Krylov-based model reduction of second-order systems with proportional damping

Chris Beattie and **Serkan Gugercin**
Department of Mathematics, Virginia Tech, USA

44th IEEE Conference on Decision and Control and European Control Conference CDC-ECC 2005
Seville, Spain, 12-15 December 2005

Introduction

- Consider the second-order system of the form

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{G}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$

- $\mathbf{H}(s) = \mathbf{C}(\mathbf{M}s^2 + \mathbf{G}s + \mathbf{K})^{-1}\mathbf{B}$
- $\mathbf{M}, \mathbf{G}, \mathbf{K} \in \mathbb{R}^{n\times n}$, and $\mathbf{B}, \mathbf{C}^T \in \mathbb{R}^n$.
- In many cases, n is too large for efficient simulation and control
- Generate, for some $r \ll n$, an r^{th} order reduced system

$$\mathbf{M}_r\ddot{\mathbf{x}}_r(t) + \mathbf{G}_r\dot{\mathbf{x}}_r(t) + \mathbf{K}_r\mathbf{x}_r(t) = \mathbf{B}_r\mathbf{u}(t), \quad \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t)$$

- $\mathbf{M}_r, \mathbf{G}_r, \mathbf{K}_r \in \mathbb{R}^{r\times r}$, and $\mathbf{B}_r, \mathbf{C}_r^T \in \mathbb{R}^r$
- $\mathbf{y}_r(t)$ approximates $\mathbf{y}(t)$ for a wide range of inputs $\mathbf{u}(t)$.

Krylov-based model reduction

$$\mathbf{H}(s) := \left\{ \begin{array}{rcl} \mathcal{E}\,\dot{\mathbf{q}}(t) & = & \mathcal{A}\mathbf{q}(t) + \mathcal{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) & = & \mathcal{C}\,\mathbf{q}(t) \end{array} \right. \Leftrightarrow \mathbf{H}(s) = \mathcal{C}\,(s\mathcal{E} - \mathcal{A})^{-1}\,\mathcal{B}$$

- Find $\mathbf{H}_r(s) = \mathcal{C}_r\,(s\mathcal{E}_r - \mathcal{A}_r)^{-1}\,\mathcal{B}_r$ so that

$$\boxed{\left.\frac{d^j\mathbf{H}(s)}{ds^j}\right|_{s=\sigma_k} = \left.\frac{d^j\mathbf{H}_r(s)}{ds^j}\right|_{s=\sigma_k}}, \quad \begin{array}{l} k = 1, \ldots, K \\ j = 0, \ldots, J-1 \end{array}.$$

- K : Number of interpolation points
- J : Number of interpolation conditions at each σ_k.
- $\mathcal{E}_r := \mathcal{Z}^T\mathcal{E}\mathcal{V}, \quad \mathcal{A}_r := \mathcal{Z}^T\mathcal{A}\mathcal{V}, \quad \mathcal{B}_r := \mathcal{Z}^T\mathcal{B}, \quad \mathcal{C}_r := \mathcal{C}\mathcal{V}$

- Given $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{N \times N}$ and $\mathbf{g} \in \mathbb{C}^{N}$, define the Krylov subspace:

$$\mathcal{K}_J(\boldsymbol{\mathcal{F}}, \mathbf{g}) = \text{span}\{\, \mathbf{g},\, \boldsymbol{\mathcal{F}}\mathbf{g},\, \boldsymbol{\mathcal{F}}^2\mathbf{g},\, \cdots,\, \boldsymbol{\mathcal{F}}^{J-1}\mathbf{g} \,\}$$

- Construct $\boldsymbol{\mathcal{V}}$ and $\boldsymbol{\mathcal{Z}}$ such that

$$\begin{aligned}
&\text{Ran}(\boldsymbol{\mathcal{V}}) = \text{span}\left\{\mathcal{K}_{J_1}(\boldsymbol{\mathcal{F}}_1, \mathbf{g}_1),\, \cdots,\, \mathcal{K}_{J_K}(\boldsymbol{\mathcal{F}}_K, \mathbf{g}_K)\right\}, \text{ and}\\
&\text{Ran}(\boldsymbol{\mathcal{Z}}) = \text{span}\left\{\mathcal{K}_{J_{K+1}}(\boldsymbol{\mathcal{F}}_{K+1}, \mathbf{g}_{K+1}), \cdots, \mathcal{K}_{J_{2K}}(\boldsymbol{\mathcal{F}}_{2K}, \mathbf{g}_{2K})\right\},
\end{aligned}$$

$$\begin{aligned}
\text{where}\quad &\boldsymbol{\mathcal{F}}_i = (\sigma_i\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{E}}, \quad \mathbf{g}_i = (\sigma_i\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}}, \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{for } i = 1, \ldots, K\\
&\boldsymbol{\mathcal{F}}_i = (\sigma_i\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-T}\boldsymbol{\mathcal{E}}^T, \quad \mathbf{g}_i = (\sigma_i\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-T}\boldsymbol{\mathcal{C}}^T, \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{for } i = K+1, \ldots, 2K
\end{aligned}$$

- $\mathbf{H}_r(s)$ satisfies the interpolation conditions

Krylov reduction of second-order systems

- Convert $\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{G}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$
  into $\quad \mathcal{E}\dot{\mathbf{q}} = \mathcal{A}\mathbf{q}(t) + \mathcal{B}u(t), \quad \mathbf{y}(t) = \mathcal{C}\mathbf{q}(t)$ where

$$\mathcal{E} = \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{M} \end{bmatrix}, \ \mathcal{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\alpha\mathbf{M} - \beta\mathbf{K} \end{bmatrix}, \ \mathcal{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B} \end{bmatrix}, \mathcal{C}^T = \begin{bmatrix} \mathbf{C}^T \\ \mathbf{0} \end{bmatrix}.$$

- Apply first-order Krylov reduction and transfer back.
- Not always possible and structure is lost $\Longrightarrow$
- Apply reduction directly in the second-order system framework
- Find $\mathbf{W} \in \mathbb{R}^{n\times r}$ such that $\mathbf{W}^T\mathbf{W} = \mathbf{I}_r$ and

$$\mathbf{M}_r = \mathbf{W}^T\mathbf{M}\mathbf{W}, \ \mathbf{G}_r = \mathbf{W}^T\mathbf{G}\mathbf{W}, \ \mathbf{K}_r = \mathbf{W}^T\mathbf{K}\mathbf{W},$$
$$\mathbf{B}_r = \mathbf{W}^T\mathbf{B}, \ \text{and } \mathbf{C}_r = \mathbf{C}\mathbf{W}$$

Second-order Krylov Subspaces (Bai[2003])

- Given $\mathfrak{F}_1, \mathfrak{F}_2 \in \mathbb{C}^{n\times n}$ and $\mathbf{g} \in \mathbb{C}^n$, second-order Krylov subspace:

$$\mathcal{K}_J^{(2)}(\mathfrak{F}_1, \mathfrak{F}_2, \mathbf{g}) = \text{span}\{\, \mathbf{r}_0, \, \mathbf{r}_1, \, \cdots \, , \, \mathbf{r}_{J-1}\} \quad \text{where}$$

$$\mathbf{r}_0 = \mathbf{g}, \quad \mathbf{r}_1 = \mathfrak{F}_1 \mathbf{r}_0, \quad \mathbf{r}_j = \mathfrak{F}_1 \mathbf{r}_{j-1} + \mathfrak{F}_2 \mathbf{r}_{j-2}, \quad \text{for} \quad j \geq 2.$$

- For second-order reduction with interpolation point σ:

$$\mathbf{W}_r = \mathcal{K}_r^{(2)}(-\widetilde{\mathbf{K}}\widetilde{\mathbf{G}}, -\widetilde{\mathbf{K}}\mathbf{M}, \mathbf{r}_0) \qquad \text{where}$$

$$\widetilde{\mathbf{K}} = \sigma^2 \mathbf{M} + \sigma \mathbf{G} + \mathbf{K}, \quad \widetilde{\mathbf{D}} = 2\sigma \mathbf{M} + \mathbf{G}, \quad \mathbf{r}_0 = \widetilde{\mathbf{K}}^{-1}\mathbf{B}$$

- Reduction directly in the second-order system framework
- Reduced model matches the first r moments at σ.

Second-order systems with proportional damping

- $\mathbf{M}\ddot{\mathbf{x}}(t) + (\alpha\mathbf{M} + \beta\mathbf{K})\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$ where $\alpha, \beta > 0$, and $\alpha\beta < 1$.

- First-order equivalent system: $\mathbf{H}(s) = \mathcal{C}\,(s\mathcal{E} - \mathcal{A})^{-1}\,\mathcal{B}$ with

$$\mathcal{E} = \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{M} \end{bmatrix}, \; \mathcal{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\alpha\mathbf{M} - \beta\mathbf{K} \end{bmatrix}, \; \mathcal{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B} \end{bmatrix}, \mathcal{C}^T = \begin{bmatrix} \mathbf{C}^T \\ \mathbf{0} \end{bmatrix}.$$

- To match the first r moments at σ, construct

$$\mathcal{V}_{2n\times r} = \text{span}\left\{\mathbf{g},\, \mathcal{F}\mathbf{g},\, \cdots,\, \mathcal{F}^{r-1}\mathbf{g}\right\}$$

where

$$\mathcal{F} = (\sigma\mathcal{E} - \mathcal{A})^{-1}\mathcal{E} \quad \text{and} \quad \mathbf{g} = (\sigma\mathcal{E} - \mathcal{A})^{-1}\mathcal{B}.$$

Theorem: Given the above set-up, $\boxed{\mathcal{V}_{2n\times r} \subset \mathcal{W}_{n\times r} \oplus \mathcal{W}_{n\times r}}$ where

$$\boxed{\mathcal{W}_{n\times r} = \mathcal{K}_r(\mathbf{K}_\sigma^{-1}\mathbf{M},\, \mathbf{K}_\sigma^{-1}\mathbf{B})} \text{ with } \boxed{\mathbf{K}_\sigma = \sigma^2\mathbf{M} + \sigma(\alpha\mathbf{M} + \beta\mathbf{K}) + \mathbf{K}}.$$

- Model reduction directly in the second-order framework using

$$\mathrm{Ran}(\mathbf{W}_r) = \mathcal{W} \quad \text{with} \quad \mathbf{W}_r^T\mathbf{W}_r = \mathbf{I}_r, \quad \mathbf{W}_r \in \mathbb{R}^{n\times r}$$

- $\mathbf{H}_r(s)$ matches r moments at σ.

$$\boxed{\text{Second-order reduction using } \textit{regular first-order Krylov subspace } \mathbf{W}_r}$$

- If $\sigma = 0$, $\mathcal{W}_{n\times r} = \mathrm{span}\{\mathbf{K}^{-1}\mathbf{B}, \cdots, (\mathbf{K}^{-1}\mathbf{M})^{r-1}\mathbf{K}^{-1}\mathbf{B}\}$
- If $\sigma = \infty$, $\mathcal{W}_{n\times r} = \mathrm{span}\{\mathbf{M}^{-1}\mathbf{B}, \cdots, (\mathbf{M}^{-1}\mathbf{K})^{r-1}\mathbf{M}^{-1}\mathbf{B}\}$

Proof: $\mathbf{g} = (\sigma\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}} = \begin{bmatrix} \mathbf{K}_\sigma^{-1}\mathbf{B} \\ \sigma\mathbf{K}_\sigma^{-1}\mathbf{B} \end{bmatrix}$ and $\mathbf{K}_\sigma^{-1}\mathbf{B} \in \mathcal{W}$.

- Let $\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}$ and $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{K}_p(\mathbf{K}_\sigma^{-1}\mathbf{M}, \mathbf{K}_\sigma^{-1}\mathbf{B})$.
- The next vector in $\mathcal{W}$ is

$$\bar{\mathbf{v}} = (\sigma\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{E}}\mathbf{v} = \begin{bmatrix} \bar{\mathbf{v}}_1 \\ \bar{\mathbf{v}}_2 \end{bmatrix} = \begin{bmatrix} \kappa_1\mathbf{v}_1 + \kappa_2\mathbf{K}_\sigma^{-1}\mathbf{M}\mathbf{v}_1 + \kappa_3\mathbf{K}_\sigma^{-1}\mathbf{M}\mathbf{v}_2 \\ \kappa_4\mathbf{v}_1 + \kappa_5\mathbf{K}_\sigma^{-1}\mathbf{M}\mathbf{v}_1 + \kappa_6\mathbf{K}_\sigma^{-1}\mathbf{M}\mathbf{v}_2 \end{bmatrix}$$

- $\Rightarrow \bar{\mathbf{v}}_1, \bar{\mathbf{v}}_2 \in \mathcal{K}_{p+1}(\mathbf{K}_\sigma^{-1}\mathbf{M}, \mathbf{K}_\sigma^{-1}\mathbf{B})$

- Moment matching: Second-order reduction using $\mathrm{Ran}(\mathbf{W}) = \mathcal{W}$ is equivalent to first-order reduction with $\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{Z}} = \begin{bmatrix} \mathbf{W} & 0 \\ 0 & \mathbf{W} \end{bmatrix}$.

Approximation by Interpolation (Beattie[2004])

- Best uniform approximation $\mathbf{H}_r(s)$ that minimizes $\max_{\omega \in \mathbb{R}} |\mathbf{H}(\jmath\omega) - \mathbf{H}_r(\jmath\omega)|$ should make $|\mathbf{H}(\jmath\omega) - \mathbf{H}_r(\jmath\omega)| \approx$ constant as $r \to \infty$ (”near-circularity” of best uniform rational approximation error (Trefethen, 1981) – related to Chebyshev equioscillation theorem.
- Best uniform approximations are hard to calculate.
- Interpolants are easy to calculate.
- How to choose the shifts ? Pick interpolation points carefully to recover a ‘good’ uniform approximant.
- $\log |\mathbf{H}(z) - \mathbf{H}_r(z)|$ has

$$\begin{cases} \text{positive singularities at system eigenvalues.} \\ \text{negative singularities at interpolation points.} \end{cases}$$

- Pick interpolation points to balance the contours of $\log|H(z) - H_r(z)|$ (makes $\log|H(z) - H_r(z)|$ nearly constant along the imaginary axis)
- Interpolation at Ritz values mirrored across the imaginary axis is good.
- Mirror Ritz values: optimal choice for $\mathcal{H}_2$ minimization as well (G./Antoulas/Beattie [2004])
- If system spectra is structured (e.g., circular), interpolation at 'equivalent lumped charge' locations will have much the same effect.

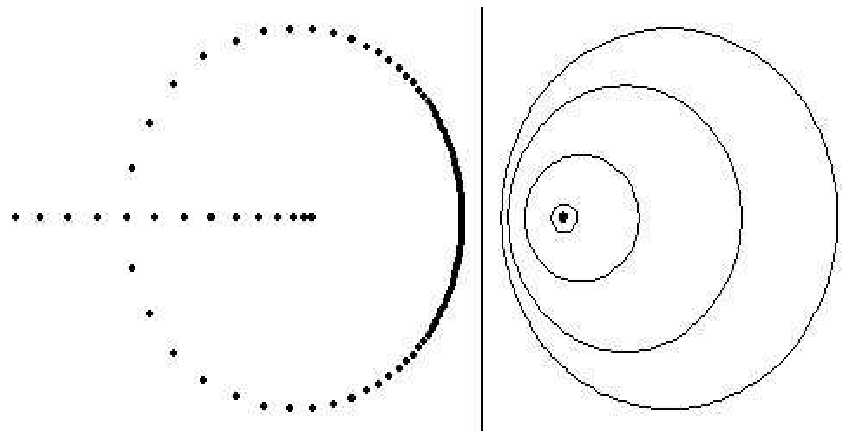
Shift Selection for Proportional Damping

- $\mathbf{M}\ddot{\mathbf{x}} + (\alpha\mathbf{M} + \beta\mathbf{K})\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{p}$.

  Proposition: All damped eigenvalues (the system poles) are on circle with center: $-\frac{1}{\beta}$, radius: $\frac{\sqrt{1-\alpha\beta}}{\beta}$ and on ray $(\infty, -\frac{1}{\beta}]$.

- Distribution depends on undamped natural frequencies, but usual elastic vibration models lead to distributions that are close to "equilibrium condenser distributions"
- Interpret $\log|\mathbf{H}(z) - \mathbf{H}_r(z)|$ as potential function associated with charge distribution (poles have net charge of $+1$, interpolation points have net charge of -1).

- Only ONE shift is necessary - replace aggregate of interpolation points (negative charge distribution) with single shift (an equivalent lumped charge) at $\boxed{\sigma_* = \sqrt{\frac{\alpha}{\beta}}.}$



- Optimal choice for condenser distribution of system poles; pretty good choice for most **K** and **M**).
- Regular first-order Krylov subspace using a *single* shift: $\sqrt{\frac{\alpha}{\beta}}$.

Exact Condenser Distribution

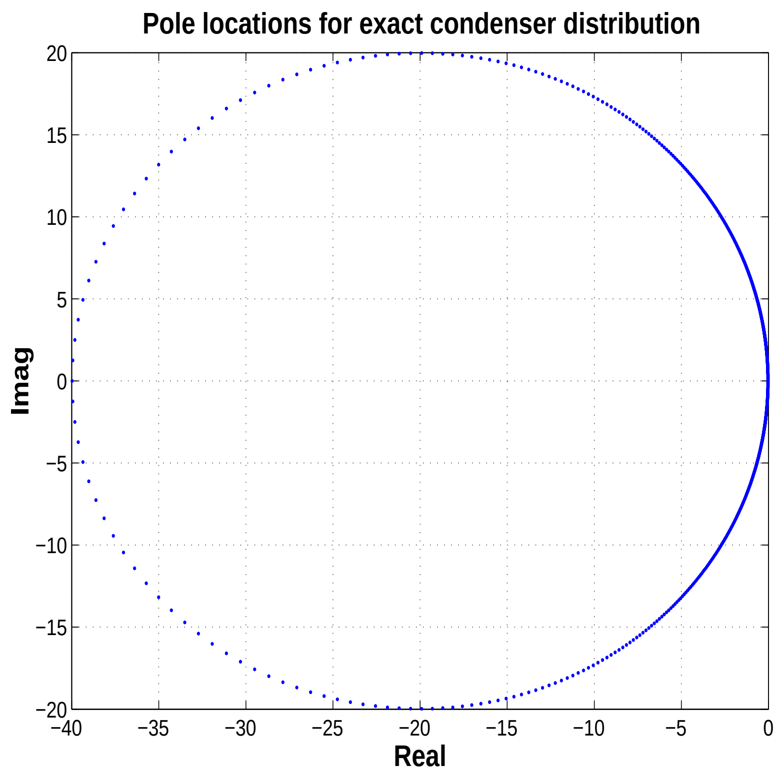
- Pick $\alpha,\ \beta \in (0,1)$

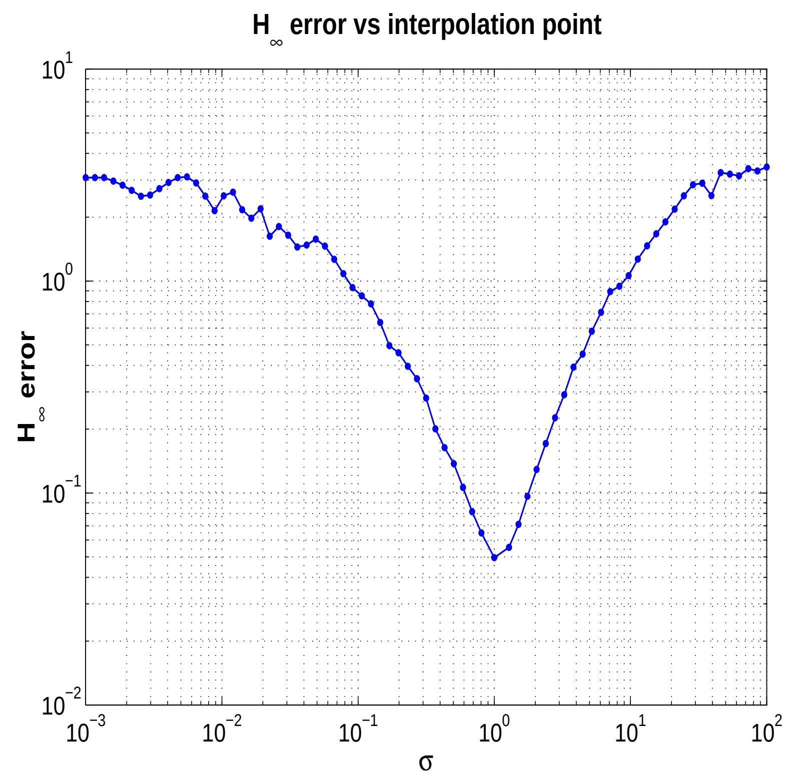
$$\mathbf{K} = \frac{\alpha}{\beta}\begin{bmatrix} \frac{2-\sqrt{1-\alpha\beta}}{\sqrt{1-\alpha\beta}} & -1 & 0 & \cdots & \\ -1 & \frac{2}{\sqrt{1-\alpha\beta}} & & & \\ 0 & & \ddots & & 0 \\ \vdots & & & \frac{2}{\sqrt{1-\alpha\beta}} & -1 \\ & & 0 & -1 & \frac{2-\sqrt{1-\alpha\beta}}{\sqrt{1-\alpha\beta}} \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} \frac{2+\sqrt{1-\alpha\beta}}{\sqrt{1-\alpha\beta}} & 1 & 0 & \cdots & \\ 1 & \frac{2}{\sqrt{1-\alpha\beta}} & & & \\ 0 & & \ddots & & 0 \\ \vdots & & & \frac{2}{\sqrt{1-\alpha\beta}} & 1 \\ & & 0 & 1 & \frac{2+\sqrt{1-\alpha\beta}}{\sqrt{1-\alpha\beta}} \end{bmatrix}.$$

- $\mathbf{G} = \alpha\,\mathbf{M} + \beta\,\mathbf{K}, \quad \alpha = \beta = 0.05$
- $\mathbf{B} = \mathbf{C}^T = [\,1 \;\; 0 \;\; 0 \;\; \cdots \;\; 0\,]^T.$

- Reduction from $n = 2000$ to $r = 30$ using a single shift

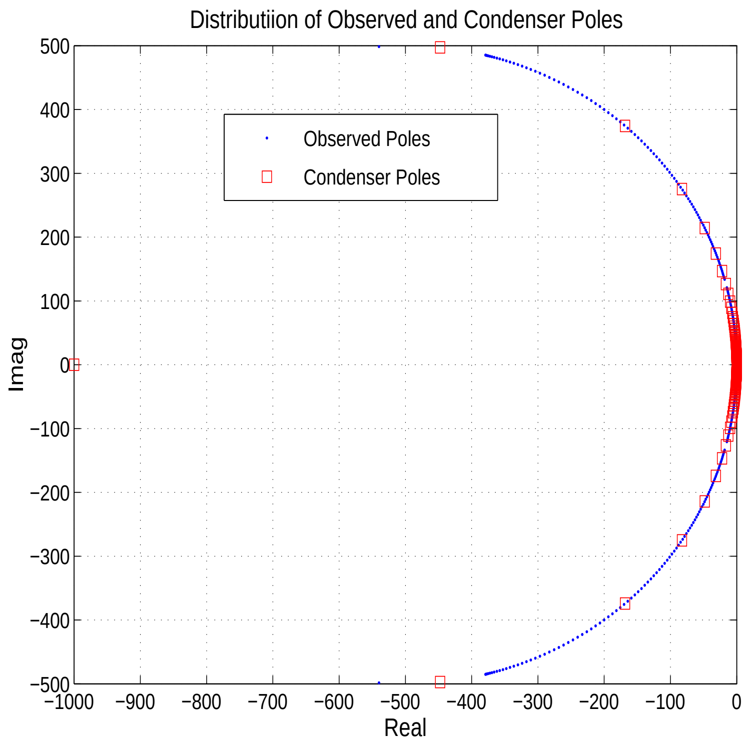


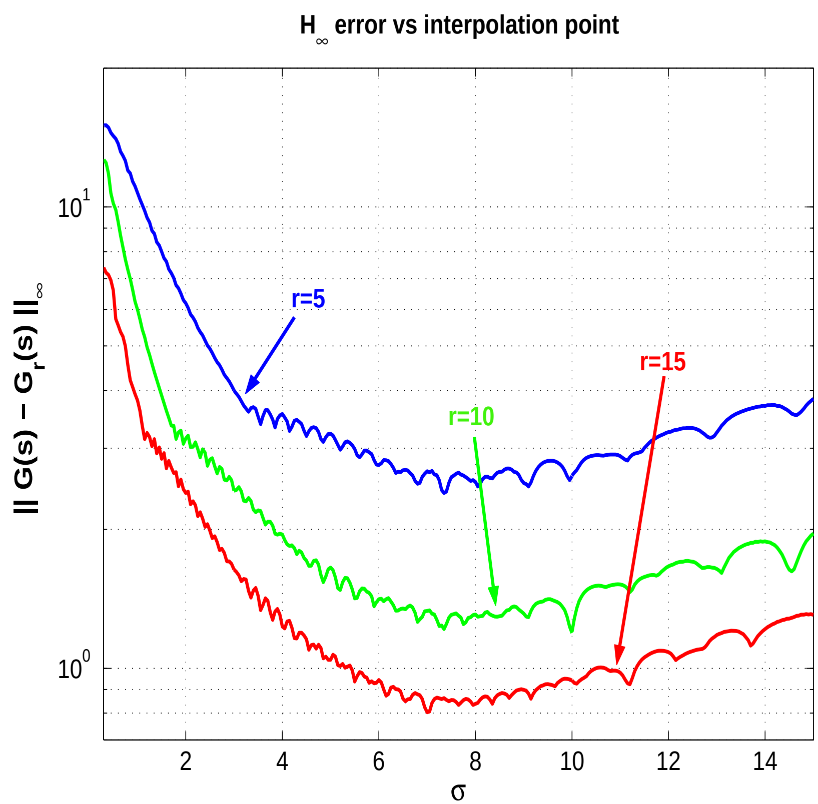


- $\sigma_* = \sqrt{\frac{\alpha}{\beta}} = 1$ is the optimal shift.

A 1-D Beam Model:

- $n = 2000$. $\alpha = 1/10$, $\beta = 1/500$, $\mathbf{B} = \mathbf{e}_1$, $\mathbf{C} = \mathbf{e}_{200}^T$.

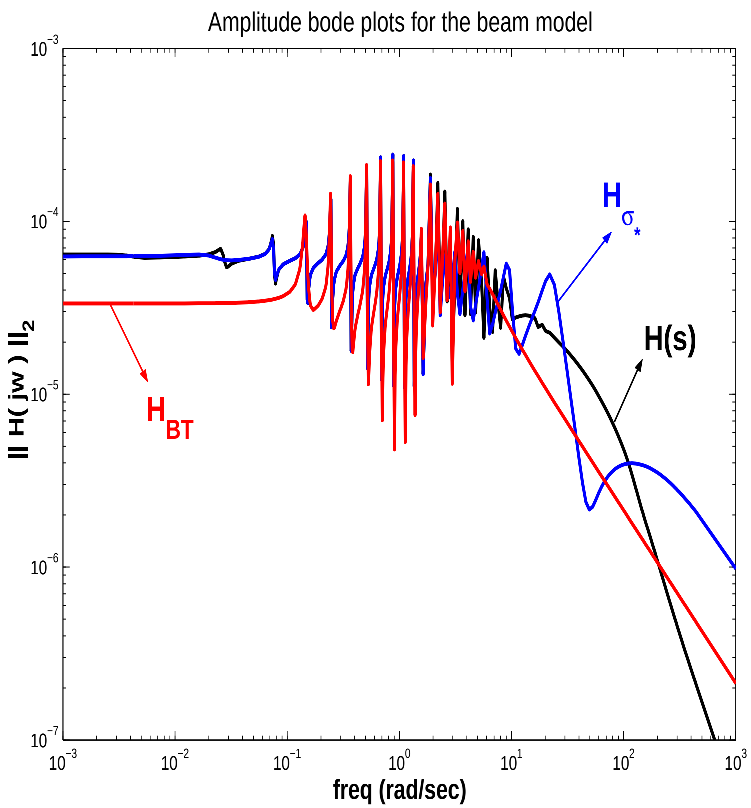


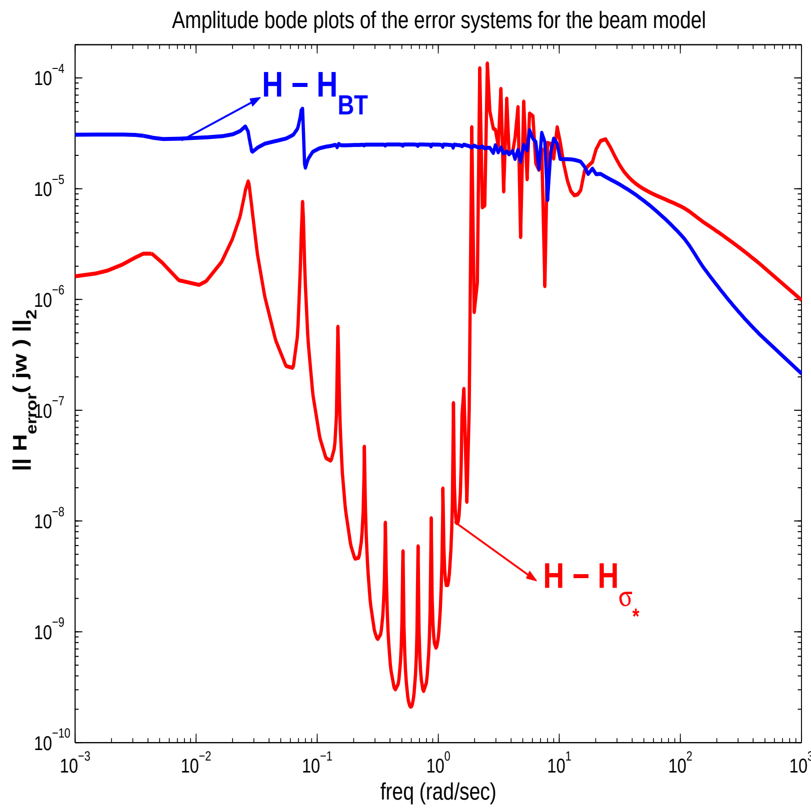


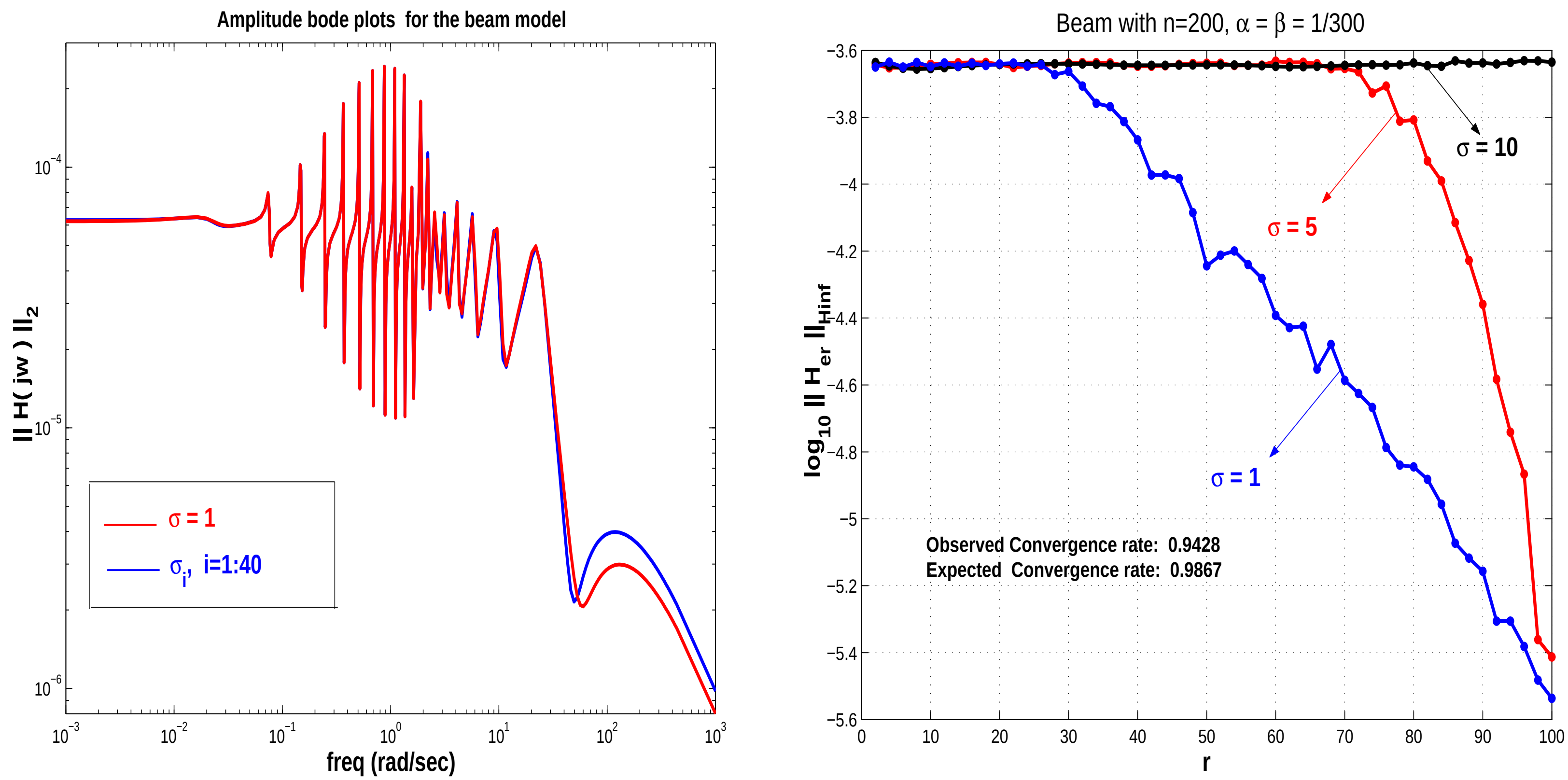
- $\sigma_* = \sqrt{\frac{\alpha}{\beta}} = 7.0711$: Very close to being optimal.

Another 1-D Beam Model:

- $n = 200$ and $\alpha = \beta = 1/300$, $\mathbf{B} = \mathbf{C}^T = \mathbf{e}_1$.
- Compare with balanced truncation and other shift selections
- Reduction done in the first-order framework to $r^{(1)} = 40$ ($r = 20$).






Amplitude bode plots for the beam model
|| H(jw) ||2
freq (rad/sec)
σ = 1
σi, i=1:40
Beam with n=200, α = β = 1/300
log10 || Her ||Hinf
r
σ = 10
σ = 5
σ = 1
Observed Convergence rate: 0.9428
Expected Convergence rate: 0.9867

Conclusions

- Considered $\mathbf{M}\ddot{\mathbf{x}} + (\alpha\mathbf{M} + \beta\mathbf{K})\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$
- Second-order Krylov reduction using first-order Krylov subspaces
- Pole locations for proportional damping
- Equivalent single shift (lumped charge) for aggregate of interpolation points (negative charge distribution)
- Optimal single shift $\sigma = \sqrt{\frac{\beta}{\alpha}}$ for condenser distribution
- Close to optimal in general
- Future work: Extensions to other types of damping.