not* requiring $\mathbf{H}_r(s)$ to (fully) interpolate $\mathbf{H}(s)$ at $s = \sigma$ i.e., we are not requiring full matrix interpolation: $\mathbf{H}(\sigma) = \mathbf{H}_r(\sigma)$ (this would result in $q \times m$ interpolation conditions at every interpolation point, $s = \sigma$).

- Instead, we are requiring $\mathbf{H}_r(s)$ to match $\mathbf{H}(s)$ at $s = \sigma$ only along a direction, b: $\mathbf{H}(\sigma)$b $= \mathbf{H}_r(\sigma)$b.

- This results in only m interpolation conditions at every interpolation point, $s = \sigma$.

- Later, we will see that this type of interpolation, *tangential interpolation*, is necessary for *optimal* model reduction.

Interpolatory Projections

- How to enforce tangential interpolation via projection?
- First case: $\mathbf{D} = \mathbf{D}_r$ (so wlog take $\mathbf{D} = \mathbf{D}_r = 0$).

Theorem

Let $\sigma,\ \mu \in \mathbb{C}$ *be such that* $s\,\mathbf{E} - \mathbf{A}$ *and* $s\,\mathbf{E}_r - \mathbf{A}_r$ *are invertible for* $s = \sigma,\ \mu$. *Assume* $\mathsf{b} \in \mathbb{C}^m$ *and* $\mathsf{c} \in \mathbb{C}^q$ *are nontrivial vectors.*

(a) *if* $(\sigma\,\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathsf{b} \in \mathrm{Ran}(\mathbf{V}_r)$, *then* $\mathbf{H}(\sigma)\mathsf{b} = \mathbf{H}_r(\sigma)\mathsf{b}$;

(b) *if* $\left(\mathsf{c}^T\mathbf{C}\,(\mu\,\mathbf{E} - \mathbf{A})^{-1}\right)^T \in \mathrm{Ran}(\mathbf{W}_r)$, *then* $\mathsf{c}^T\mathbf{H}(\mu) = \mathsf{c}^T\mathbf{H}_r(\mu)$;

(c) *and if both (a) and (b) hold, and* $\sigma = \mu$, *then*

$$\mathsf{c}^T\mathbf{H}'(\sigma)\mathsf{b} = \mathsf{c}^T\mathbf{H}_r'(\sigma)\mathsf{b} \quad \textit{as well.}$$

[Skelton *et. al.*, 87], [Grimme, 97], [Gallivan *et. al.*, 05]

Consequences:

- Given $\{\sigma_i\}_{i=1}^r$, $\{\mu_j\}_{j=1}^r$, $\{\mathsf{b}_i\}_{i=1}^r \in \mathbb{C}^m$, and $\{\mathsf{c}_j\}_{j=1}^r \in \mathbb{C}^q$, set

$$\mathbf{V}_r = \left[(\sigma_1\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\mathsf{b}_1,\ \cdots,\ (\sigma_r\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\mathsf{b}_r\right] \in \mathbb{C}^{n\times r} \text{ and}$$
$$\mathbf{W}_r = \left[(\mu_1\mathbf{E}-\mathbf{A}^T)^{-1}\mathbf{C}^T\mathsf{c}_1\ \cdots\ (\mu_r\mathbf{E}-\mathbf{A}^T)^{-1}\mathbf{C}^T\mathsf{c}_r\ \right] \in \mathbb{C}^{n\times r}$$

- Obtain $\mathbf{H}_r(s)$ via projection as before

$$\mathbf{E}_r = \mathbf{W}_r{}^T\mathbf{E}\mathbf{V}_r \quad \mathbf{A}_r = \mathbf{W}_r{}^T\mathbf{A}\mathbf{V}_r, \quad \mathbf{B}_r = \mathbf{W}_r{}^T\mathbf{B}, \quad \mathbf{C}_r = \mathbf{C}\mathbf{V}_r, \quad \mathbf{D}_r = \mathbf{D}$$

- Then

$$\begin{aligned} \mathbf{H}(\sigma_i)\mathsf{b}_i &= \mathbf{H}_r(\sigma_i)\mathsf{b}_i, && \text{for } i = 1,\cdots,r, \\ \mathsf{c}_j^T\mathbf{H}(\mu_j) &= \mathsf{c}_j^T\mathbf{H}_r(\mu_j), && \text{for } j = 1,\cdots,r, \\ \mathsf{c}_k^T\mathbf{H}'(\sigma_k)\mathsf{b}_k &= \mathsf{c}_k^T\mathbf{H}_r'(\sigma_k)\mathsf{b}_k && \text{if } \sigma_k = \mu_k \end{aligned}$$

bitangential Hermite interpolation where $\sigma_k = \mu_k$

Reduction from $n = 2$ to $r = 1$ (?!) ☺

- Recall the simple example $n = m = q = 2$ case with $\mathbf{D} = \mathbf{0}$,

$$\mathbf{E} = \mathbf{I}_2, \quad \mathbf{A} = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix},$$

- $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} = \dfrac{1}{s^2 + 3s + 2} \begin{bmatrix} 1 & s+4 \\ s-1 & -6 \end{bmatrix}$
- Let $\sigma_1 = \mu_1 = 0$, $\mathsf{b}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, and $\mathsf{c}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$,
- $\mathbf{V}_r = (\sigma_1 \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathsf{b}_1 = \begin{bmatrix} 1 \\ -1.5 \end{bmatrix}$
- $\mathbf{W}_r = (\sigma_1 \mathbf{E} - \mathbf{A})^{-T}\mathbf{C}^T\mathsf{c}_1 = \begin{bmatrix} -0.5 \\ -3.5 \end{bmatrix}$

- $\mathbf{E}_r = \mathbf{W}_r{}^T\mathbf{E}\mathbf{V}_r = 4.75, \qquad \mathbf{A}_r = \mathbf{W}_r{}^T\mathbf{A}\mathbf{V}_r = -3.5,$
- $\mathbf{B}_r = \mathbf{W}_r{}^T\mathbf{B} = \begin{bmatrix} -0.5 & -4 \end{bmatrix}, \qquad \mathbf{C}_r = \mathbf{C}\mathbf{V}_r = \begin{bmatrix} -1.5 \\ 2.5 \end{bmatrix},$
- $\mathbf{H}_r(s) = \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r = \frac{1}{s+0.7368}\begin{bmatrix} 0.1579 & 1.2630 \\ -0.2632 & -2.105 \end{bmatrix}$
- $\mathbf{H}(\sigma_1)\mathsf{b}_1 = \mathbf{H}_r(\sigma_1)\mathsf{b}_1 = \begin{bmatrix} -1.5 \\ 2.5 \end{bmatrix} \quad \checkmark$
- $\mathsf{c}_1^T\mathbf{H}(\sigma_1) = \mathsf{c}_1^T\mathbf{H}_r(\sigma_1) = [\,-0.5 \quad -4\,] \quad \checkmark$
- $\mathsf{c}_1^T\mathbf{H}'(\sigma_1)\mathsf{b}_1 = \mathsf{c}_1^T\mathbf{H}_r'(\sigma_1)\mathsf{b}_1 = 4.75 \quad \checkmark$

Interpolation Proof:

- Recall $\mathcal{V}_r = \text{Ran}(\mathbf{V}_r)$ and $\mathcal{W}_r = \text{Ran}(\mathbf{W}_r)$. Define

$$\mathcal{P}_r(z) = \mathbf{V}_r(z\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{W}_r^T(z\mathbf{E} - \mathbf{A}) \quad \text{and}$$

$$\mathcal{Q}_r(z) = (z\mathbf{E} - \mathbf{A})\mathbf{V}_r(z\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{W}_r^T = (z\mathbf{E} - \mathbf{A})\mathcal{P}_r(z)(z\mathbf{E} - \mathbf{A})^{-1}$$

- $\mathcal{P}_r^2(z) = \mathcal{P}_r(z)$ with $\mathcal{V}_r = \text{Ran}(\mathcal{P}_r(z)) = \text{Ker}\left(\mathbf{I} - \mathcal{P}_r(z)\right)$
- $\mathcal{Q}_r^2(z) = \mathcal{Q}_r(z)$ with $\mathcal{W}_r^\perp = \text{Ker}(\mathcal{Q}_r(z)) = \text{Ran}\left(\mathbf{I} - \mathcal{Q}_r(z)\right)$

$$\mathbf{H}(z) - \mathbf{H}_r(z) = \mathbf{C}(z\mathbf{E} - \mathbf{A})^{-1}\left(\mathbf{I} - \mathcal{Q}_r(z)\right)(z\mathbf{E} - \mathbf{A})\Big(\mathbf{I} - \mathcal{P}_r(z)\Big)(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$$

- Evaluate at $z = \sigma_i$ and postmultiply by $\mathbf{b}_i$: $\mathbf{H}(\sigma_i)\mathbf{b}_i = \mathbf{H}_r(\sigma_i)\mathbf{b}_i$
- Evaluate at $z = \sigma_i$ and premultiply by $\mathbf{c}^T$: $\mathbf{c}_i^T\mathbf{H}(\sigma_i) = \mathbf{c}_i^T\mathbf{H}_r(\sigma_i)$
- Evaluate at $z = \sigma + \varepsilon$, premultiply by $\mathbf{c}^T$ and postmultiply by $\mathbf{b}$:

$$\mathbf{c}_i^T\mathbf{H}(\sigma_i + \varepsilon)\mathbf{b}_i - \mathbf{c}_i^T\mathbf{H}_r(\sigma_i + \varepsilon)\mathbf{b}_i = \mathcal{O}(\varepsilon^2).$$

Since $\mathbf{c}_i^T\mathbf{H}(\sigma_i)\mathbf{b}_i = \mathbf{c}_i^T\mathbf{H}_r(\sigma_i)\mathbf{b}_i$,

$$\frac{1}{\varepsilon}\left(\mathbf{c}_i^T\mathbf{H}(\sigma_i + \varepsilon)\mathbf{b}_i - \mathbf{c}_i^T\mathbf{H}(\sigma_i)\mathbf{b}_i\right) - \frac{1}{\varepsilon}\left(\mathbf{c}_i^T\mathbf{H}_r(\sigma_i + \varepsilon)\mathbf{b}_i - \mathbf{c}_i^T\mathbf{H}_r(\sigma_i)\mathbf{b}_i\right) \to 0, \quad \text{as} \quad \varepsilon \to 0.$$

Higher-order Interpolation

Theorem

Let $\sigma \in \mathbb{C}$ be such that both $\sigma\mathbf{E} - \mathbf{A}$ and $\sigma\mathbf{E}_r - \mathbf{A}_r$ are invertible. If $\mathsf{b} \in \mathbb{C}^m$ *and* $\mathsf{c} \in \mathbb{C}^q$ *are fixed nontrivial vectors then*

(a) if $\left((\sigma\mathbf{E} - \mathbf{A})^{-1}\mathbf{E}\right)^{j-1}(\sigma\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathsf{b} \in \mathrm{Ran}(\mathbf{V}_r)$ *for* $j = 1, ., N$

then $\mathbf{H}^{(\ell)}(\sigma)\mathsf{b} = \mathbf{H}_r^{(\ell)}(\sigma)\mathsf{b}$ *for* $\ell = 0, 1, \ldots, N-1$

(b) if $\left((\mu\mathbf{E} - \mathbf{A})^{-T}\mathbf{E}^T\right)^{j-1}(\mu\mathbf{E} - \mathbf{A})^{-T}\mathbf{C}^T\mathsf{c} \in \mathrm{Ran}(\mathbf{W}_r)$ *for* $j = 1, ., M$,

then $\mathsf{c}^T\mathbf{H}^{(\ell)}(\mu) = \mathsf{c}^T\mathbf{H}_r^{(\ell)}(\mu)\mathsf{b}$ *for* $\ell = 0, 1, \ldots, M-1$;

(c) if both (a) and (b) hold, and if $\sigma = \mu$, *then* $\mathsf{c}^T\mathbf{H}^{(\ell)}(\sigma)\mathsf{b} = \mathsf{c}^T\mathbf{H}_r^{(\ell)}(\sigma)\mathsf{b}$,

for $\ell = 1, \ldots, M+N+1$

- The proof follows similarly.

Constructing interpolants with $\mathbf{D}_r \neq \mathbf{D}$

- For optimal $\mathcal{H}_\infty$ approximants, typically $\lim_{s\to\infty} \mathbf{H}_r(s) \neq \lim_{s\to\infty} \mathbf{H}(s)$

Theorem ([B/Gugercin,09] [Mayo/Antoulas,07])

Given $\{\mu_i\}_{i=1}^r \cup \{\sigma_j\}_{j=1}^r$, $\{\mathsf{c}_i\}_{i=1}^r \subset \mathbb{C}^q$ *and* $\{\mathsf{b}_j\}_{j=1}^r \subset \mathbb{C}^m$, *let* $\mathbf{V}_r \in \mathbb{C}^{n\times r}$ *and* $\mathbf{W}_r \in \mathbb{C}^{n\times r}$ *be as before. Define* $\widetilde{\mathsf{B}}$ *and* $\widetilde{\mathsf{C}}$ *as*

$$\widetilde{\mathsf{B}} = [\mathsf{b}_1,\, \mathsf{b}_2,\, \ldots,\, \mathsf{b}_r\,] \quad \textit{and} \quad \widetilde{\mathsf{C}}^T = [\mathsf{c}_1,\, \mathsf{c}_2,\, \ldots,\, \mathsf{c}_r]^T$$

For any $\mathbf{D}_r \in \mathbb{C}^{p\times m}$, *define*

$$\mathbf{E}_r(s) = \mathbf{W}_r^T\mathbf{E}\mathbf{V}_r, \qquad \mathbf{A}_r = \mathbf{W}_r^T\mathbf{A}\mathbf{V}_r + \widetilde{\mathsf{C}}^T\mathbf{D}_r\widetilde{\mathsf{B}},$$
$$\mathbf{B}_r = \mathbf{W}_r^T\mathbf{B} - \widetilde{\mathsf{C}}^T\mathbf{D}_r, \quad \text{and} \quad \mathbf{C}_r = \mathbf{C}\mathbf{V}_r - \mathbf{D}_r\widetilde{\mathsf{B}}.$$

Then with $\mathbf{H}_r(s) = \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r + \mathbf{D}_r$, *we have*

$$\mathbf{H}(\sigma_i)\mathsf{b}_i = \mathbf{H}_r(\sigma_i)\mathsf{b}_i \quad \textit{and} \quad \mathsf{c}_i^T\mathbf{H}(\mu_i) = \mathsf{c}_i^T\mathbf{H}_r(\mu_i) \quad \textit{for } i = 1,\, \ldots,\, r.$$

Interpolation from Data: Loewner Framework

- In some applications, dynamics are not available; but an abundant amount of input/output measurements are available.
- The goal: Construct a reduced-order model directly from data.

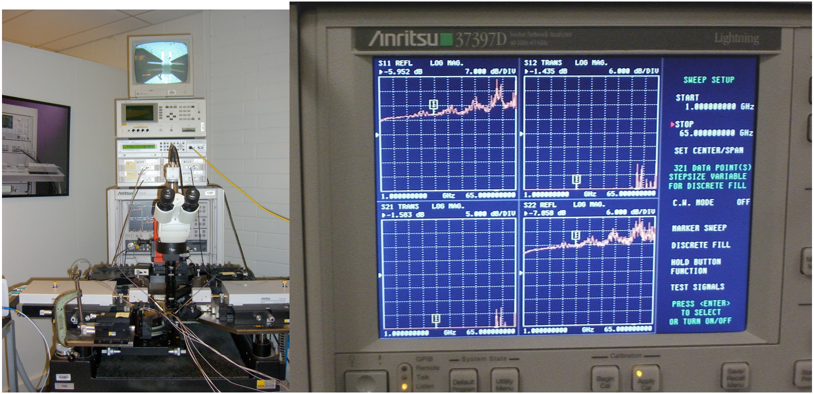


Figure: Vector Network Analyzer. (Data: A.C. Antoulas)

A more general problem setting

- Consider the following example ([Antoulas, 2005])

$$\frac{\partial T}{\partial t}(z,t) = \frac{\partial^2 T}{\partial z^2}(z,t), \quad t \geq 0, \quad z \in [0,1]$$

with the boundary conditions $\frac{\partial T}{\partial t}(0,t) = 0 \quad \text{and} \quad \frac{\partial T}{\partial z}(1,t) = u(t)$

- $u(t)$: supplied heat, $\quad y(t) = T(0,t)$
- Transfer function: $\quad \mathbf{H}(s) = \frac{Y(s)}{U(s)} = \frac{1}{\sqrt{s}\sinh\sqrt{s}} \neq \mathbf{C}(s\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}$
- New goal: Given the ability to evaluate $\mathbf{H}(s)$:

$$\boxed{\mathcal{H}(s)} \overset{?}{\approx} \boxed{\begin{array}{c} \mathbf{E}_r\dot{\mathbf{x}} = \mathbf{A}_r\mathbf{x}_r(t) + \mathbf{B}_r\mathbf{u}(t) \\ \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t) \end{array}}$$

Problem Set-up

- Given a set of input-output response measurements on $\mathbf{H}(s)$:

left driving frequencies:		*right driving frequencies*:
$\{\mu_i\}_{i=1}^r \subset \mathbb{C}$,		$\{\sigma_i\}_{i=1}^r \subset \mathbb{C}$
using *left input directions*:	and	using *right input directions*:
$\{\tilde{\mathsf{c}}_i\}_{i=1}^r \subset \mathbb{C}^q$,		$\{\tilde{\mathsf{b}}_i\}_{i=1}^r \subset \mathbb{C}^m$
producing *left responses*:		producing *right responses*:
$\{\tilde{\mathsf{z}}_i\}_{i=1}^r \subset \mathbb{C}^m$,		$\{\tilde{\mathsf{y}}_i\}_{i=1}^r \subset \mathbb{C}^q$

- Find a reduced model by determining (reduced) system matrices $\mathbf{E}_r$, $\mathbf{A}_r$, $\mathbf{B}_r$, $\mathbf{C}_r$, and $\mathbf{D}_r$ such that the associated transfer function, $\mathbf{H}_r(s)$ is a *tangential interpolant* to the given data:

$$\tilde{\mathsf{c}}_i^T \mathbf{H}_r(\mu_i) = \tilde{\mathsf{z}}_i^T \quad \text{for } i = 1, \cdots, r, \qquad \text{and} \qquad \mathbf{H}_r(\sigma_j)\tilde{\mathsf{b}}_j = \tilde{\mathsf{y}}_j, \quad \text{for } j = 1, \cdots, r,$$

Main Ingredients

- The *Loewner matrix*:

$$\mathbb{L} = \begin{bmatrix} \dfrac{\tilde{z}_1^T \tilde{b}_1 - \tilde{c}_1^T \tilde{y}_1}{\mu_1 - \sigma_1} & \cdots & \dfrac{\tilde{z}_1^T \tilde{b}_r - \tilde{c}_1^T \tilde{y}_r}{\mu_1 - \sigma_r} \\ \vdots & \ddots & \vdots \\ \dfrac{\tilde{z}_q^T \tilde{b}_1 - \tilde{c}_q^T \tilde{y}_1}{\mu_q - \sigma_1} & \cdots & \dfrac{\tilde{z}_q^T \tilde{b}_r - \tilde{c}_q^T \tilde{y}_r}{\mu_q - \sigma_r} \end{bmatrix} \in \mathbb{C}^{q \times r}.$$

- Suppose $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$:

$$\mathbb{L}_{ij} = \frac{\tilde{z}_i^T \tilde{b}_j - \tilde{c}_i^T \tilde{y}_j}{\mu_i - \sigma_j} = \frac{\tilde{c}_i^T [\mathbf{H}(\mu_i) - \mathbf{H}(\sigma_j)] \tilde{b}_j}{\mu_i - \sigma_j}$$

- What does $\mathbb{L}$ represent?

$$\widetilde{\mathsf{B}} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \tilde{\mathsf{b}}_1 & \tilde{\mathsf{b}}_2 & \dots & \tilde{\mathsf{b}}_r \\ \vdots & \vdots & & \vdots \end{bmatrix} \qquad \widetilde{\mathsf{Y}} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \tilde{\mathsf{y}}_1 & \tilde{\mathsf{y}}_2 & \dots & \tilde{\mathsf{y}}_r \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

$$\widetilde{\mathsf{Z}}^T = \begin{bmatrix} \dots & \tilde{\mathsf{z}}_1^T & \dots \\ \dots & \tilde{\mathsf{z}}_2^T & \dots \\ & \vdots & \\ \dots & \tilde{\mathsf{z}}_q^T & \dots \end{bmatrix} \qquad \widetilde{\mathsf{C}}^T = \begin{bmatrix} \dots & \tilde{\mathsf{c}}_1^T & \dots \\ \dots & \tilde{\mathsf{c}}_2^T & \dots \\ & \vdots & \\ \dots & \tilde{\mathsf{c}}_q^T & \dots \end{bmatrix}$$

Theorem (Mayo/Antoulas,2007)

The Loewner matrix $\mathbb{L}$ *satisfies the Sylvester equation*

$$\mathbb{L}\Sigma - M\mathbb{L} = \widetilde{\mathsf{C}}^T\widetilde{\mathsf{Y}} - \widetilde{\mathsf{Z}}^T\widetilde{\mathsf{B}},$$

where $\Sigma = \mathit{diag}(\sigma_1, \ \dots, \ \sigma_r) \in \mathbb{C}^{r\times r}$, *and* $M = \mathit{diag}(\mu_1, \ \dots, \ \mu_q) \in \mathbb{C}^{q\times q}$.

- Proof by direct substitution.

- The *shifted Loewner matrix*:

$$\mathbb{M} = \begin{bmatrix} \frac{\mu_1 \tilde{\mathsf{z}}_1^T \tilde{\mathsf{b}}_1 - \sigma_1 \tilde{\mathsf{c}}_1^T \tilde{\mathsf{y}}_1}{\mu_1 - \sigma_1} & \cdots & \frac{\mu_1 \tilde{\mathsf{z}}_1^T \tilde{\mathsf{b}}_r - \sigma_r \tilde{\mathsf{c}}_1^T \tilde{\mathsf{y}}_r}{\mu_1 - \sigma_r} \\ \vdots & \ddots & \vdots \\ \frac{\mu_q \tilde{\mathsf{z}}_q^T \tilde{\mathsf{b}}_1 - \sigma_1 \tilde{\mathsf{c}}_q^T \tilde{\mathsf{y}}_1}{\mu_q - \sigma_1} & \cdots & \frac{\mu_q \tilde{\mathsf{z}}_q^T \tilde{\mathsf{b}}_r - \sigma_r \tilde{\mathsf{c}}_q^T \tilde{\mathsf{y}}_r}{\mu_q - \sigma_r} \end{bmatrix} \in \mathbb{C}^{q \times r}$$

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

$$\mathbb{M}_{ij} = \frac{\tilde{\mathsf{c}}_i^T [\mu_i \mathbf{H}(\mu_i) - \sigma_j \mathbf{H}(\sigma_j)] \tilde{\mathsf{b}}_j}{\mu_i - \sigma_j}$$

- What does $\mathbb{M}$ represent?

Theorem (Mayo/Antoulas,2007)

$\mathbb{M}$ *satisfies the Sylvester equation*

$$\mathbb{M}\Sigma - M\mathbb{M} = \widetilde{\mathsf{C}}^T\widetilde{\mathsf{Y}}\Sigma - M\widetilde{\mathsf{Z}}^T\widetilde{\mathsf{B}}.$$

- Proof by direct substitution.

Theorem (Mayo/Antoulas,2007)

Assume that $\mu_i \neq \sigma_j$ *for all* $i, j = 1, \ldots, r$. *Suppose that* $\mathbb{M} - s\,\mathbb{L}$ *is invertible for all* $s \in \{\sigma_i\} \cup \{\mu_j\}$. *Then, with*

$$\mathbf{E}_r = -\mathbb{L}, \quad \mathbf{A}_r = -\mathbb{M}, \quad \mathbf{B}_r = \widetilde{\mathsf{Z}}^T, \quad \mathbf{C}_r = \widetilde{\mathsf{Y}}, \quad \mathbf{D}_r = 0,$$

$$\mathbf{H}_r(s) = \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r = \widetilde{\mathsf{Z}}^T(\mathbb{M} - s\,\mathbb{L})^{-1}\widetilde{\mathsf{Y}}$$

interpolates the data and furthermore is a minimal realization.

Sketch of the proof

- Assume $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$ (convenient but not necessary).
- $\mathbf{H}(\mu_i) - \mathbf{H}(\sigma_j) = (\sigma_j - \mu_i)\,\mathbf{C}(\mu_i\mathbf{E} - \mathbf{A})^{-1}\mathbf{E}(\sigma_j\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}.$

  $\Longrightarrow \mathbb{L} = -\mathbf{W}_r^T\mathbf{E}\mathbf{V}_r$ (resolvent identity !)
- $\mu_i\,\mathbf{H}(\mu_i) - \sigma_j\,\mathbf{H}(\sigma_j) = (\sigma_j - \mu_i)\,\mathbf{C}(\mu_i\mathbf{E} - \mathbf{A})^{-1}\mathbf{A}(\sigma_j\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}.$

  $\Longrightarrow \mathbb{M} = -\mathbf{W}_r^T\mathbf{A}\mathbf{V}_r$ (resolvent identity !)
- Also $\widetilde{\mathsf{Z}}^T = \mathbf{W}_r^T\mathbf{B}$ and $\widetilde{\mathsf{Y}} = \mathbf{C}\mathbf{V}_r$ by definition.

  $\Rightarrow$ $\mathbf{H}_r(s) = \widetilde{\mathsf{Y}}(\mathbb{M} - s\,\mathbb{L})^{-1}\widetilde{\mathsf{Z}}^T$ is a tangential interpolant to $\mathbf{H}(s)$.
- Proof without assuming $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$ uses the Sylvester equations.

Rank deficient case

- Assume

$$\mathrm{rank}\,(s\mathbb{L}-\mathbb{M}) = \mathrm{rank}\,[\mathbb{L}\ \mathbb{M}] = \mathrm{rank}\begin{bmatrix}\mathbb{L}\\ \mathbb{M}\end{bmatrix} \geq \rho, \text{ for all } s \in \{\sigma_i\}\cup\{\mu_j\}.$$

- Compute the SVD: $s\mathbb{L} - \mathbb{M} = \mathbf{Y}\Theta\mathbf{X}^*$, for some $s \in \{\sigma_i\} \cup \{\mu_j\}$

Theorem (Mayo/Antoulas,2007)

A realization $[\mathbf{E}_\rho, \mathbf{A}_\rho, \mathbf{B}_\rho, \mathbf{C}_\rho]$, *of a minimal solution is given as follows:*

$$\mathbf{E}_\rho = -\mathbf{Y}_\rho^*\mathbb{L}\mathbf{X}_\rho,\ \ \mathbf{A}_\rho = -\mathbf{Y}_\rho^*\mathbb{M}\mathbf{X}_\rho,\ \ \mathbf{B}_\rho = \mathbf{Y}_\rho^*\widetilde{\mathsf{Y}},\ \ \mathbf{C}_\rho = \widetilde{\mathsf{Z}}^T\mathbf{X}_\rho.$$

- Depending on whether ρ is the exact or approximate rank, either an interpolant or an approximate interpolant, respectively.

- There is no need for $\mathbf{H}(s)$ itself to be a finite-order rational function.

  All that is required is the ability of computing $\mathbf{H}(s)$ at any $s \in \mathbb{C}$; for example, $\mathbf{H}(s) = \dfrac{1}{\sqrt{s}\sinh\sqrt{s}}$ can be handled easily.

- Once data is collected, only a minimal amount of computation is necessary.

- For Hermite interpolation, choose $\sigma_i = \mu_i$ and then modify

$$\mathbb{L}_{ii} = \tilde{\mathsf{c}}_i \mathbf{H}'(\sigma_i)\tilde{\mathsf{b}}_i \;\text{ and }\; \mathbb{M}_{ii} = \tilde{\mathsf{c}}_i [s\mathbf{H}(s)]'_{s=\sigma_i}\tilde{\mathsf{b}}_i$$

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 may be reduced to an equivalent first order system"   Might not be the best approach ...
- For example

$$\mathbf{C}(s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}\mathbf{B} = \mathcal{C}(s\mathcal{E} - \mathcal{A})^{-1}\mathcal{B}$$

where

$$\mathcal{E} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{bmatrix}, \ \mathcal{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\mathbf{D} \end{bmatrix}, \ \mathcal{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B} \end{bmatrix}, \ \mathcal{C} = \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix}$$

- Disadvantages???

- 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; maintain structural relationships among the variables.

  “Structure-preserving model reduction”
- For the second-order systems, see: [Craig Jr.,1981], [Chahlaoui et.al, 2005], [Bai,2002], [Su/Craig,(1991)], [Meyer/Srinivasan,1996], ...
- We will be investigating a much more general framework.

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 $\mathbf{y}(t) = [\varpi(x_1,t),\,\ldots,\,\varpi(x_p,t)]^T$

- [Leitman and Fisher, 1973]
- $\mathbf{w}(x,t)$ is the displacement field; $\varpi(x,t)$ is the pressure field; $\rho(\tau)$ is a “relaxation function”

Example 1: Incompressible viscoelastic vibration

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

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

- [Leitman and Fisher, 1973]
- $\mathbf{w}(x,t)$ 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 !):

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

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

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

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

  with symmetric positive semidefinite $\mathbf{M}_r$, $\mathbf{K}_r \in \mathbb{R}^{r\times r}$, $\mathbf{B}_r \in \mathbb{R}^{r\times m}$, $\mathbf{C}_r \in \mathbb{R}^{p\times r}$, 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 !):

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

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

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

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

  with symmetric positive semidefinite $\mathbf{M}_r,\ \mathbf{K}_r \in \mathbb{R}^{r\times r}$, $\mathbf{B}_r \in \mathbb{R}^{r\times m}$, $\mathbf{C}_r \in \mathbb{R}^{p\times r}$, 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 !):

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

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

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

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

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

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

Checkpoint - Where are we ?

- Basic framework for interpolatory model reduction:
  - Rational Krylov spaces are natural projecting (test/trial) subspaces for canonical first-order realizations of SISO systems — but not for general (coprime) realizations or MIMO systems (tangential interpolation).
- Data-driven Interpolation - the Loewner framework
  - Reduced models are obtained directly from response measurements
- Importance of maintaining ancillary system structure
  - Foreshadowing of generalized coprime realizations for structure-preserving model reduction
- Open questions (so far)
  - Where do we interpolate ? . . . and in what directions ? ($\mathcal{H}_2$-optimal methods)
  - Extensions / Applications ? (e.g., DAEs, portHamiltonian/passive systems, bilinear/quadratic systems, parameterized systems, time-domain data-driven interpolation,...)

Outline

Lecture 1: (Beattie)

Linear (time-invariant, nonparametric) case: $\begin{cases} \mathbf{E}\,\dot{\mathbf{x}}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t), \\ \mathbf{y}(t) = \mathbf{C}\,\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{cases}$

Rational Krylov subspaces
Tangential interpolation

The Loewner Framework: Nonintrusive model reduction directly from observations of system response without access to $\mathbf{E}, \mathbf{A}, \mathbf{B}, \mathbf{C}$.

Reducing structured dynamical systems

Lecture 2: (Beattie)

a. More on structure-preserving model reduction
b. *Optimal* model reduction by interpolation and IRKA

Lecture 3: (Antoulas)

- Data-driven interpolatory methods for nonlinear systems
- Chef's surprize

Generalized Coprime Realizations

$$\mathbf{u}(t) \longrightarrow \boxed{\mathcal{H}(s) = \mathcal{C}(s)\mathcal{K}(s)^{-1}\mathcal{B}(s)} \longrightarrow \mathbf{y}(t)$$

- $\mathcal{C}(s) \in \mathbb{C}^{q\times n}$ and $\mathcal{B}(s) \in \mathbb{C}^{n\times m}$ are analytic in the right half plane;
- $\mathcal{K}(s) \in \mathbb{C}^{n\times n}$ is analytic and full rank throughout the right half plane with $n \approx 10^4 - 10^7$ or higher.
- "Internal state" $\mathbf{x}(t)$ is not itself important.
- How much state space detail is needed to replicate the map "$\mathbf{u} \mapsto \mathbf{y}$" ?

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

A General Projection Framework

- Select $\mathcal{V}_r \in \mathbb{R}^{n\times r}$ and $\mathcal{W}_r \in \mathbb{R}^{n\times r}$.
- The the reduced model $\mathcal{H}_r(s) = \mathcal{C}_r(s)\mathcal{K}_r(s)^{-1}\mathcal{B}_r(s)$ is

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

$$\mathbf{u}(t) \longrightarrow \boxed{\mathcal{H}_r(s) = \mathcal{C}_r(s)\mathcal{K}_r(s)^{-1}\mathcal{B}_r(s)} \longrightarrow \mathbf{y}_r(t) \approx \mathbf{y}(t)$$

- The generic case: $\mathcal{K}(s) = s\mathbf{E} - \mathbf{A}$, $\mathcal{B}(s) = \mathbf{B}$, $\mathcal{C}(s) = \mathbf{C}$,
- We choose $\mathcal{V}_r \in \mathbb{R}^{n\times r}$ and $\mathcal{W}_r \in \mathbb{R}^{n\times r}$ to enforce (tangential) interpolation.

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"
- (Spatial) discretization yields:

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

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

- $\mathbf{x}\in\mathbb{R}^{n_1}$ 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 $\mathbf{D}\in\mathbb{R}^{n_1\times n_2}$.

  We want a reduced model having the same "viscoelastic" structure.

Example 1: Incompressible viscoelastic vibration

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

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

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

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

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

  with symmetric positive semidefinite $\mathbf{M}_r$, $\mathbf{K}_r \in \mathbb{R}^{r\times r}$, $\mathbf{B}_r \in \mathbb{R}^{r\times m}$, $\mathbf{C}_r \in \mathbb{R}^{p\times r}$, and $\mathbf{D}_r \in \mathbb{R}^{r\times r}$.
- Because of the shared 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 processes that create a time lag from processing, communication, material transport, or inertial effects occuring at a finer scale than is explicitly modeled.

$$\dot{\mathbf{x}}(t) = \mathbf{A}_1\mathbf{x}(t) + \mathbf{A}_2\mathbf{x}(t-\tau) + \mathbf{B}\,\mathbf{u}(t), \qquad \mathbf{y}(t) \quad = \mathbf{C}\mathbf{x}(t)$$

$$\mathcal{H}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A}_1 - e^{-\tau s}\mathbf{A}_2)^{-1}\mathbf{B}.$$

- 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:

$$\dot{\mathbf{x}}_r(t) = \mathbf{A}_{1r}\mathbf{x}_r(t) + \mathbf{A}_{2r}\mathbf{x}_r(t-\tau) + \mathbf{B}_r\,\mathbf{u}(t), \qquad \mathbf{y}_r(t) \quad = \mathbf{C}_r\mathbf{x}_r(t)$$

$$\mathcal{H}_r(s) = \mathbf{C}_r(s\mathbf{I} - \mathbf{A}_{1r} - e^{-\tau s}\mathbf{A}_{2r})^{-1}\mathbf{B}_r$$

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 processes that create a time lag from processing, communication, material transport, or inertial effects occuring at a finer scale than is explicitly modeled.

$$\dot{\mathbf{x}}(t) = \mathbf{A}_1\mathbf{x}(t) + \mathbf{A}_2\mathbf{x}(t-\tau) + \mathbf{B}\,\mathbf{u}(t), \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$

$$\mathcal{H}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A}_1 - e^{-\tau s}\mathbf{A}_2)^{-1}\mathbf{B}.$$

- 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:

$$\dot{\mathbf{x}}_r(t) = \mathbf{A}_{1r}\mathbf{x}_r(t) + \mathbf{A}_{2r}\mathbf{x}_r(t-\tau) + \mathbf{B}_r\,\mathbf{u}(t), \qquad \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t)$$

$$\mathcal{H}_r(s) = \mathbf{C}_r(s\mathbf{I} - \mathbf{A}_{1r} - e^{-\tau s}\mathbf{A}_{2r})^{-1}\mathbf{B}_r$$

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 processes that create a time lag from processing, communication, material transport, or inertial effects occuring at a finer scale than is explicitly modeled.

$$\dot{\mathbf{x}}(t) = \mathbf{A}_1\mathbf{x}(t) + \mathbf{A}_2\mathbf{x}(t-\tau) + \mathbf{B}\,\mathbf{u}(t), \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$
$$\mathcal{H}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A}_1 - e^{-\tau s}\mathbf{A}_2)^{-1}\mathbf{B}.$$

- 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:

$$\dot{\mathbf{x}}_r(t) = \mathbf{A}_{1r}\mathbf{x}_r(t) + \mathbf{A}_{2r}\mathbf{x}_r(t-\tau) + \mathbf{B}_r\,\mathbf{u}(t), \qquad \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t)$$
$$\mathcal{H}_r(s) = \mathbf{C}_r(s\mathbf{I} - \mathbf{A}_{1r} - e^{-\tau s}\mathbf{A}_{2r})^{-1}\mathbf{B}_r$$

Interpolatory Projection for GenCoP Realizations

Galerkin/Petrov-Galerkin reduction preserves GenCoP structure.
Can we interpolate by specifying subspaces ?

- For selected points $\{\sigma_1, \sigma_2, \ldots \sigma_r\}$ in $\mathbb{C}$; and vectors $\{\mathsf{b}_1, \ldots \mathsf{b}_r\} \in \mathbb{C}^m$ and $\{\mathsf{c}_1, \ldots \mathsf{c}_r\} \in \mathbb{C}^q$, find $\mathcal{H}_r(s)$ so that

$$
\begin{aligned}
\mathsf{c}_i^T \mathcal{H}(\sigma_i) &= \mathsf{c}_i^T \mathcal{H}_r(\sigma_i) \\
\mathcal{H}(\sigma_i)\mathsf{b}_i &= \mathcal{H}_r(\sigma_i)\mathsf{b}_i, \text{ and} \\
\mathsf{c}_i^T \mathcal{H}'(\sigma_i)\mathsf{b}_i &= \mathcal{H}_r(\sigma_i)\mathsf{b}_i
\end{aligned}
$$

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

- Interpolation points: $\sigma_k \in \mathbb{C}$.
- Tangent directions: $\mathsf{c}_k \in \mathbb{C}^q$, and $\mathsf{b}_k \in \mathbb{C}^m$.

  Now what ?

Interpolatory Projection for GenCoP Realizations

Theorem (B/Gugercin,09)

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 $\mathsf{b} \in \mathbb{C}^p$ *and* $\mathsf{c} \in \mathbb{C}^q$ *are arbitrary nontrivial vectors.*

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

- Tangential interpolation via projection, as before.
- Proof follows similarly as the canonical first-order case.
- Shows the flexibility of the interpolation framework.

Interpolatory projections in model reduction

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

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

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

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

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

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

GenCoP Interpolation Proof (sketch):

- Recall $\mathcal{V}_r = \text{Ran}(\mathbf{V}_r)$ and $\mathcal{W}_r = \text{Ran}(\mathbf{W}_r)$. Define

$$\mathcal{P}_r(z) = \mathbf{V}_r \mathcal{K}_r(z)^{-1} \mathbf{W}_r^T \mathcal{K}(z) \quad \text{and}$$
$$\mathcal{Q}_r(z) = \mathcal{K}(z) \mathbf{V}_r \mathcal{K}_r(z)^{-1} \mathbf{W}_r^T = \mathcal{K}(z) \mathcal{P}_r(z) \mathcal{K}(z)^{-1}$$

- $\mathcal{P}_r^2(z) = \mathcal{P}_r(z)$ with $\mathcal{V}_r = \text{Ran}(\mathcal{P}_r(z)) = \text{Ker}\,(\mathbf{I} - \mathcal{P}_r(z))$
- $\mathcal{Q}_r^2(z) = \mathcal{Q}_r(z)$ with $\mathcal{W}_r^\perp = \text{Ker}(\mathcal{Q}_r(z)) = \text{Ran}\,(\mathbf{I} - \mathcal{Q}_r(z))$

  $\mathcal{H}(z) - \mathcal{H}_r(z) = \mathcal{C}(z)\mathcal{K}(z)^{-1}\,(\mathbf{I} - \mathcal{Q}_r(z))\,\mathcal{K}(z)\Big(\mathbf{I} - \mathcal{P}_r(z)\Big)\mathcal{K}(z)^{-1}\mathcal{B}(z)$

- Evaluate at $z = \sigma_i$ and postmultiply by $\mathbf{b}_i$: $\mathcal{H}(\sigma_i)\mathbf{b}_i = \mathcal{H}_r(\sigma_i)\mathbf{b}_i$
- Evaluate at $z = \sigma_i$ and premultiply by $\mathbf{c}^T$: $\mathbf{c}_i^T \mathcal{H}(\sigma_i) = \mathbf{c}_i^T \mathcal{H}_r(\sigma_i)$
- For Hermite condition, expand around $\sigma + \epsilon$ as before.

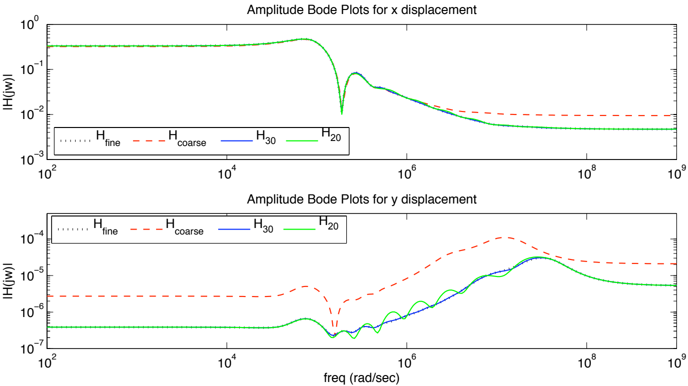
Incompressible Viscoelastic System Example

- 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 $\hat{x} = (0.5, 0.5)$.

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

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

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



$\mathcal{H}_{\text{fine}}$: $n_x = 51,842$ and $n_p = 6,651$ $\mathcal{H}_{30}$: $n_x = n_p = 30$
$\mathcal{H}_{\text{coarse}}$: $n_x = 13,122$ $n_p = 1,681$ $\mathcal{H}_{20}$: $n_x = n_p = 20$

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

Delay Example

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

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}_1\mathbf{x}(t) + \mathbf{A}_2\mathbf{x}(t-\tau) + \mathbf{B}\,\mathbf{u}(t), \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$

$$\mathcal{H}(s) = \underbrace{\mathbf{C}}_{\mathcal{C}(s)}\underbrace{\left(s\mathbf{E} - \mathbf{A}_1 - e^{-\tau s}\mathbf{A}_2\right)^{-1}}_{\mathcal{K}(s)}\underbrace{\mathbf{B}}_{\mathcal{B}(s)}.$$

- Find a reduced order model retaining the same delay structure:

$$\mathbf{E}_r\dot{\mathbf{x}}_r(t) = \mathbf{A}_{1r}\mathbf{x}_r(t) + \mathbf{A}_{2r}\mathbf{x}_r(t-\tau) + \mathbf{B}_r\,\mathbf{u}(t), \qquad \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t)$$

$$\mathcal{H}_r(s) = \underbrace{\mathbf{C}_r}_{\mathcal{C}_r(s)}\underbrace{\left(s\mathbf{E}_r - \mathbf{A}_{1r} - e^{-\tau s}\mathbf{A}_{2r}\right)^{-1}}_{\mathcal{K}_r(s)}\underbrace{\mathbf{B}_r}_{\mathcal{B}_r(s)}.$$

Comparison with other approaches

- Direct (generalized) interpolation:
  $\mathcal{H}_r(s) = \mathbf{e}^T\mathcal{V}_r\left(s\mathcal{W}_r^T\mathbf{E}\mathcal{V}_r - \mathcal{W}_r^T\mathbf{A}_1\mathcal{V}_r - \mathcal{W}_r^T\mathbf{A}_2\mathcal{V}_r\, e^{-s\tau}\right)^{-1}\mathcal{W}_r^T\mathbf{e}.$

- Approximate delay term with rational function:

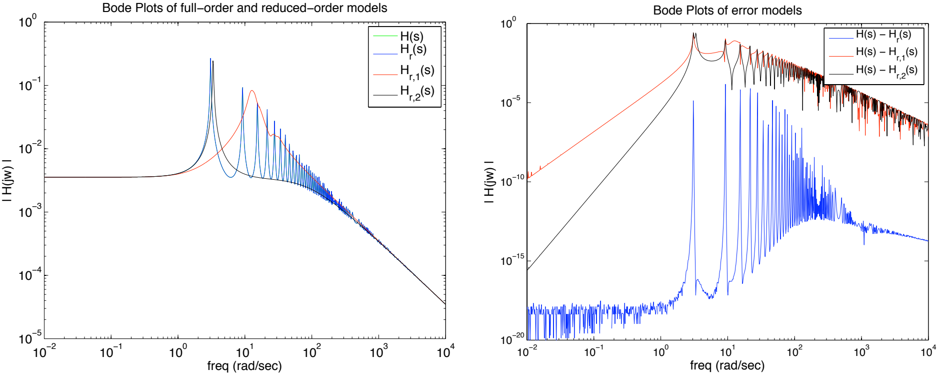
$$e^{-\tau s} \approx \frac{p_\ell(-\tau s)}{p_\ell(\tau s)}$$

  - 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!)

Second Example: Delay System

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

$\mathcal{H}_{r,2}(s)$ - Second-order Padé;



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

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

	$\mathcal{H}_\infty$ error
$\mathcal{H} - \mathcal{H}_r$	2.42×10^{-4}
$\mathcal{H} - \mathcal{H}_{r,1}$	2.65×10^{-1}
$\mathcal{H} - \mathcal{H}_{r,2}$	2.61×10^{-1}

- Consider $\mathcal{H}_{p,70}(s)$.
- $\|\mathcal{H}(s) - \mathcal{H}_{p,70}(s)\|_{\mathcal{H}_\infty} = 1.57 \times 10^{-3}$.
- Reducing $\mathcal{H}_{p,70}(s)$ requires solving linear systems of order $(500 \times 70) \times (500 \times 70)$.
- Preserving the delay structure is crucial.
- Multiple delays could also be handled similarly.

$\mathcal{H}_2$ Space

- $\mathcal{H}_2$: Set of matrix-valued functions, $\mathbf{H}(z)$, with components that are analytic for z in the open right half plane, $Re(z) > 0$, such that

$$\sup_{x>0} \int_{-\infty}^{\infty} \|\mathbf{H}(x+\imath y)\|_F^2 \, dy < \infty.$$

- $\mathcal{H}_2$ is a Hilbert space and transfer functions associated with stable finite dimensional dynamical systems are elements of $\mathcal{H}_2$.
- For stable $\mathbf{G}(s)$ and $\mathbf{H}(s)$ with the same m and q

$$\langle \mathbf{G},\ \mathbf{H}\rangle_{\mathcal{H}_2} \stackrel{\text{def}}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{Tr}\left(\overline{\mathbf{G}(\imath\omega)}\mathbf{H}(\imath\omega)^T\right)\, d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{Tr}\left(\mathbf{G}(-\imath\omega)\mathbf{H}(\imath\omega)^T\right) d\omega$$

- with a norm defined as

$$\|\mathbf{G}\|_{\mathcal{H}_2} \stackrel{\text{def}}{=} \left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} \|\mathbf{G}(\imath\omega)\|_F^2 \, d\omega\right)^{1/2}.$$

- For matrix-valued meromorphic functions, $\mathbf{F}(s)$,

$$\text{res}[\mathbf{F}(s),\lambda] = \lim_{s\to\lambda}(s-\lambda)\mathbf{F}(s) \;\; \text{has rank-}1 \text{ if } \lambda \text{ is a simple pole}$$

- We assume simple poles; the theory applies to the general case.
- Pole-residue expansion of $\mathbf{F}(s)$ of dimension-r:

$$\mathbf{F}(s) = \sum_{i=1}^{r} \frac{1}{s-\lambda_i}\mathsf{c}_i\mathsf{b}_i^T,$$

- where

$$\lambda_i \in \mathbb{C}_-, \quad \mathsf{c}_i \in \mathbb{C}^q, \quad \text{and} \quad \mathsf{b}_i \in \mathbb{C}^m \quad \text{for} \quad i = 1,\ldots,r.$$

Lemma

Suppose that $\mathbf{G}(s)$ *and* $\mathbf{H}(s) = \sum_{i=1}^{m} \frac{1}{s-\mu_i} \mathsf{c}_i \mathsf{b}_i^T$ *are real, stable and suppose that* $\mathbf{H}(s)$ *has simple poles at* $\mu_1, \mu_2, \ldots \mu_m$. *Then*

$$\langle \mathbf{G}, \mathbf{H} \rangle_{\mathcal{H}_2} = \sum_{k=1}^{m} \mathsf{c}_k^T \mathbf{G}(-\mu_k) \mathsf{b}_k$$

$$\text{and} \quad \|\mathbf{H}\|_{\mathcal{H}_2} = \left(\sum_{k=1}^{m} \mathsf{c}_k^T \mathbf{H}(-\mu_k) \mathsf{b}_k \right)^{1/2}.$$

- Proof: Application of the residue theorem:

$$\langle \mathbf{G}, \mathbf{H} \rangle_{\mathcal{H}_2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{Tr}\left(\mathbf{G}(-\imath\omega) \mathbf{H}(\imath\omega)^T \right) d\omega \quad = \lim_{R\to\infty} \frac{1}{2\pi\imath} \int_{\Gamma_R} \mathrm{Tr}\left(\mathbf{G}(-s) \mathbf{H}(s)^T \right) ds$$

- where

$$\Gamma_R = \{ z \,|\, z = \imath\omega \text{ with } \omega \in [-R, R] \} \cup \left\{ z \,\middle|\, z = R\, e^{\imath\theta} \text{ with } \theta \in [\frac{\pi}{2}, \frac{3\pi}{2}] \right\}.$$

Pole-residue based $\mathcal{H}_2$ error expression

Theorem

Given a full-order real system, $\mathbf{H}(s)$ *and a reduced model,* $\mathbf{H}_r(s)$*, having the form* $\mathbf{H}_r(s) = \sum_{i=1}^{r} \frac{1}{s-\hat{\lambda}_i} \mathsf{c}_i \mathsf{b}_i^T$ *(*$\mathbf{H}_r$ *has simple poles at* $\hat{\lambda}_1, \hat{\lambda}_2, \ldots \hat{\lambda}_r$ *and rank-1 residues* $\mathsf{c}_1\mathsf{b}_1^T, \ldots, \mathsf{c}_r\mathsf{b}_r^T$*.), the* $\mathcal{H}_2$ *norm of the error system is given by*

$$\|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_2}^2 = \|\mathbf{H}\|_{\mathcal{H}_2}^2 - 2\sum_{k=1}^{r} \mathsf{c}_k^T \mathbf{H}(-\hat{\lambda}_k)\mathsf{b}_k + \sum_{k,\ell=1}^{r} \frac{\mathsf{c}_k^T \mathsf{c}_\ell \, \mathsf{b}_\ell^T \mathsf{b}_k}{-\hat{\lambda}_k - \hat{\lambda}_\ell}$$

- SISO Case: [Krajewski et al.,1995], [Gugercin/Antoulas,2003]
- MIMO Case: [B./Gugercin,2008],
- Can be used in developing a trust region - descent $\mathcal{H}_2$ optimal model reduction algorithm [B./Gugercin,2009]

Optimal $\mathcal{H}_2$ approximation

Problem

Given $\mathbf{H}(s)$, *find* $\mathbf{H}_r(s)$ *of order* r *which solves:* $\min_{degree(\mathbf{G}_r)=r} \|\mathbf{H} - \mathbf{G}_r\|_{\mathcal{H}_2}$.

- The goal is to minimize (a bound for) $\max_{t \geq 0} \|\mathbf{y}(t) - \mathbf{y}_r(t)\|_\infty$ over all possible unit energy inputs.
- Non-convex optimization problem. Finding a global minimum is, at best, a formidable task.
- [Wilson,1970], [Hyland/Bernstein,1985]: Sylvester-equation based optimality conditions
- Wilson [1970]: Solution is obtained by projection.

  Is it interpolatory projection?

Interpolatory $\mathcal{H}_2$ optimality conditions

Theorem ([Gugercin/Antoulas/B.,08])

Given $\mathbf{H}(s)$, *let* $\mathbf{H}_r(s) = \sum_{i=1}^{r} \frac{1}{s-\hat{\lambda}_i} \hat{\mathsf{c}}_i \hat{\mathsf{b}}_i^T$ *be the best stable* r^{th} *order approximation of* $\mathbf{H}$ *with respect to the* $\mathcal{H}_2$ *norm. Assume* $\mathbf{H}_r$ *has simple poles at* $\hat{\lambda}_1, \hat{\lambda}_2, \ldots \hat{\lambda}_r$. *Then*

$$\mathbf{H}(-\hat{\lambda}_k)\hat{\mathsf{b}}_k = \mathbf{H}_r(-\hat{\lambda}_k)\hat{\mathsf{b}}_k, \quad \hat{\mathsf{c}}_k^T\mathbf{H}(-\hat{\lambda}_k) = \hat{\mathsf{c}}_k^T\mathbf{H}_r(-\hat{\lambda}_k),$$

$$\text{and} \quad \hat{\mathsf{c}}_k^T\mathbf{H}'(-\hat{\lambda}_k)\hat{\mathsf{b}}_k = \hat{\mathsf{c}}_k^T\mathbf{H}_r'(-\hat{\lambda}_k)\hat{\mathsf{b}}_k \quad \text{for } k = 1, 2, \ldots, r.$$

Interpolatory $\mathcal{H}_2$ optimality conditions

Theorem ([Gugercin/Antoulas/B.,08])

Given $\mathbf{H}(s)$, *let* $\mathbf{H}_r(s) = \sum_{i=1}^{r} \frac{1}{s-\hat{\lambda}_i}\hat{c}_i\hat{b}_i^T$ *be the best stable* r^{th} *order approximation of* $\mathbf{H}$ *with respect to the* $\mathcal{H}_2$ *norm. Assume* $\mathbf{H}_r$ *has simple poles at* $\hat{\lambda}_1, \hat{\lambda}_2, \ldots \hat{\lambda}_r$. *Then*

$$\mathbf{H}(-\hat{\lambda}_k)\hat{b}_k = \mathbf{H}_r(-\hat{\lambda}_k)\hat{b}_k, \quad \hat{c}_k^T\mathbf{H}(-\hat{\lambda}_k) = \hat{c}_k^T\mathbf{H}_r(-\hat{\lambda}_k),$$

$$\textit{and} \quad \hat{c}_k^T\mathbf{H}'(-\hat{\lambda}_k)\hat{b}_k = \hat{c}_k^T\mathbf{H}_r'(-\hat{\lambda}_k)\hat{b}_k \quad \textit{for } k = 1, 2, \ldots, r.$$

- Tangential Hermite interpolation for $\mathcal{H}_2$ optimality
- $\boxed{\text{Optimal interpolation points}: \ \sigma_i = -\hat{\lambda}_i}$
- The SISO conditions: [Meier /Luenberger,67]
- Other MIMO works: [van Dooren et al..08], [Bunse-Gernster et al.,09]

Proof:

- Let $\widetilde{\mathbf{H}}_r(s)$ be a stable r-th order dynamical system. Then,

$$\begin{aligned} \|\mathbf{H}-\mathbf{H}_r\|_{\mathcal{H}_2}^2 \leq\ & \|\mathbf{H}-\widetilde{\mathbf{H}}_r\|_{\mathcal{H}_2}^2 = \|\mathbf{H}-\mathbf{H}_r+\mathbf{H}_r-\widetilde{\mathbf{H}}_r\|_{\mathcal{H}_2}^2 \\ & = \|\mathbf{H}-\mathbf{H}_r\|_{\mathcal{H}_2}^2 + 2\,\Re e\,\langle \mathbf{H}-\mathbf{H}_r, \mathbf{H}_r-\widetilde{\mathbf{H}}_r\rangle_{\mathcal{H}_2} + \|\mathbf{H}_r-\widetilde{\mathbf{H}}_r\|_{\mathcal{H}_2}^2 \\ \text{so that} \quad & 0 \leq 2\,\Re e\,\langle \mathbf{H}-\mathbf{H}_r, \mathbf{H}_r-\widetilde{\mathbf{H}}_r\rangle_{\mathcal{H}_2} + \|\mathbf{H}_r-\widetilde{\mathbf{H}}_r\|_{\mathcal{H}_2}^2 \end{aligned}$$

- Choose $\widetilde{\mathbf{H}}_r(s)$ so that $\mathbf{H}_r(s)-\widetilde{\mathbf{H}}_r(s) = \dfrac{\varepsilon e^{\imath\theta}}{s-\hat{\lambda}_\ell}\xi \mathsf{b}_\ell^T$, $\boldsymbol{\xi}\in\mathbb{C}^q$: arbitrary

$$\Longrightarrow \quad \langle \mathbf{H}-\mathbf{H}_r, \mathbf{H}_r-\widetilde{\mathbf{H}}_r\rangle_{\mathcal{H}_2} = -\varepsilon\,|\boldsymbol{\xi}^T\left(\mathbf{H}(-\hat{\lambda}_\ell)-\mathbf{H}_r(-\hat{\lambda}_\ell)\right)\mathsf{b}_\ell|.$$

$$\Longrightarrow \quad 0 \leq |\boldsymbol{\xi}^T\left(\mathbf{H}(-\hat{\lambda}_\ell)-\mathbf{H}_r(-\hat{\lambda}_\ell)\right)\mathsf{b}_\ell| \leq \varepsilon\frac{\|\mathsf{b}_\ell\|_2^2}{-2\Re e(\hat{\lambda}_\ell)}$$

$$\Longrightarrow \quad \boldsymbol{\xi}^T\left(\mathbf{H}(-\hat{\lambda}_\ell)-\mathbf{H}_r(-\hat{\lambda}_\ell)\right)\mathsf{b}_\ell = 0$$

$$\Longrightarrow \quad \left(\mathbf{H}(-\hat{\lambda}_\ell)-\mathbf{H}_r(-\hat{\lambda}_\ell)\right)\mathsf{b}_\ell = 0.$$

- A similar arguments leads to left-tangential conditions.
- For the Hermite condition, choose $\widetilde{\mathbf{H}}_r(s)$ so that

$$\mathbf{H}_r(s) - \widetilde{\mathbf{H}}_r(s) = \left(\frac{1}{s - \hat{\lambda}_\ell} - \frac{1}{s - \mu} \right) \mathsf{c}_\ell \mathsf{b}_\ell^T.$$

- After various manipulations

$$0 \leq -2\varepsilon |\mathsf{c}_\ell^T \left(\mathbf{H}'(-\hat{\lambda}_\ell) - \mathbf{H}_r'(-\hat{\lambda}_\ell) \right) \mathsf{b}_\ell| + \mathcal{O}(\varepsilon^2).$$

- As $\varepsilon \to 0$, we obtain $|\mathsf{c}_\ell^T \left(\mathbf{H}'(-\hat{\lambda}_\ell) - \mathbf{H}_r'(-\hat{\lambda}_\ell) \right) \mathsf{b}_\ell| = 0.$
- Some analogous necessary $\mathcal{H}_2$-optimality conditions are known for GenCoP (B./Benner, 2014).
- $\hat{\lambda}_i, \hat{\mathsf{b}}_i, \hat{\mathsf{c}}_i$ NOT known a priori $\Longrightarrow$ Need iterative steps

An Iterative Rational Krylov Algorithm (IRKA):

Algorithm (Gugercin/Antoulas/B. [2008])

1. *Choose* $\{\sigma_1, \ldots, \sigma_r\}$, $\{\hat{\mathsf{b}}_1, \ldots, \hat{\mathsf{b}}_r\}$ *and* $\{\hat{\mathsf{c}}_1, \ldots, \hat{\mathsf{c}}_r\}$
2. $\mathbf{V}_r = \left[(\sigma_1 \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\hat{\mathsf{b}}_1 \;\cdots\; (\sigma_r \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\hat{\mathsf{b}}_r \right]$
   $\mathbf{W}_r = \left[(\sigma_1 \mathbf{E} - \mathbf{A}^T)^{-1}\mathbf{C}^T\hat{\mathsf{c}}_1 \;\cdots\; (\sigma_r \mathbf{E} - \mathbf{A}^T)^{-1}\mathbf{C}^T\hat{\mathsf{c}}_r \right]$.
3. *while (not converged)*
   1. $\mathbf{A}_r = \mathbf{W}_r^T\mathbf{A}\mathbf{V}_r$, $\mathbf{E}_r = \mathbf{W}_r^T\mathbf{E}\mathbf{V}_r$, $\mathbf{B}_r = \mathbf{W}_r^T\mathbf{B}$, *and* $\mathbf{C}_r = \mathbf{C}\mathbf{V}_r$
   2. *Compute* $\mathbf{H}_r(s) = \sum_{i=1}^{r} \frac{\hat{\mathsf{c}}_i\hat{\mathsf{b}}_i^T}{s - \hat{\lambda}_i}$, *and set* $\{\sigma_i\} \longleftarrow \{-\hat{\lambda}_i\}$,
   3. $\mathbf{V}_r = \left[(\sigma_1 \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\hat{\mathsf{b}}_1 \;\cdots\; (\sigma_r \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\hat{\mathsf{b}}_r \right]$
   4. $\mathbf{W}_r = \left[(\sigma_1 \mathbf{E} - \mathbf{A}^T)^{-1}\mathbf{C}^T\hat{\mathsf{c}}_1 \;\cdots\; (\sigma_r \mathbf{E} - \mathbf{A}^T)^{-1}\mathbf{C}^T\hat{\mathsf{c}}_r \right]$.
4. $\mathbf{A}_r = \mathbf{W}_r^T\mathbf{A}\mathbf{V}_r$, $\mathbf{E}_r = \mathbf{W}_r^T\mathbf{E}\mathbf{V}_r$, $\mathbf{B}_r = \mathbf{W}_r^T\mathbf{B}$, $\mathbf{C}_r = \mathbf{C}\mathbf{V}_r$, $\mathbf{D}_r = \mathbf{D}$.

- In its simplest form, IRKA is a fixed point iteration.
- IRKA is not a descent method and global convergence is not guaranteed despite overwhelming numerical evidence.
- Newton formulation is possible [Gugercin/Antoulas/B.,08]
- Guaranteed convergence for state-space symmetric systems [Flagg/B./Gugercin,2012]
- Globally convergent descent version: [B./Gugercin (2009)]
- Implementation with iterative solves:
  - w/ Krylov subspace recycling [Ahuja/deSturler/Gugercin/Chang (2010)]
  - w/ general iterative system solves [B/Gugercin/Wyatt (2010)]
  - w/ preconditioned multishift BiCG [Ahmad/Szyld/vanGijzen(2016)]

Small order benchmark examples

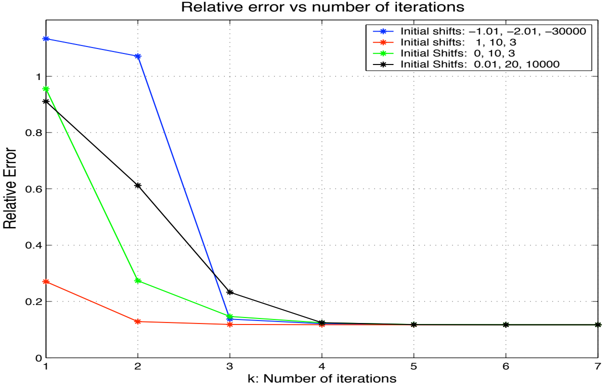
Model	r	IRKA	GFM	OPM
FOM-1	1	4.2683×10^{-1}	4.2709×10^{-1}	4.2683×10^{-1}
FOM-1	2	3.9290×10^{-2}	3.9299×10^{-2}	3.9290×10^{-2}
FOM-1	3	1.3047×10^{-3}	1.3107×10^{-3}	1.3047×10^{-3}
FOM-2	3	1.171×10^{-1}	1.171×10^{-1}	Divergent
FOM-2	4	8.199×10^{-3}	8.199×10^{-3}	8.199×10^{-3}
FOM-2	5	2.132×10^{-3}	2.132×10^{-3}	Divergent
FOM-2	6	5.817×10^{-5}	5.817×10^{-5}	5.817×10^{-5}
FOM-3	1	4.818×10^{-1}	4.818×10^{-1}	4.818×10^{-1}
FOM-3	2	2.443×10^{-1}	2.443×10^{-1}	Divergent
FOM-3	3	5.74×10^{-2}	5.98×10^{-2}	5.74×10^{-2}
FOM-4	1	9.85×10^{-2}	9.85×10^{-2}	9.85×10^{-2}

- **GFM**: Gradient Flow Method of Yan and Lam [1999]
- **OPM**: Optimal Projection Method of Hyland and Bernstein [1985]
- FOM-1: $n = 4$, FOM-2: $n = 7$, FOM-3: $n = 4$, FOM-4: $n = 2$,

- FOM-3: $\mathbf{H}(s) = \frac{s^2+15s+50}{s^4+5s^3+22s^2+79s+50}$

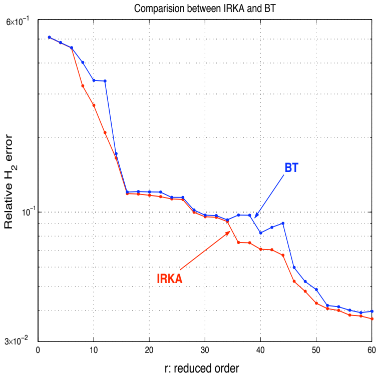
- $\mathbf{H}_3(s) = \frac{2.155s^2+3.343s+33.8}{(s+6.2217)(s+0.61774+\jmath 1.5628)(s+0.61774+\jmath 1.5628)}$

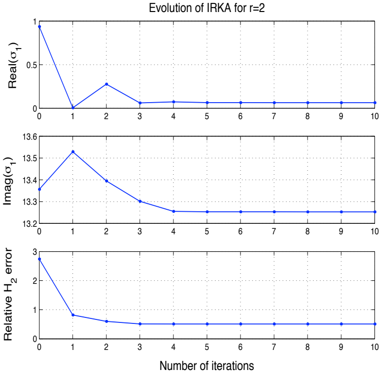
- $\mathcal{S}_1 = \{-1.01,\ -2.01,\ -30000\}$, $\mathcal{S}_2 = \{0,\ 10,\ 3\}$, $\mathcal{S}_3 = \{1,\ 10,\ 3\}$, and $\mathcal{S}_4 = \{0.01,\ 20,\ 10000\}$



ISS 12a Module

- $n = 1412$. Reduce to $r = 2 : 2 : 60$
- Compare with balanced truncation





Indoor-air environment in a conference room

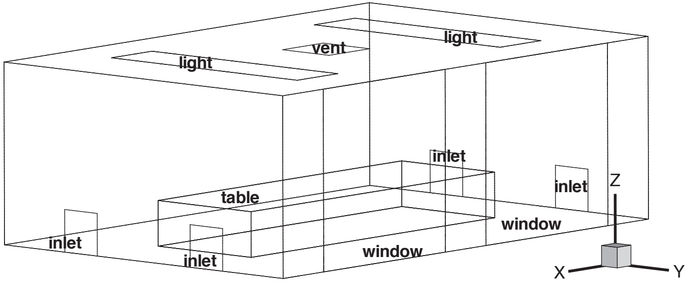


Figure: Geometry for our Indoor-air Simulation

- Four inlets, one return vent
- Thermal loads: two windows, two overhead lights and occupants
- FLUENT to simulate the indoor-air velocity, temperature and moisture.

- Modeled by

$$\begin{aligned}
\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} &= -\nabla P + \frac{1}{\mathrm{Re}} \Delta \mathbf{v} + \frac{\mathrm{Gr}}{\mathrm{Re}^2} T \hat{k} \\
\nabla \cdot \mathbf{v} &= 0 \\
\frac{\partial T}{\partial t} + \mathbf{v} \cdot \nabla T &= \frac{1}{\mathrm{RePr}} \Delta T + Bu, \\
\frac{\partial S}{\partial t} + \mathbf{v} \cdot \nabla S &= \frac{1}{Pe} \Delta S,
\end{aligned}$$

- $\mathbf{v}$: the velocity vector, P: the pressure,
  T: the temperature, S: the moisture concentration.
- Adiabatic boundary conditions on all surfaces except the inlets, windows and lights.
- FLUENT simulations with varying inlet temperature, occupant loads, as well as solar and lighting loads $\Rightarrow \overline{\mathbf{v}}$ was computed.

Finite Element Model of Convection/Diffusion

- A finite element model for thermal energy transfer with *frozen* velocity field $\overline{\mathbf{v}}$,

$$\frac{\partial T}{\partial t} + \overline{\mathbf{v}} \cdot \nabla T = \frac{1}{\mathrm{RePr}} \Delta T + Bu,$$

- leading to

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

with $n = 202140$, $m = 2$ inputs, and $p = 2$ outputs.
Inputs:

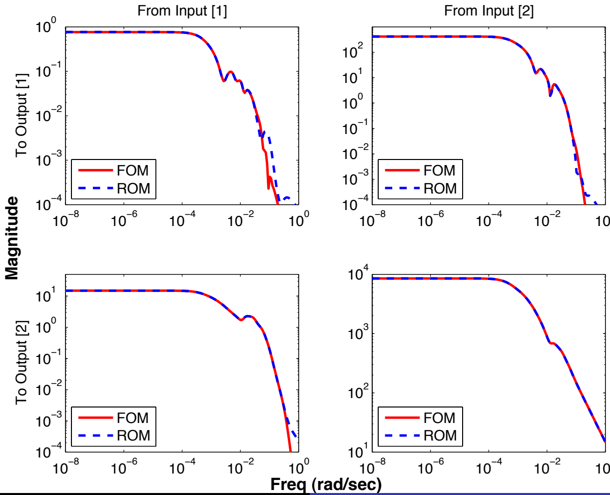
1. the temperature of the inflow air at all four vents, and
2. a disturbance caused by occupancy around the conference table,

Outputs:

1. the temperature at a sensor location on the *max x* wall,
2. the average temperature in an occupied volume around the table,

Revisit the conference room example

- Recall $n = 202140$, $m = 2$ and $p = 2$
- Reduced the order to $r = 30$ using IRKA.



- The $(2,2)$ block is associated with the dominant subsystem.
- Relative $\mathcal{H}_\infty$ errors in each subsystem by IRKA

	From Input [1]	From Input [2]
To Output [1]	6.62×10^{-3}	1.82×10^{-5}
To Output [2]	4.86×10^{-4}	5.40×10^{-7}

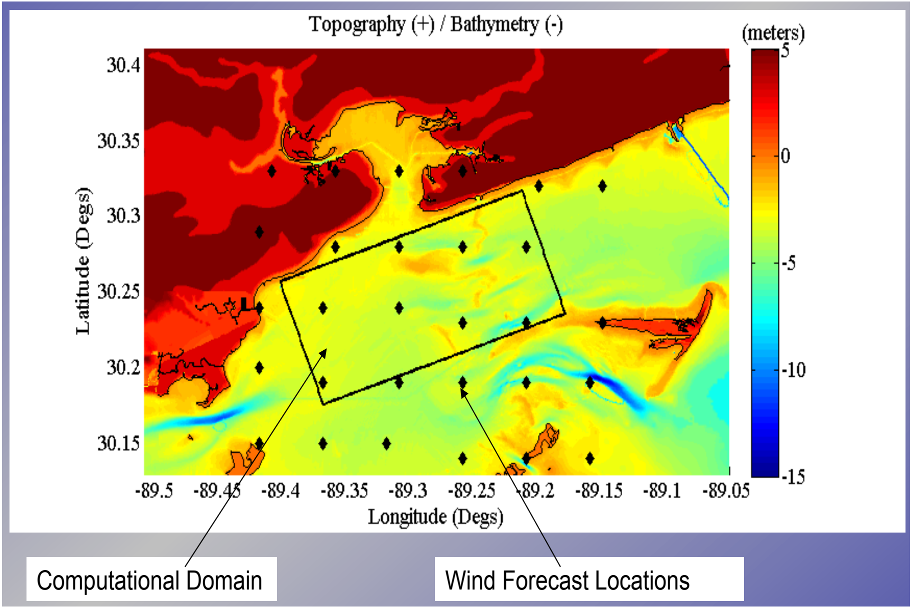
- Does IRKA pay off? How about some ad hoc selections:

	From Input [1]	From Input [2]
To Output [1]	9.19×10^{-2}	8.38×10^{-2}
To Output [2]	5.90×10^{-2}	2.22×10^{-2}

- One can keep trying different ad hoc selections but this is exactly what we want to avoid.

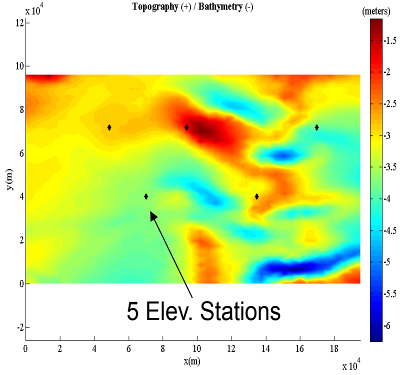
Storm Surge Modeling of Bay St. Louis, MS, USA

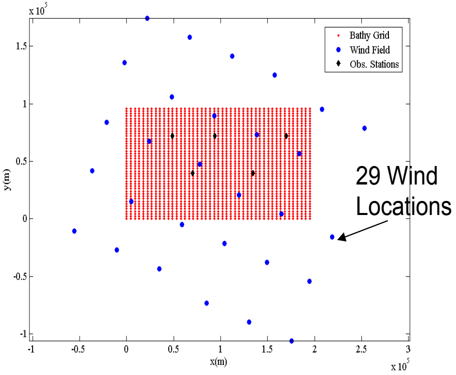
- Data: Chris Massey, US Army Corps of Eng. Res. & Dev. Ctr.


Topography (+) / Bathymetry (-)
(meters)
5
0
-5
-10
-15
30.4
30.35
30.3
30.25
30.2
30.15
Latitude (Degs)
-89.5
-89.45
-89.4
-89.35
-89.3
-89.25
-89.2
-89.15
-89.1
-89.05
Longitude (Degs)
Computational Domain
Wind Forecast Locations

Storm Surge Modeling of Bay St. Louis

- 29 wind-forecast locations
- Surface elevation measurements at five measurement stations.
- A model of the form $\begin{cases} \mathbf{E}\,\dot{\mathbf{x}} = \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t), \\ \mathbf{y} = \mathbf{C}\,\mathbf{x}(t) \end{cases}$
  results from linearization of Shallow Water Equations
  with $n = 5808$
- Reduced-order model to predict surface elevation given the wind-forecast data.

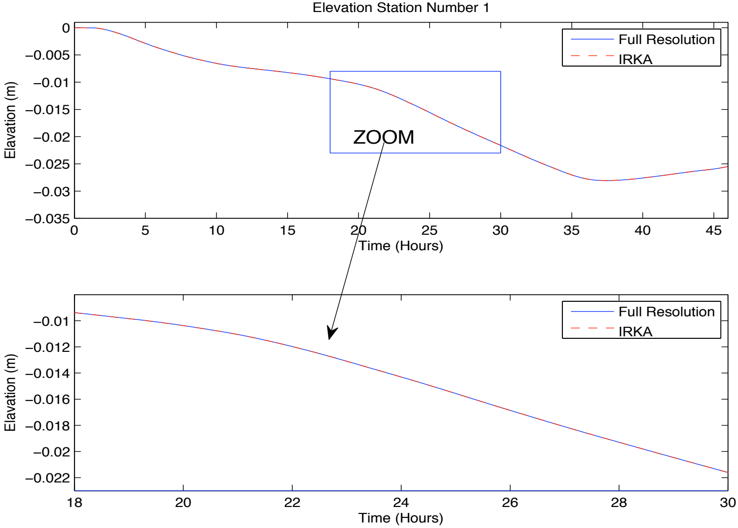




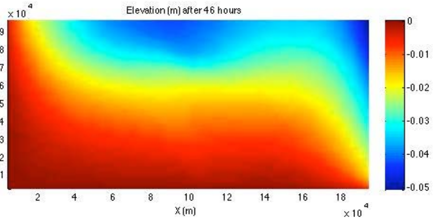
- Recall the model: $\begin{cases} \mathbf{E}\,\dot{\mathbf{x}} = \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t), \\ \mathbf{y} = \mathbf{C}\,\mathbf{x}(t) \end{cases}$

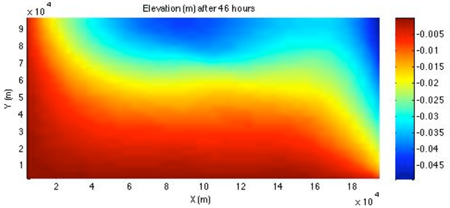
  with $n = 5808$, $m = 58$ and $\ell = 5$.
- Reduce the order to $r = 30$ with IRKA and compare with half-resolution discretization.

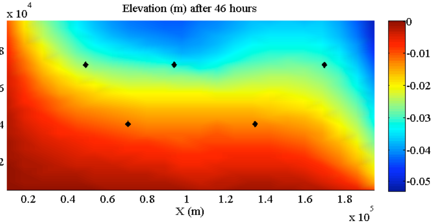
Elevation Station 1



Surface elevation after 46 hours

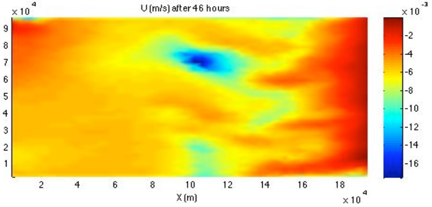


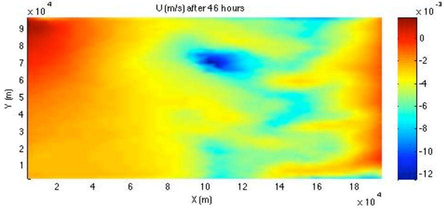


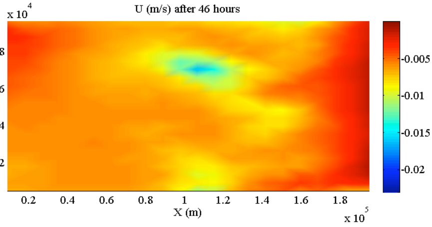


(1,1) plot: Full-resolution
(1.2) plot: r=30 IRKA reduction
(2,1) plot: Half-resolution

U Component of Velocity after 46 Hours

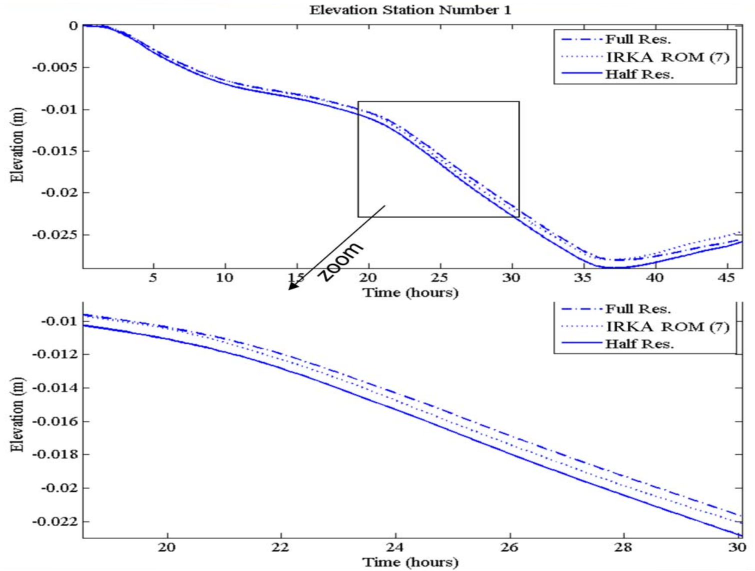






(1,1) plot: Full-resolution
(1.2) plot: r=30 IRKA reduction
(2,1) plot: Half-resolution

How about $r = 7$


Elevation Station Number 1
Full Res.
IRKA ROM (7)
Half Res.
Elevation (m)
Time (hours)
zoom
0
-0.005
-0.01
-0.015
-0.02
-0.025
5
10
15
20
25
30
35
40
45
Full Res.
IRKA ROM (7)
Half Res.
Elevation (m)
Time (hours)
-0.01
-0.012
-0.014
-0.016
-0.018
-0.02
-0.022
20
22
24
26
28
30

IRKA in other settings and application

- Cellular neurophysiology: [Kellems,Roos,Xiao,Cox (2009)].
- Bilinear Systems: [Benner/Breiten (2011)], [Flagg/Gugercin (2012)]

$$\dot{\mathbf{y}}(t) = \mathbf{A}\mathbf{y}(t) + \sum_{k=1}^{n_d} \mathbf{N}_k \mathbf{u}_k(t)\mathbf{y}(t) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{m}(t) = \mathbf{C}\,\mathbf{y}(t)$$

- Inverse Problems: [Druskin/Simoncini/Zaslavsky (2011)]
- $\mathcal{H}_\infty$-model reduction: [Flagg/B/Gugercin (2011)]
- Energy-efficient building design: [Borggard/Cliff/Gugercin (2012)]
- Aerospace Applications [Poussat-Vassal (2011)].
- Structural Models [Bonin et.al (2010)], [Wyatt, (2012)], [Polyuga et.al. (2012)]

Data-Driven IRKA: Freedom from realizations in $\mathbf{H}(S)$

- Recall the optimality conditions.

Theorem ([Gugercin/Antoulas/B,08])

Given $\mathbf{H}(s)$, *let* $\mathbf{H}_r(s) = \sum_{i=1}^{r} \frac{1}{s-\hat{\lambda}_i}\hat{\mathsf{c}}_i\hat{\mathsf{b}}_i^T$ *be the best stable* r^{th} *order approximation of* $\mathbf{H}$ *with respect to the* $\mathcal{H}_2$ *norm. Assume* $\mathbf{H}_r$ *has simple poles at* $\hat{\lambda}_1, \hat{\lambda}_2, \ldots \hat{\lambda}_r$*. Then*

$$\mathbf{H}(-\hat{\lambda}_k)\hat{\mathsf{b}}_k = \mathbf{H}_r(-\hat{\lambda}_k)\hat{\mathsf{b}}_k, \quad \hat{\mathsf{c}}_k^T\mathbf{H}(-\hat{\lambda}_k) = \hat{\mathsf{c}}_k^T\mathbf{H}_r(-\hat{\lambda}_k),$$

$$\text{and} \quad \hat{\mathsf{c}}_k^T\mathbf{H}'(-\hat{\lambda}_k)\hat{\mathsf{b}}_k = \hat{\mathsf{c}}_k^T\mathbf{H}_r'(-\hat{\lambda}_k)\hat{\mathsf{b}}_k \quad \text{for } k = 1, 2, \ldots, r.$$

- No assumption that $\mathbf{H}(s)$ needs to be rational, only that $\mathbf{H}_r(s)$ is.
- The conditions are valid for general non-rational $\mathbf{H}(s)$.
- IRKA iteratively corrects Hermite interpolants.

Recall (regular) IRKA:

Algorithm (Gugercin/Antoulas/B [2008])

1. *Choose* $\{\sigma_1, \ldots, \sigma_r\}$, $\{\hat{\mathsf{b}}_1, \ldots, \hat{\mathsf{b}}_r\}$ *and* $\{\hat{\mathsf{c}}_1, \ldots, \hat{\mathsf{c}}_r\}$
2. $\mathbf{V}_r = \left[(\sigma_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\hat{\mathsf{b}}_1 \ \cdots \ (\sigma_r\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\hat{\mathsf{b}}_r \right]$
   $\mathbf{W}_r = \left[(\sigma_1\mathbf{E} - \mathbf{A}^T)^{-1}\mathbf{C}^T\hat{\mathsf{c}}_1 \ \cdots \ (\sigma_r\mathbf{E} - \mathbf{A}^T)^{-1}\mathbf{C}^T\hat{\mathsf{c}}_r \right]$.
3. *while (not converged)*
   1. $\mathbf{A}_r = \mathbf{W}_r^T\mathbf{A}\mathbf{V}_r$, $\mathbf{E}_r = \mathbf{W}_r^T\mathbf{E}\mathbf{V}_r$, $\mathbf{B}_r = \mathbf{W}_r^T\mathbf{B}$, *and* $\mathbf{C}_r = \mathbf{C}\mathbf{V}_r$
   2. *Compute* $\mathbf{H}_r(s) = \sum_{i=1}^{r} \frac{\hat{\mathsf{c}}_i\hat{\mathsf{b}}_i^T}{s - \hat{\lambda}_i}$, *and set* $\{\sigma_i\} \longleftarrow \{-\hat{\lambda}_i\}$,
   3. $\mathbf{V}_r = \left[(\sigma_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\hat{\mathsf{b}}_1 \ \cdots \ (\sigma_r\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\hat{\mathsf{b}}_r \right]$
   4. $\mathbf{W}_r = \left[(\sigma_1\mathbf{E} - \mathbf{A}^T)^{-1}\mathbf{C}^T\hat{\mathsf{c}}_1 \ \cdots \ (\sigma_r\mathbf{E} - \mathbf{A}^T)^{-1}\mathbf{C}^T\hat{\mathsf{c}}_r \right]$.
4. $\mathbf{A}_r = \mathbf{W}_r^T\mathbf{A}\mathbf{V}_r$, $\mathbf{E}_r = \mathbf{W}_r^T\mathbf{E}\mathbf{V}_r$, $\mathbf{B}_r = \mathbf{W}_r^T\mathbf{B}$, $\mathbf{C}_r = \mathbf{C}\mathbf{V}_r$, $\mathbf{D}_r = \mathbf{D}$.

- Replace Hermite interpolation via projection with Loewner

Realization Independent IRKA (TF-IRKA)

Algorithm (Realization Independent IRKA B/Gugercin, 2012)

1. *Choose initial* σ_i, $\{\tilde{\mathsf{c}}_i\}$, *and* $\{\tilde{\mathsf{b}}_i\}$ *for* $i = 1, \ldots, r$.
2. *Evaluate* $\mathcal{H}(\sigma_i)$ *and* $\mathcal{H}'(\sigma_i)$ *for* $i = 1, \ldots, r$.
3. *while not converged*
   1. *Construct* $\mathbf{E}_r = -\mathbb{L}$, $\mathbf{A}_r = -\mathbb{M}$, $\mathbf{B}_r = \widetilde{\mathsf{Z}}^T$ *and* $\mathbf{C}_r = \widetilde{\mathsf{Y}}$
   2. *Construct* $\mathbf{H}_r(s) = \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r = \widetilde{\mathsf{Z}}^T(\mathbb{M} - s\,\mathbb{L})^{-1}\widetilde{\mathsf{Y}} = \sum_{i=1}^{r} \frac{\mathsf{c}_i \mathsf{b}_i^T}{s-\lambda_i}$
   3. $\sigma_i \longleftarrow -\lambda_i$, $\tilde{\mathsf{c}}_i \longleftarrow \mathsf{c}_i$, *and* $\tilde{\mathsf{b}}_i \longleftarrow \mathsf{b}_i$ *for* $i = 1, \ldots, r$
   4. *Evaluate* $\mathcal{H}(\sigma_i)$ *and* $\mathcal{H}'(\sigma_i)$ *for* $i = 1, \ldots, r$.
4. *Construct* $\mathbf{H}_r(s) = \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r = \widetilde{\mathsf{Z}}^T(\mathbb{M} - s\,\mathbb{L})^{-1}\widetilde{\mathsf{Y}}$

- Allows infinite order transfer functions !!
  e.g., $\mathcal{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A}_0 - e^{-\tau_1 s}\mathbf{A}_1 - e^{-\tau_2 s}\mathbf{A}_2)^{-1}\mathbf{B}$

Revisit: One-dimensional heat equation

- $\frac{\partial T}{\partial t}(z,t) = \frac{\partial^2 T}{\partial z^2}(z,t)$, $\frac{\partial T}{\partial t}(0,t) = 0$, $\frac{\partial T}{\partial z}(1,t) = u(t)$, and $y(t) = T(0,t)$
- $\mathcal{H}(s) = \dfrac{1}{\sqrt{s}\sinh\sqrt{s}}$
- Apply TF-IRKA. Cost: Evaluate $\mathcal{H}(s)$ and $\mathcal{H}'(s)$!!!
- Optimal points upon convergence: $\sigma_1 = 20.9418$, $\sigma_2 = 10.8944$.

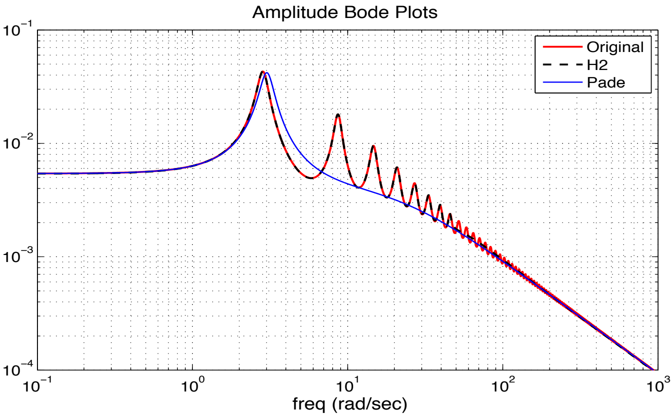
$$\mathcal{H}_r(s) = \frac{-0.9469s - 37.84}{s^2 + 31.84s + 228.1} + \frac{1}{s}$$

- $\|\mathcal{H} - \mathcal{H}_r\|_{\mathcal{H}_2} = 5.84 \times 10^{-3}$, $\|\mathcal{H} - \mathcal{H}_r\|_{\mathcal{H}_\infty} = 9.61 \times 10^{-4}$
- Balanced truncation of the discretized model:
  - $n = 1000$: $\|\mathcal{H} - \mathcal{H}_r\|_{\mathcal{H}_2} = 5.91 \times 10^{-3}$, $\|\mathcal{H} - \mathcal{H}_r\|_{\mathcal{H}_\infty} = 1.01 \times 10^{-3}$

Delay Example

- $\mathbf{E}\,\dot{\mathbf{x}}(t) = \mathbf{A}_1\mathbf{x}(t) + \mathbf{A}_2\mathbf{x}(t-\tau) + \mathbf{B}\,\mathbf{u}(t), \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$
- $\mathbf{E}, \mathbf{A}_1, \mathbf{A}_2 \in \mathbb{R}^{1000\times1000}$, $\mathbf{B}, \mathbf{C}^T \in \mathbb{R}^{1000}$
- $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A}_1 - e^{-\tau s}\mathbf{A}_2)^{-1}\mathbf{B}.$
- $\mathbf{H}'(s) = -\mathbf{C}(s\mathbf{E} - \mathbf{A}_1 - e^{-\tau s}\mathbf{A}_2)^{-1}(\mathbf{E} + \tau e^{-\tau s}\mathbf{A}_2)(s\mathbf{E} - \mathbf{A}_1 - e^{-\tau s}\mathbf{A}_2)^{-1}\mathbf{B}.$
- Obtain an order $r = 20$ optimal $\mathcal{H}_2$ rational approximation directly using $\mathbf{H}(s)$ and $\mathbf{H}'(s)$
- $\mathbf{H}_r(s)$ exactly interpolates $\mathbf{H}(s)$. This will not be the case if $e^{-\tau s}$ is approximated by a rational function.
- Moreover, the rational approximation of $e^{-\tau s}$ increases the order drastically.
- Multiple state-delays, delays in the input/output mappings are welcome.

Delay Example



- Relative $\mathcal{H}_\infty$ errors:
  $\mathcal{H}_2$-model: 8.63×10^{-3} Pade approx: 5.40×10^{-1}
- Pade Model has dimension $N = 3000$!!!

Conclusions

- Basic framework for interpolatory model reduction:
  - Focus on interpolatory projections instead of rational Krylov spaces.
  - Can create locally optimal reduced models effective.
  - Characterization of where to interpolate and in which direction — discussion of associated fixed point algorithm, IRKA.
- Data-driven Interpolation - the Loewner framework
  - Reduced models obtained directly from response measurements
  - Nonintrusive, locally optimal reduced models using only data.
- Importance of maintaining ancillary system structure
  - Generalized coprime realizations to preserve structure.
  - Important structural properties can be easily retained
- Open questions:
  - Extensions / Applications ? (e.g., DAEs, portHamiltonian/passive systems, bilinear/quadratic systems, parameterized systems, time-domain data-driven interpolation,...)

URL: www.math.vt.edu/people/gugercin/Publications.html

Related Papers:

1. S. Gugercin, A.C. Antoulas, and C.A. Beattie, *$\mathcal{H}_2$ model reduction for large-scale linear dynamical systems*, SIMAX, 2008.
2. C.A. Beattie and S. Gugercin, *Interpolatory Projection Methods for Structure-preserving Model Reduction*, Systems and Control Letters, 2009.
3. C.A. Beattie and S. Gugercin, *A Trust Region Method for Optimal $\mathcal{H}_2$ Model Reduction*, Proceedings of the 48th IEEE Conference on Decision and Control, 2009.
4. A.C. Antoulas, C.A. Beattie and S. Gugercin, *Interpolatory Model Reduction of Large-scale Dynamical Systems*, Efficient Modeling and Control of Large-Scale System, Springer, 2010.
5. C. Beattie and S. Gugercin, Model reduction by rational interpolation, in *Model Reduction and Approximation: Theory and Algorithms*. (Benner, P., Ohlberger, M., Cohen, A. and Willcox, K. eds), SIAM, 2017