The role of mirror-image points in system theory

A.C. Antoulas

aca@rice.edu
URL: www-ece.rice.edu/˜aca

Rice University

Sde Boker, September 2005

Contents

Contents

1. Positive real interpolation and model reduction.

Contents

1. Positive real interpolation and model reduction.

2. Optimal H_2 model reduction.

Contents

1. Positive real interpolation and model reduction.

2. Optimal H_2 model reduction.

3. Hankel norm model reduction.

Passive model reduction

Passive systems:

$Re \int_{-\infty}^{t} u(\tau)^* y(\tau) \mathrm{d}\tau \geq 0, \forall\, t \in \mathbb{R}, \forall\, u \in \mathbf{L}_2(\mathbb{R})$.

Positive real rational functions:

(1) $G(s) = D + C(sI - A)^{-1}B$ is analytic for $Re(s) > 0$,

(2) $Re\, G(s) \geq 0$ for $Re(s) \geq 0, \quad s$ not a pole of $G(s)$.

Theorem: $\mathbf{\Sigma} = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$ is passive $\Leftrightarrow$ $G(s)$ is positive real.

The new approach: Ingredients

- Positive realness of $G(s)$ implies the existence of a *spectral factorization* $G(s) + G^*(-s) = W(s)W^*(-s)$, where $W(s)$ is the stable rational and $W(s)^{-1}$ is also stable. The *spectral zeros* λ_i of the system are the zeros of the spectral factor $W(\lambda_i) = 0$, $i = 1, \cdots, n$.

The new approach: Ingredients

- Positive realness of $G(s)$ implies the existence of a *spectral factorization* $G(s) + G^*(-s) = W(s)W^*(-s)$, where $W(s)$ is the stable rational and $W(s)^{-1}$ is also stable. The *spectral zeros* λ_i of the system are the zeros of the spectral factor $W(\lambda_i) = 0$, $i = 1, \cdots, n$.

- *Rational Lanczos.* Skelton – de Villemagne and Grimme – Van Dooren have shown (assuming for simplicity $m = p = 1$) the following. Let

The new approach: Ingredients

- Positive realness of $G(s)$ implies the existence of a *spectral factorization* $G(s) + G^*(-s) = W(s)W^*(-s)$, where $W(s)$ is the stable rational and $W(s)^{-1}$ is also stable. The *spectral zeros* λ_i of the system are the zeros of the spectral factor $W(\lambda_i) = 0$, $i = 1, \cdots, n$.

- *Rational Lanczos.* Skelton – de Villemagne and Grimme – Van Dooren have shown (assuming for simplicity $m = p = 1$) the following. Let

$$V = \left[(\lambda_1 I - A)^{-1}B \;\cdots\; (\lambda_k I - A)^{-1}B\right] \in \mathbb{R}^{n\times k}$$

$$\bar{W}^* = \begin{bmatrix} C(\lambda_{k+1} I - A)^{-1} \\ \vdots \\ C(\lambda_{2k} \;\; I - A)^{-1} \end{bmatrix} \in \mathbb{R}^{k\times n}$$

where $\lambda_i \neq \lambda_{k+j}$, $i, j = 1, \cdots, k$; let the $k \times k$ matrix $\Gamma = W^* V$ be non-singular: $\det \Gamma \neq 0$; $W^* = \Gamma^{-1} \bar{W}^*$. The transfer function of $\hat{\mathbf{\Sigma}}$:

$$\hat{A} = W^* A V,\ \hat{B} = W^* B,\ \hat{C} = C V,\ \hat{D} = D$$

interpolates the transfer function of the original system $\mathbf{\Sigma}$ at the interpolation points λ_i:

$$G(\lambda_i) = D + C(\lambda_i I_n - A)^{-1} B = \hat{D} + \hat{C}(\lambda_i I_k - \hat{A})^{-1} \hat{B} = \hat{G}(\lambda_i)$$

for $i = 1, \cdots, n$.

Positive real interpolation

Problem. Given (λ_i, h_i), find a positive real rational function $G(s)$ such that: $G(\lambda_i) = h_i$, $i = 1, \cdots, n$.

Positive real interpolation

Problem. Given (λ_i, h_i), find a positive real rational function $G(s)$ such that: $G(\lambda_i) = h_i$, $i = 1, \cdots, n$.

Theorem 1. The problem is solvable if and only if the Pick matrix $P = \left[\frac{h_i + h_j^*}{\lambda_i + \lambda_j^*}\right]$, is positive semi-definite.

Positive real interpolation

Problem. Given (λ_i, h_i), find a positive real rational function $G(s)$ such that: $G(\lambda_i) = h_i$, $i = 1, \cdots, n$.

Theorem 1. The problem is solvable if and only if the Pick matrix $P = \left[\frac{h_i + h_j^*}{\lambda_i + \lambda_j^*}\right]$, is positive semi-definite.

Theorem 2. If the problem is solvable, a minimal rational interpolant of the pairs (λ_i, h_i), together with the MIRROR IMAGE pairs $(-\lambda_i^*, -h_i^*)$, is positive real.

Main result

- **Method**: Rational Lanczos
- **Difficulty**: Data $(A, B, C, D) \Rightarrow$ all interpolation points must be samples of **same** rational function: $G(s) = D + C(sI - A)^{-1}B$.
- **Solution**: interpolation points = spectral zeros

Main result

- **Method**: Rational Lanczos
- **Difficulty**: Data $(A, B, C, D) \Rightarrow$ all interpolation points must be samples of **same** rational function: $G(s) = D + C(sI - A)^{-1}B$.
- **Solution**: interpolation points = spectral zeros

Recall:

$$V = \left[(\lambda_1 I - A)^{-1}B \ \cdots \ (\lambda_k I - A)^{-1}B\right] \in \mathbb{R}^{n\times k}$$

$$\bar{W}^* = \begin{bmatrix} C(\lambda_{k+1} I - A)^{-1} \\ \vdots \\ C(\lambda_{2k} \ \ I - A)^{-1} \end{bmatrix} \in \mathbb{R}^{k\times n}, \ W^* = (\bar{W}^* V)^{-1}\bar{W}^*$$

Main result

- **Method**: Rational Lanczos
- **Difficulty**: Data $(A, B, C, D) \Rightarrow$ all interpolation points must be samples of **same** rational function: $G(s) = D + C(sI - A)^{-1}B$.
- **Solution**: interpolation points = spectral zeros

Main result. Given is the stable and passive system $\boldsymbol{\Sigma}$. If V, W are defined as above, where $\lambda_1, \cdots, \lambda_k$ are **spectral zeros**, and in addition $\boxed{\lambda_{k+i} = -\lambda_i^*}$, the reduced system $\hat{\boldsymbol{\Sigma}}$ satisfies:

(i) the interpolation constraints, **(ii)** it is stable, and **(iii)** it is passive.

Spectral zeros as eigenvalues

The *zeros* of the spectral factors can be computed by solving an *eigenvalue problem* involving A, B, C, D. If $D + D^*$ is non-singular the computation of the zeros of W is reduced to the computation of the eigenvalues of the matrix:

$$\mathbf{H} = \begin{bmatrix} A & 0 \\ 0 & -A^* \end{bmatrix} - \begin{bmatrix} B \\ -C^* \end{bmatrix} (D + D^*)^{-1} \begin{bmatrix} C & B^* \end{bmatrix} \in \mathbf{R}^{2n \times 2n}$$

which are in the left-half of the complex plane. Notice that $\mathbf{H}$ is a **Hamiltonian** matrix $\iff$ $(\mathbf{H}J)^* = \mathbf{H}J$, where $J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$. ■

Spectral zeros as generalized eigenvalues

If $D + D^*$ is singular, the spectral zeros are the finite generalized eigenvalues $\lambda \in \sigma(\mathcal{A}, \mathcal{E})$: $\text{rank}\,(\mathcal{A} - \lambda\mathcal{E}) < 2n + p$, where

$$\mathcal{A} = \begin{pmatrix} A & & B \\ & -A^* & -C^* \\ C & B^* & D + D^* \end{pmatrix} \quad \text{and} \quad \mathcal{E} = \begin{pmatrix} I & & \\ & I & \\ & & \mathbf{0} \end{pmatrix}$$

Due to the structure, the following reflection property is satisfied:
$\lambda \in \sigma(\mathcal{A}, \mathcal{E}) \iff -\lambda^* \in \sigma(\mathcal{A}, \mathcal{E})$.

Interpolation via invariant subspaces: Sorensen

$$\begin{pmatrix} A & & B \\ & -A^* & -C^* \\ C & B^* & D+D^* \end{pmatrix} \overbrace{\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}}^{k} = \overbrace{\begin{bmatrix} X \\ Y \\ \mathbf{0} \end{bmatrix}}^{k} R$$

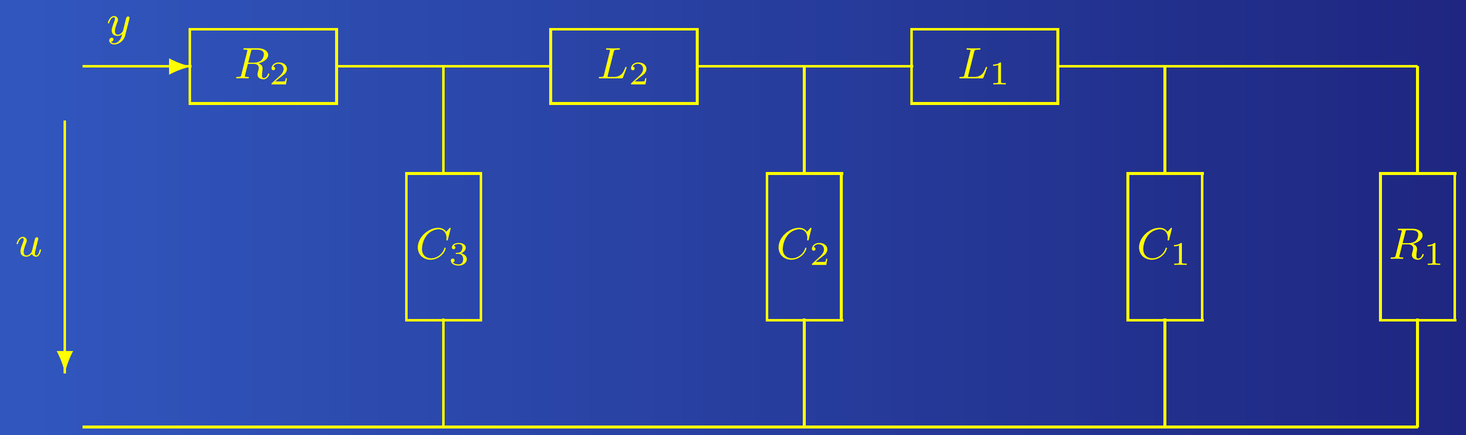
Lemma. X and Y are both full rank with $X^*Y = Y^*X$.

Construction of projectors:

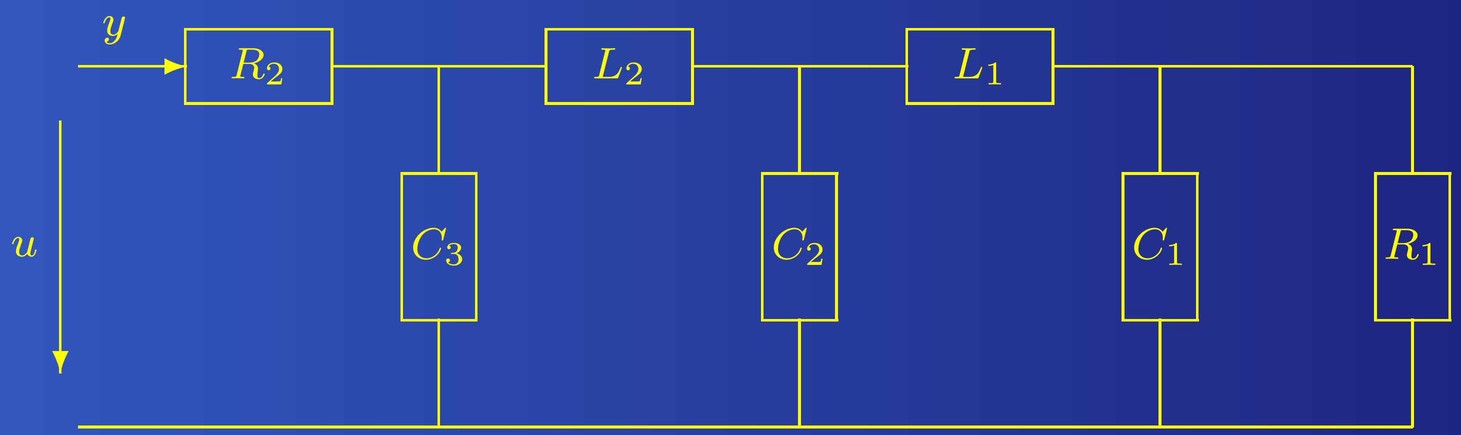
$$\mathbf{V} = X, \quad \bar{\mathbf{W}} = Y$$

Theorem. *The transfer function $\hat{G}$ of the reduced system $(\hat{A}, \hat{B}, \hat{C}, D)$ is positive real and thus the reduced system is both stable and passive.*

Example: RLC Ladder Circuit

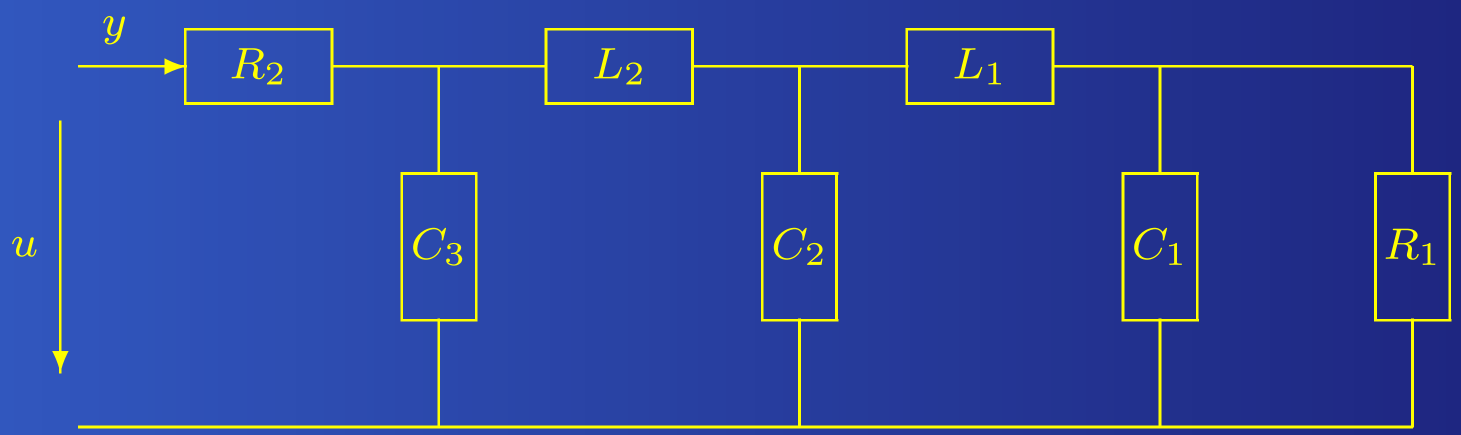


Example: RLC Ladder Circuit



$$A = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & -5 \end{bmatrix}, \ B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \ C = [0 \ \ 0 \ \ 0 \ \ 0 \ \ -2], \ D = 1$$

Example: RLC Ladder Circuit



$$A = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & -5 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \; C = [0 \;\; 0 \;\; 0 \;\; 0 \;\; -2], \; D = 1$$

$$H(s) = \frac{s^5 + 3s^4 + 6s^3 + 9s^2 + 7s + 3}{s^5 + 7s^4 + 14s^3 + 21s^2 + 23s + 7}$$

$$H(s)+H(-s)=\frac{2(s^{10}-s^8-12s^6+5s^4+35s^2-21)}{(s^5+7s^4+14s^3+21s^2+23s+7)(s^5-7s^4+14s^3-21s^2+23s-7)}.$$

The zeros of the stable spectral factor are:

$$\boxed{\begin{array}{c} -.1833 \pm 1.5430i \\ -.7943 \\ -1.3018 \\ -1.8355 \end{array}}$$

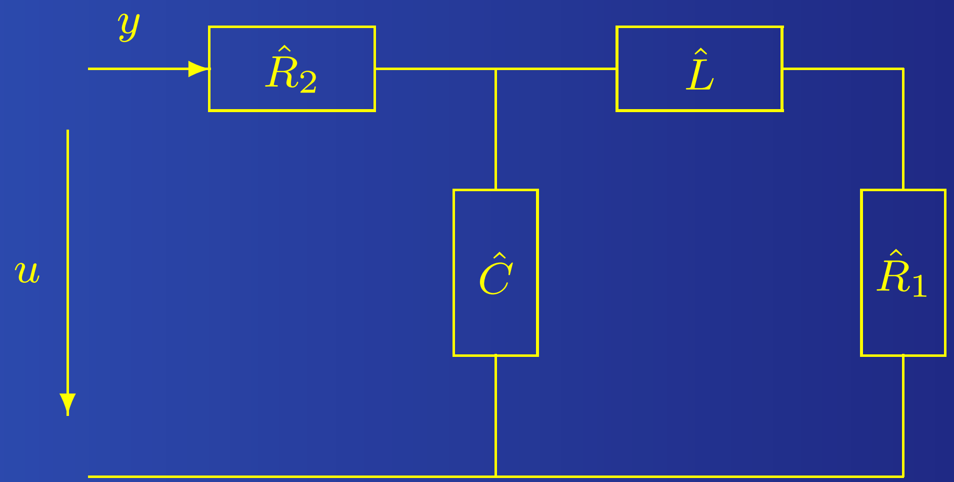
As interpolation points we choose: $\lambda_1 = 1.8355$, $\lambda_2 = -1.8355$, $\lambda_3 = 1.3018$, $\lambda_4 = -1.3018$.

$$\bar{V} = [(\lambda_1 I_5 - A)^{-1}B \quad (\lambda_2 I_5 - A)^{-1}B], \quad \bar{W} = \begin{pmatrix} C(\lambda_3 I_5 - A)^{-1} \\ C(\lambda_4 I_5 - A)^{-1} \end{pmatrix}$$

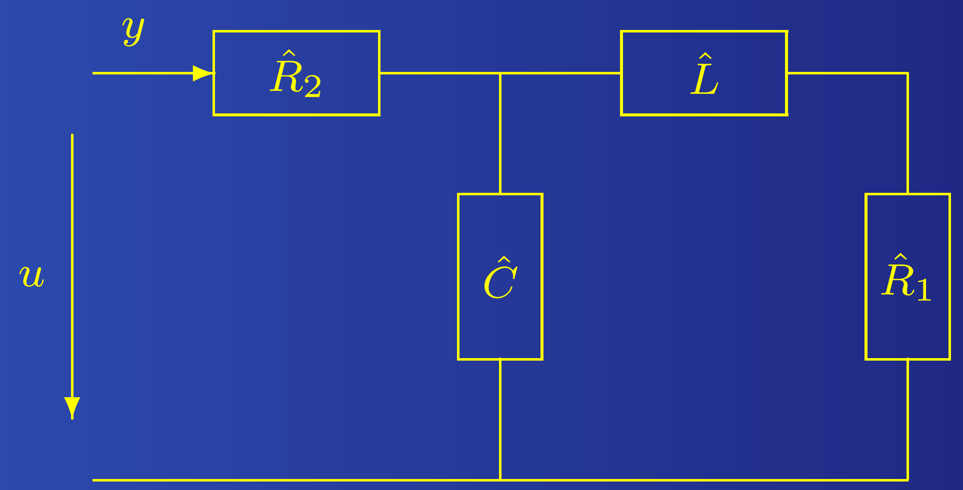
The reduced system is

$$\hat{A} = WAV = -\begin{pmatrix} 3.2923 & 5.0620 \\ 0.9261 & 2.5874 \end{pmatrix}, \quad \hat{B} = WB = -\begin{pmatrix} 1.4161 \\ 0.2560 \end{pmatrix},$$

$\hat{C} = CV = [1.9905 \;\; 5.0620]$, and $\hat{D} = D$. $\Rightarrow \hat{H}(\lambda_i) = H(\lambda_i)$, $i = 1, 2, 3, 4$.



y
$\hat{R}_2$
$\hat{L}$
u
$\hat{C}$
$\hat{R}_1$



The reduced order model can be realized by means of the second-order LRC circuit as shown, with the following values of the parameters:

$$\hat{R}_1 = 4.8258, \; \hat{R}_2 = 1.2907, \; \hat{C} = 0.1459, \; \hat{L} = 8.4796.$$

The corresponding transfer function turns out to be

$$\hat{H}(s) = \frac{0.7784\, s^2 + 0.4409\, s + 0.6263}{s^2 + 5.8797\, s + 3.8306} \, .$$

Optimal $\mathcal{H}_2$ model reduction

Consider a stable single-input/single-output linear system

$$\dot{x}(t) = Ax(t) + Bu(t), \;\; y(t) = Cx(t) \Leftrightarrow \mathbf{\Sigma} = \left(\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right),$$

where $A \in \mathbb{R}^{n\times n}$, $B, C^* \in \mathbb{R}^n$. We seek to produce a system $\mathbf{\Sigma}_r$ of smaller order obtained by *projection*:

$$\dot{\xi}(t) = A_r\xi(t) + B_r u(t), \;\; y_r(t) = C_r\xi(t) \Leftrightarrow \mathbf{\Sigma}_r = \left(\begin{array}{c|c} A_r & B_r \\ \hline C_r & 0 \end{array}\right)$$

where $A_r \in \mathbb{R}^{r\times r}$, $B_r, C_r^* \in \mathbb{R}^r$, with $r < n$, such that:

1. the approximation error $\|\mathbf{\Sigma} - \mathbf{\Sigma}_r\|_{\mathcal{H}_2}$ is *small* (*minimized*)
2. *stability* is preserved
3. method is *numerically efficient*

Krylov projection methods. Given $\Pi = VZ^*$, where $V, Z \in \mathbb{R}^{n\times r}$ satisfy $Z^*V = I_r$, then

$$\boxed{A_r = Z^*AV, \;\; B_r = Z^*B, \;\; C_r = CV}$$

The $\mathcal{H}_2$ norm of the system is:

$$\|\boldsymbol{\Sigma}\|_{\mathcal{H}_2} = \left(\int_{-\infty}^{+\infty} h^*(t)h(t)dt \right)^{1/2}$$

Our goal is be to construct a *Krylov projection method* such that

$$\boldsymbol{\Sigma}_r = \arg \min_{\substack{\deg(\hat{\boldsymbol{\Sigma}}) = r \\ \hat{\boldsymbol{\Sigma}} \,:\, \text{stable}}} \left\| \boldsymbol{\Sigma} - \hat{\boldsymbol{\Sigma}} \right\|_{\mathcal{H}_2}.$$

Facts on moment matching and rational Krylov

- Let

$$V = [(\sigma_1 I - A)^{-1}B \ \cdots \ (\sigma_r I - A)^{-j}B], \quad \sigma_k \in \mathbb{C}, \ k = 1, \cdots, r$$

and Z such that $Z^*V = I_r$; then

$$G(s)|_{s=\sigma_k} = G_r(s)|_{s=\sigma_k}, \ k = 1, \cdots, r.$$

- If we choose

$$\bar{Z} = [(\sigma_1 I - A^*)^{-1}C^* \ \cdots \ (\sigma_r I - A^*)^{-j}C^*]$$

and $Z^* = (\bar{Z}^*V)^{-1}\bar{Z}$, then in addition

$$\left.\frac{dG(s)}{ds}\right|_{s=\sigma_k} = \left.\frac{dG_r(s)}{ds}\right|_{s=\sigma_k}, \quad \text{for} \ \ k = 1, \ldots, r.$$

First-order necessary conditions: Wilson

Let $G_r(s) = \left(\begin{array}{c|c} A_r & B_r \\ \hline C_r & 0 \end{array}\right)$ solve the optimal $\mathcal{H}_2$ problem. Define the error system

$$G_e(s) := G(s) = G_r(s) := \left(\begin{array}{c|c} A_e & B_e \\ \hline C_e & 0 \end{array}\right) := \left(\begin{array}{c|c} \begin{pmatrix} A & 0 \\ 0 & A_r \end{pmatrix} & \begin{pmatrix} B \\ B_r \end{pmatrix} \\ \hline \begin{pmatrix} C & -C_r \end{pmatrix} & 0 \end{array}\right)$$

Let $\mathcal{P}_e$ and $\mathcal{Q}_e$ be the gramians for the error system $G_e(s)$, i.e. $\mathcal{P}_e$ and $\mathcal{Q}_e$ solve

$$A_e\mathcal{P}_e + \mathcal{P}_e A_e^* + B_e B_e^* = 0, \;\; \mathcal{Q}_e A_e + A_e^* \mathcal{Q}_e + C_e^* C_e = 0$$

Partition $\mathcal{P}_e$ and $\mathcal{Q}_e$:

$$\mathcal{P}_e = \begin{pmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{12}^* & \mathcal{P}_{22} \end{pmatrix}, \;\; \mathcal{Q}_e = \begin{pmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12} \\ \mathcal{Q}_{12}^* & \mathcal{Q}_{22} \end{pmatrix}$$

The first-order necessary conditions imply that

$$\begin{aligned} \mathcal{P}_{12}^* \mathcal{Q}_{12} + \mathcal{P}_{22} \mathcal{Q}_{22} &= 0 \\ \mathcal{Q}_{12}^* B + \mathcal{Q}_{22} B_r &= 0 \\ C_r \mathcal{P}_{22} - C \mathcal{P}_{12} &= 0. \end{aligned}$$

Then it follows that, the reduced order model

$$G_r(s) = \left(\begin{array}{c|c} A_r & B_r \\ \hline C_r & 0 \end{array} \right) = \left(\begin{array}{c|c} Z^* A V & Z^* B \\ \hline C V & 0 \end{array} \right),$$

$$\text{where } \boxed{V := \mathcal{P}_{12} \mathcal{P}_{22}^{-1}} \text{ and } \boxed{Z = -\mathcal{Q}_{12} \mathcal{Q}_{22}^{-1}}$$

satisfies the first-order conditions of the optimal $\mathcal{H}_2$ problem (it can be shown that $Z^* V = I_r$).

First-order necessary conditions: Hyland and Bernstein

Suppose $G_r(s) = \left(\begin{array}{c|c} A_r & B_r \\ \hline C_r & 0 \end{array}\right)$ solves the optimal $\mathcal{H}_2$ problem. Then there exist positive semi-definite matrices $\mathcal{P}, \mathcal{Q} \in \mathbb{R}^{n\times n}$:

$$\mathcal{P}\mathcal{Q} = VMZ^*, Z^*V = I_r$$

such that

$$G_r(s) = \left(\begin{array}{c|c} A_r & B_r \\ \hline C_r & 0 \end{array}\right) = \left(\begin{array}{c|c} Z^*AV & Z^*B \\ \hline CV & 0 \end{array}\right),$$

with $\Pi = VZ^*$. Furthermore the following conditions are satisfied:

$$\begin{aligned} \mathrm{rank}(\mathcal{P}) = \mathrm{rank}(\mathcal{Q}) &= \mathrm{rank}(\mathcal{P}\mathcal{Q}) \\ \Pi\left[A\mathcal{P} + \mathcal{P}A^* + BB^*\right] &= 0 \\ \left[A^*\mathcal{Q} + \mathcal{Q}A + C^*C\right]\Pi &= 0 \end{aligned}$$

First-order necessary conditions: Meier

Let $G_r(s) = \left(\begin{array}{c|c} A_r & B_r \\ \hline C_r & 0 \end{array}\right)$ solve the optimal $\mathcal{H}_2$ problem and let $\hat{\lambda}_i$ denote the eigenvalues of A_r (assumed to have multiplicity one). The necessary conditions become

$$\frac{d^k}{ds^k} G(s)\Big|_{s=-\hat{\lambda}_i^*} = \frac{d^k}{ds^k} G_r(s)\Big|_{s=-\hat{\lambda}_i^*}, \quad k = 0, 1.$$

Thus the reduced model has to match the first two moments of the original system at the *mirror images* of the eigenvalues of A_r.

First-order necessary conditions: Meier

Let $G_r(s) = \left(\begin{array}{c|c} A_r & B_r \\ \hline C_r & 0 \end{array}\right)$ solve the optimal $\mathcal{H}_2$ problem and let $\hat{\lambda}_i$ denote the eigenvalues of A_r (assumed to have multiplicity one). The necessary conditions become

$$\frac{d^k}{ds^k} G(s) \Big|_{s=-\hat{\lambda}_i^*} = \frac{d^k}{ds^k} G_r(s) \Big|_{s=-\hat{\lambda}_i^*}, \quad k = 0, 1.$$

Thus the reduced model has to match the first two moments of the original system at the *mirror images* of the eigenvalues of A_r.

It can be shown that these three sets of necessary conditions are equivalent.

An expression for the $\mathcal{H}_2$ norm

New Result. Let $G(s) = \sum_{k=1}^{n} \frac{\phi_k}{s-\lambda_k}$ $\Rightarrow$

$$\|G\|_{\mathcal{H}_2}^2 = \sum_{k=1}^{n} c_k \, G(-\lambda_i^*)$$

An expression for the $\mathcal{H}_2$ norm

New Result. Let $G(s) = \sum_{k=1}^{n} \frac{\phi_k}{s-\lambda_k}$ $\Rightarrow$ $$\|G\|_{\mathcal{H}_2}^2 = \sum_{k=1}^{n} c_k \, G(-\lambda_i^*)$$

Corollary. Let in addition $G_r(s) = \sum_{k=1}^{r} \frac{\hat{\phi}_k}{s-\hat{\lambda}_k}$. The $\mathcal{H}_2$ norm of the error system, is

$$\mathcal{J} = \|G - G_r\|_{\mathcal{H}_2}^2 = \sum_{i=1}^{n} \phi_i \left[G(-\lambda_i) - G_r(-\lambda_i)\right] + \sum_{j=1}^{r} \hat{\phi}_j \left[G_r(-\hat{\lambda}_j) - G(-\hat{\lambda}_j)\right]$$

An expression for the $\mathcal{H}_2$ norm

New Result. Let $G(s) = \sum_{k=1}^{n} \frac{\phi_k}{s-\lambda_k}$ $\Rightarrow$ $\|G\|_{\mathcal{H}_2}^2 = \sum_{k=1}^{n} c_k\, G(-\lambda_i^*)$

Corollary. Let in addition $G_r(s) = \sum_{k=1}^{r} \frac{\hat{\phi}_k}{s-\hat{\lambda}_k}$. The $\mathcal{H}_2$ norm of the error system, is

$$\mathcal{J} = \|G - G_r\|_{\mathcal{H}_2}^2 = \sum_{i=1}^{n} \phi_i \, [G(-\lambda_i) - G_r(-\lambda_i)] + \sum_{j=1}^{r} \hat{\phi}_j \left[G_r(-\hat{\lambda}_j) - G(-\hat{\lambda}_j)\right]$$

Conclusion. The $\mathcal{H}_2$ error is due to the *mismatch* of the transfer functions $G - G_r$ at the *mirror images* of the full-order and reduced system poles λ_i, $\hat{\lambda}_i$.

A new derivation of the first-order conditions

Let

$$G(s) = \sum_{k=1}^{n} \frac{\phi_k}{s-\lambda_k} \text{ and } G_r(s) = \sum_{k=1}^{r} \frac{\hat{\phi}_k}{s-\hat{\lambda}_k}$$

denote the full-order and reduced-order models, respectively. For simplicity, we assume that both $G(s)$ and $G_r(s)$ have distinct poles.

The Following string of equalities holds:

$$G(-\lambda_i) = \sum_{k=1}^{n} \frac{\phi_k}{-\lambda_i-\lambda_k},\ i=1,\ldots,n \quad \text{and} \quad G(-\hat{\lambda}_j) = \sum_{k=1}^{n} \frac{\phi_k}{-\hat{\lambda}_j-\lambda_k},\ j=1,\ldots,r,$$

$$G_r(-\lambda_i) = \sum_{k=1}^{r} \frac{\hat{\phi}_k}{-\lambda_i-\hat{\lambda}_k},\ i=1,\ldots,n \quad \text{and} \quad G_r(-\hat{\lambda}_j) = \sum_{k=1}^{r} \frac{\hat{\phi}_k}{-\hat{\lambda}_j-\hat{\lambda}_k},\ j=1,\ldots,r,$$

$$\frac{\partial}{\partial\hat{\phi}_m} G(-\lambda_i) = 0 \quad \text{and} \quad \frac{\partial}{\partial\hat{\phi}_m} G(-\hat{\lambda}_j) = 0,\ m=1,\ldots,r,$$

$$\frac{\partial}{\partial\hat{\phi}_m} G_r(-\lambda_i) = \frac{1}{-\lambda_i-\hat{\lambda}_m}, \quad \text{and} \quad \frac{\partial}{\partial\hat{\phi}_m} G_r(-\hat{\lambda}_j) = \frac{1}{-\hat{\lambda}_j-\hat{\lambda}_m},\ m=1,\ldots,r,$$

To find the first-order necessary conditions, we differentiate the cost function $\mathcal{J}$ with respect to the reduced-order parameters $\hat{\phi}_m$ and $\hat{\lambda}_m$, $i = 1, \ldots, r$.

$$\frac{\partial \mathcal{J}}{\partial \hat{\phi}_m} = \underbrace{\sum_{k=1}^{n} \phi_i \left(\frac{-1}{-\lambda_i - \hat{\lambda}_m} \right)}_{=-G(-\hat{\lambda}_m)} + G_r(-\hat{\lambda}_m) - G(-\hat{\lambda}_m) + \underbrace{\sum_{j=1}^{r} \hat{\phi}_j \frac{1}{-\hat{\lambda}_j - \hat{\lambda}_m}}_{=G_r(-\hat{\lambda}_m)}$$

$$= -2G(-\hat{\lambda}_m) + 2G_r(-\hat{\lambda}_m), \quad m = 1, \ldots, r.$$

Therefore, setting $\frac{\partial \mathcal{J}}{\partial \hat{\phi}_m}$ to zero leads to $G(-\hat{\lambda}_m) = G_r(-\hat{\lambda}_m), \;\; m = 1, \ldots, r$. These are half of Meier's first-order conditions.

For the remaining ones we need to differentiate $\mathcal{J}$ with respect to $\hat{\lambda}_m$.

$$\frac{\partial \mathcal{J}}{\partial \hat{\lambda}_m} = -\hat{\phi}_m \left[\underbrace{\sum_{i=1}^{n} \frac{\phi_i}{(\lambda_i + \hat{\lambda}_m)^2}}_{= -\frac{dG}{ds}|_{s=-\hat{\lambda}_m}} + \underbrace{\sum_{k=1}^{n} \frac{\phi_k}{(\hat{\lambda}_m + \lambda_k)^2}}_{= -\frac{dG}{ds}|_{s=-\hat{\lambda}_m}} \right] + \hat{\phi}_m \left[\underbrace{\sum_{k=1}^{r} \frac{\hat{\phi}_k}{(\hat{\lambda}_m + \hat{\lambda}_k)^2}}_{= -\frac{dG_r}{ds}|_{s=-\hat{\lambda}_m}} + \underbrace{\sum_{j=1}^{r} \frac{\hat{\phi}_j}{(\hat{\lambda}_m + \hat{\lambda}_j)^2}}_{= -\frac{dG_r}{ds}|_{s=-\hat{\lambda}_m}} \right]$$

$$= \quad 2\hat{\phi}_m \left(\left.\frac{dG}{ds}\right|_{s=-\hat{\lambda}_m} - \left.\frac{dG_r}{ds}\right|_{s=-\hat{\lambda}_m} \right)$$

Hence, letting $\frac{\partial \mathcal{J}}{\partial \hat{\lambda}_m} = 0$ leads to the remaining first-order conditions.

An optimality condition

Given a stable rational function $G(s)$, with fixed stable reduced poles $\alpha_1, \ldots, \alpha_r$, define

$$G_r(s) := \frac{\beta_0 + \beta_1 s + \cdots + \beta_r s^r}{(s-\alpha_1)\ldots(s-\alpha_r)}.$$

Gaier showed that $\|G - G_r\|_{\mathcal{H}_2}$ is minimized if and only if

$$G(s) = G_r(s) \quad \text{for} \quad s = -\alpha_1^*, -\alpha_2^*, \ldots, -\alpha_r^*.$$

Thus if $G_r(s)$ interpolates $G(s)$ at the mirror images of the poles of $G_r(s)$, then $G_r(s)$ is guaranteed to be an *optimal* approximant of $G(s)$ with respect to the $\mathcal{H}_2$ norm among all reduced order systems having the same reduced system poles $\{\alpha_i\}$, $i = 1, \ldots, r$. Therefore, our algorithm generates a reduced model $G_r(s)$ which is the optimal solution for a restricted $\mathcal{H}_2$ problem.

An Iterative Rational Krylov Algorithm (IRKA)

The proposed algorithm produces a reduced order model $G_r(s)$ that satisfies the interpolation-based conditions, i.e.

$$\frac{d^k}{ds^k}G(s)\Big|_{s=-\hat{\lambda}_i} = \frac{d^k}{ds^k}G_r(s)\Big|_{s=-\hat{\lambda}_i}, \quad k=0,1.$$

Since the interpolation points depend on the poles of the reduced system and are not a priori known, we propose to run the rational Krylov method iteratively; at the $(i+1)^{\mathrm{st}}$ step the interpolation points are chosen as the *mirror images* of the eigenvalues of A_r obtained at the i^{th} step.

1. Make an initial selection of σ_i, for $i=1,\dots,r$
2. $\bar{Z} = [(\sigma_1 I - A^*)^{-1}C^*, \cdots, (\sigma_r I - A^*)^{-1}C^*]$.
3. $V = [(\sigma_1 I - A)^{-1}B, \cdots, (\sigma_r I - A)^{-1}B]$, and let $Z^* = (\bar{Z}^* V)^{-1}\bar{Z}$.
4. while (not converged)
   (a) $A_r = Z^* A V$,
   (b) $\sigma_i \longleftarrow -\lambda_i(A_r)$ for $i=1,\dots,r$
   (c) $\bar{Z} = [(\sigma_1 I - A^*)^{-1}C^*, \cdots, (\sigma_r I - A^*)^{-1}C^*]$
   (d) $V = [(\sigma_1 I - A)^{-1}B, \cdots, (\sigma_r I - A)^{-1}B]$, $Z^* = (\bar{Z}^* V)^{-1}\bar{Z}$.
5. $A_r = Z^* A V$, $B_r = Z^* B$, $C_r = CV$

Numerical examples: low order models

- FOM-1: Example 6.1 of Hyland and Bernstein:

$$A = \begin{bmatrix} 0 & 0 & 0 & -150 \\ 1 & 0 & 0 & -245 \\ 0 & 1 & 0 & -113 \\ 0 & 0 & 1 & -19 \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad C^* = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

  We reduce the order to $r = 3, 2, 1$ using the proposed algorithm **IRKA** and compare our results with the gradient flow method of Yan **GFM**, the orthogonal projection method of Hyland **OPM**, and the balanced truncation method **BTM**.

- FOM-2: Example in Lepschy. Transfer function of FOM-2 is given by

$$G(s) = \frac{2s^6 + 11.5s^5 + 57.75s^4 + 178.625s^3 + 345.5s^2 + 323.625s + 94.5}{s^7 + 10s^6 + 46s^5 + 130s^4 + 239s^3 + 280s^2 + 194s + 60}$$

  We reduce the order to $r = 6, 5, 4, 3$ using **IRKA**, and compare our results with **GFM**; **OPM**; **BTM**; and the method proposed in Lepschy, denoted by **LMPV**.

- FOM-3: Example 1 in Spanos. Transfer function of FOM-3 is given by

$$G(s) = \frac{s^2 + 15s + 50}{s^4 + 5s^3 + 33s^2 + 79s + 50}$$

  We reduce the order to $r = 3, 2, 1$ using **IRKA**, and compare our results with **GFM**; **OPM**; **BTM**; and the method proposed in Spanos denoted by **SMM**.

- FOM-4: Example 2 in Spanos. Transfer function of FOM-4 is given by

$$G(s) = \frac{10000s + 5000}{s^2 + 5000s + 25}$$

The resulting relative $\mathcal{H}_2$ errors $\frac{\|G(s)-G_r(s)\|_{\mathcal{H}_2}}{\|G(s)\|_{\mathcal{H}_2}}$ are tabulated below.

Model	r	IRKA	GFM	OPM	BTM	LMPV	SMM
FOM-1	1	4.2683×10^{-1}	4.2709×10^{-1}	4.2683×10^{-1}	4.3212×10^{-1}		
FOM-1	2	3.9290×10^{-2}	3.9299×19^{-2}	3.9290×10^{-2}	3.9378×10^{-2}		
FOM-1	3	1.3047×10^{-3}	1.3107×19^{-3}	1.3047×10^{-3}	1.3107×10^{-3}		
FOM-2	3	1.171×10^{-1}	1.171×10^{-1}	Divergent	2.384×10^{-1}	1.171×10^{-1}	
FOM-2	4	8.199×10^{-3}	8.199×10^{-3}	8.199×10^{-3}	8.226×10^{-3}	8.199×10^{-3}	
FOM-2	5	2.132×10^{-3}	2.132×10^{-3}	Divergent	2.452×10^{-3}	2.132×10^{-3}	
FOM-2	6	5.817×10^{-5}	5.817×10^{-5}	5.817×10^{-5}	5.822×10^{-5}	2.864×10^{-4}	
FOM-3	1	4.818×10^{-1}	4.818×10^{-1}	4.818×10^{-1}	4.848×10^{-1}		4.818×10^{-1}
FOM-3	2	2.443×10^{-1}	2.443×10^{-1}	Divergent	3.332×10^{-1}		2.443×10^{-1}
FOM-3	3	5.74×10^{-2}	5.98×10^{-2}	5.74×10^{-2}	5.99×10^{-2}		5.74×10^{-2}
FOM-4	1	9.85×10^{-2}	9.85×10^{-2}	9.85×10^{-2}	9.949×10^{-1}	9.985×10^{-2}	

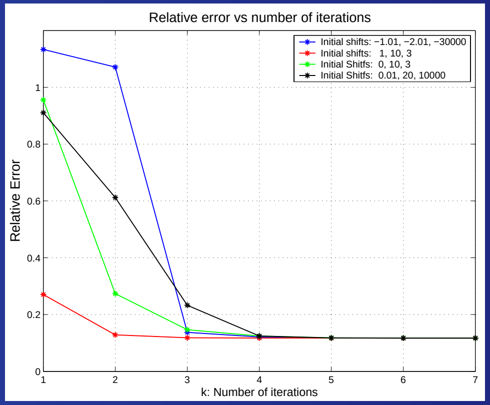
Thus the proposed method is the only one which attains the minimum in each case.

To illustrate the evolution of the $\mathcal{H}_2$ error throughout the iteration, consider the model FOM-5 with $r = 3$. The proposed method yields the following third order optimal reduced model:

$$G_3(s) = \frac{2.155s^2 + 3.343s + 33.8}{s^3 + 7.457s^2 + 10.51s + 17.57}.$$

Poles of $G_3(s)$ are $\hat{\lambda}_1 = -6.2217$ and $\hat{\lambda}_{2,3} = -6.1774 \times 10^{-1} \pm \jmath 1.5628$; it can be shown that $G_3(s)$ interpolates the first two moments of $G(s)$ at $-\hat{\lambda}_i$, for $i = 1, 2, 3$. Hence, the first-order interpolation conditions are satisfied. However, we try four random, but *bad*, initial selections (i.e. we start away from the optimal solution):

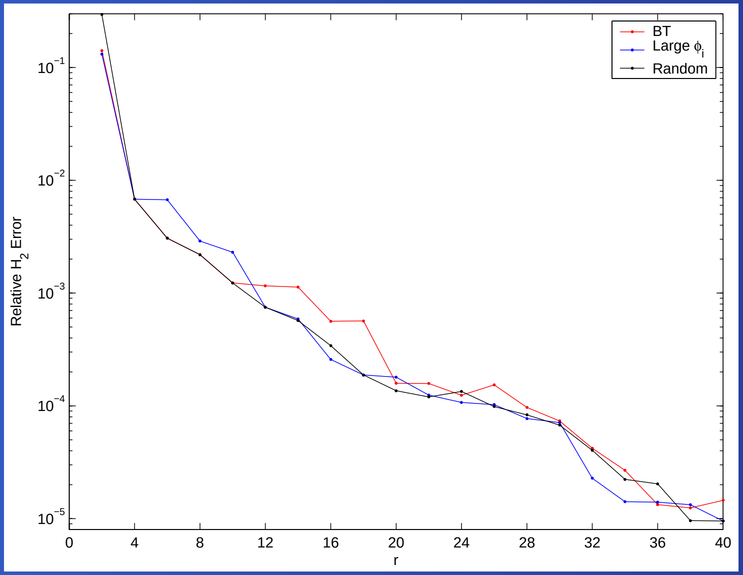
$$\begin{aligned} \mathcal{S}_1 &= \{-1.01,\ -2.01,\ -30000\} \\ \mathcal{S}_2 &= \{0,\ 10,\ 3\} \\ \mathcal{S}_3 &= \{1,\ 10,\ 3\} \\ \mathcal{S}_4 &= \{0.01,\ 20,\ 10000\} \end{aligned}$$

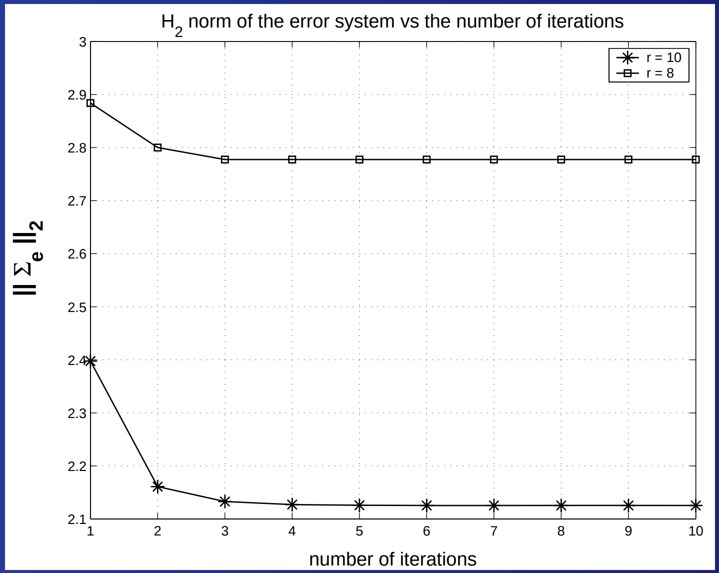


$\mathcal{H}_2$ norm of the error system vs the number of iterations

The CD Player Example

- CD player model: $n = 120$, $m = p = 1$.
- The Hankel singular values do not decay rapidly $\Rightarrow$ model is hard to reduce.
- The performance of the proposed method is compared with balanced truncation for $r = 2, \cdots, 40$.

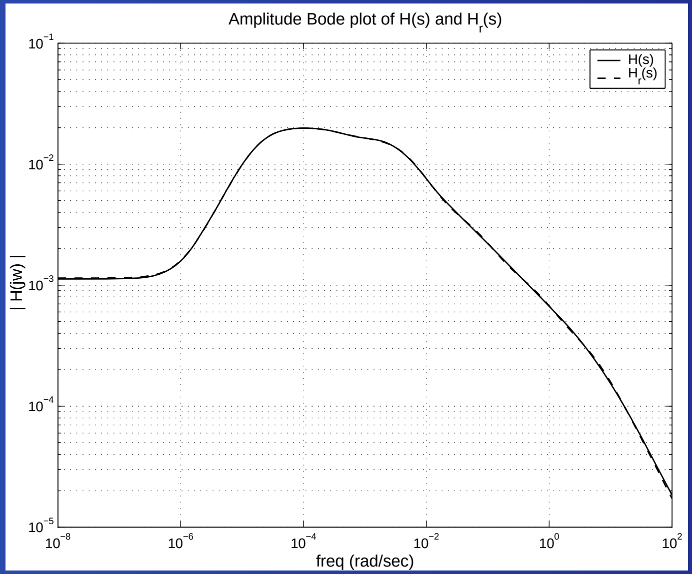




Relative $\mathcal{H}_2$ norm of the error system vs r (left) and the cases $r = 8$, $r = 10$ vs the number of iterations (right)

A heat transfer problem

The problem of cooling steel profiles is modeled as boundary control of a two dimensional heat equation. A finite element discretization with mesh width of 1.382×10^{-2} results in a system with state dimension $n = 20209$, $m = 7$ inputs and $p = 6$ outputs.



We consider the full-order SISO system that associates the sixth input of this system with the second output. The reduced order is $r = 6$. **IRKA** converged in 7 iteration steps and the relative $\mathcal{H}_\infty$ error is 7.85×10^{-3}.

Hankel approximation: Decay rate of HSVs

A new set of system invariants. Let the transfer function be

$$G(s) = C(sI - A)^{-1}B + D = \frac{p(s)}{q(s)}, \quad \lambda_i(A) \ \text{distinct}$$

Hankel singular values $\sigma_i(\mathbf{\Sigma}) = \sqrt{\lambda_i(\mathcal{PQ})}$, $i = 1, \cdots, n$, $\sigma_i \geq \sigma_{i+1}$. With $q(s)^* = q^*(-s) = \sum_{k=0}^{n} \alpha_k^*(-s)^k$, define:

$$\gamma_i = \frac{p(\lambda_i)}{q^*(\lambda_i)} = \left.\frac{p(s)}{q(s)^*}\right|_{s=\lambda_i}, \quad |\gamma_i| \geq |\gamma_{i+1}|, \ i = 1, \cdots, n-1.$$

Hankel approximation: Decay rate of HSVs

A new set of system invariants. Let the transfer function be

$$G(s) = C(sI - A)^{-1}B + D = \frac{p(s)}{q(s)}, \quad \lambda_i(A) \ \text{distinct}$$

Hankel singular values $\sigma_i(\mathbf{\Sigma}) = \sqrt{\lambda_i(\mathcal{P}\mathcal{Q})}$, $i = 1, \cdots, n$, $\sigma_i \geq \sigma_{i+1}$. With $q(s)^* = q^*(-s) = \sum_{k=0}^{n} \alpha_k^*(-s)^k$, define:

$$\gamma_i = \frac{p(\lambda_i)}{q^*(\lambda_i)} = \left.\frac{p(s)}{q(s)^*}\right|_{s=\lambda_i}, \quad |\gamma_i| \geq |\gamma_{i+1}|, \ i = 1, \cdots, n-1.$$

Theorem. **Decay rate of the Hankel singular values**. The vector of Hankel singular values σ **majorizes** the absolute value of the vector of new invariants γ **multiplicatively**:

$$|\gamma| \prec_\mu \sigma$$

Hankel approximation: the Hankel operator

Recall that

$$\mathcal{L}_2(j\mathbb{R}) = \mathcal{H}_2(\mathbb{C}_-) \oplus \mathcal{H}_2(\mathbb{C}_+)$$

where $\mathcal{H}_2(\mathbb{C}_-)$, $\mathcal{H}_2(\mathbb{C}_+)$ contain functions which are analytic in $\mathbb{C}_-$, $\mathbb{C}_+$ respectively. The projections onto $\mathcal{H}_2(\mathbb{C}_-)$, $\mathcal{H}_2(\mathbb{C}_+)$ are denoted by P_-, P_+. Given $\phi \in \mathcal{L}_2(j\mathbb{R})$, we use ϕ^* to denote

$$\phi(s)^* = \phi^*(-s)$$

The *Hankel operator with (scalar) symbol* $\phi \in \mathcal{L}_2(j\mathbb{R})$ is defined as:

$$\mathcal{H}_\phi : \mathcal{H}_2(\mathbb{C}_-) \to \mathcal{H}_2(\mathbb{C}_+), \;\; f \mapsto \mathcal{H}_\phi f = P_+(\phi f).$$

The *dual Hankel operator* is defined as:

$$\mathcal{H}_\phi^* : \mathcal{H}_2(\mathbb{C}_+) \to \mathcal{H}_2(\mathbb{C}_-), \;\; \mathbf{f} \mapsto \mathcal{H}_\phi^* f = P_-(\phi^* f).$$

In the sequel we consider Hankel operators with stable, rational symbol:

$$\mathcal{H}_\phi, \; \phi = \frac{n}{d} \in \mathcal{H}_2(\mathbb{C}_+), \; \gcd(n, d) = 1$$

The space X^d of all strictly proper rational functions with denominator d: $X^d = \left\{ \frac{a}{d} \; : \; \deg a < \deg d \right\}$, is a finite-dimensional linear space with dimension $\deg d$. The *shift operator* in this space is defined as $F^d(h(s)) = \pi_-\left[\frac{1}{d(s)}\pi_+\left[\frac{d(s)}{s}h(s)\right]\right]$, where π_+, π_- are projections on to the polynomials, strictly proper rational functions, respectively. It readily follows that the characteristic polynomial of this shift operator is d.

The Hankel operator and its SVD

Consider a Hankel operator $\mathcal{H}_\phi$ and its adjoint $\mathcal{H}_\phi^*$.

1. *The kernel of this Hankel operator is:* $\ker \mathcal{H}_\phi = \frac{d}{d^*}\mathcal{H}_2(\mathbb{C}_+)$
2. *The kernel of the adjoint Hankel operator is:* $\ker \mathcal{H}_\phi^* = \frac{d^*}{d}\mathcal{H}_2(\mathbb{C}_-)$
3. *The image of* $\mathcal{H}_\phi$ *is:* $\mathrm{im}\, \mathcal{H}_\phi = X^d$
4. *The image of the adjoint operator is:* $\mathrm{im}\, \mathcal{H}_\phi^* = X^{d^*}$

A pair of functions $f \in \mathcal{H}_2(\mathbb{C}_-)$, $g \in \mathcal{H}_2(\mathbb{C}_+)$ is a Schmidt pair of $\mathcal{H}_\phi$ corresponding to the singular value σ, provided that the following relationships are satisfied

$$\mathcal{H}_\phi f = \sigma g, \; \mathcal{H}_\phi^* g = \sigma f.$$

The Hankel operator and its SVD

Consider a Hankel operator $\mathcal{H}_\phi$ and its adjoint $\mathcal{H}_\phi^*$.

1. *The kernel of this Hankel operator is:* $\ker \mathcal{H}_\phi = \frac{d}{d*}\mathcal{H}_2(\mathbb{C}_+)$
2. *The kernel of the adjoint Hankel operator is:* $\ker \mathcal{H}_\phi^* = \frac{d*}{d}\mathcal{H}_2(\mathbb{C}_-)$
3. *The image of* $\mathcal{H}_\phi$ *is:* $\mathrm{im}\, \mathcal{H}_\phi = X^d$
4. *The image of the adjoint operator is:* $\mathrm{im}\, \mathcal{H}_\phi^* = X^{d^*}$

A pair of functions $f \in \mathcal{H}_2(\mathbb{C}_-)$, $g \in \mathcal{H}_2(\mathbb{C}_+)$ is a Schmidt pair of $\mathcal{H}_\phi$ corresponding to the singular value σ, provided that the following relationships are satisfied

$$\mathcal{H}_\phi f = \sigma g, \; \mathcal{H}_\phi^* g = \sigma f.$$

Since $g \in \mathrm{im}\, \mathcal{H}_\phi$ and $f \in \mathrm{im}\, \mathcal{H}_\phi^*$, we obtain the following equations:

$$P_+ \frac{n}{d}\frac{p}{d*} = \sigma \frac{\hat{p}}{d} \quad \Rightarrow \quad \frac{n}{d}\frac{p}{d*} = \sigma \frac{\hat{p}}{d} + \frac{\pi}{d*} \quad \Rightarrow \quad \boxed{np = \sigma d^* \hat{p} + d\pi}$$

$$P_- \frac{n^*}{d*}\frac{\hat{p}}{d} = \sigma \frac{p}{d*} \quad \Rightarrow \quad \frac{n^*}{d*}\frac{\hat{p}}{d} = \sigma \frac{p}{d*} + \frac{\xi}{d} \quad \Rightarrow \quad \boxed{n^* \hat{p} = \sigma d p + d^* \xi}$$

The quantities p, $\hat{p}$, π, ξ are polynomials having degree less than the degree of d.

The Hankel operator and its SVD

Consider a Hankel operator $\mathcal{H}_\phi$ and its adjoint $\mathcal{H}_\phi^*$.

1. *The kernel of this Hankel operator is:* $\ker \mathcal{H}_\phi = \frac{d}{d*}\mathcal{H}_2(\mathbb{C}_+)$
2. *The kernel of the adjoint Hankel operator is:* $\ker \mathcal{H}_\phi^* = \frac{d^*}{d}\mathcal{H}_2(\mathbb{C}_-)$
3. *The image of* $\mathcal{H}_\phi$ *is:* $\mathrm{im}\, \mathcal{H}_\phi = X^d$
4. *The image of the adjoint operator is:* $\mathrm{im}\, \mathcal{H}_\phi^* = X^{d^*}$

A pair of functions $f \in \mathcal{H}_2(\mathbb{C}_-)$, $g \in \mathcal{H}_2(\mathbb{C}_+)$ is a Schmidt pair of $\mathcal{H}_\phi$ corresponding to the singular value σ, provided that the following relationships are satisfied

$$\mathcal{H}_\phi f = \sigma g, \ \mathcal{H}_\phi^* g = \sigma f.$$

Since $g \in \mathrm{im}\, \mathcal{H}_\phi$ and $f \in \mathrm{im}\, \mathcal{H}_\phi^*$, we obtain the following equations:

$$P_+ \frac{n}{d}\frac{p}{d*} = \sigma \frac{\hat{p}}{d} \quad \Rightarrow \quad \frac{n}{d}\frac{p}{d*} = \sigma \frac{\hat{p}}{d} + \frac{\pi}{d*} \quad \Rightarrow \quad \boxed{np = \sigma d^* \hat{p} + d\pi}$$

$$P_- \frac{n^*}{d*}\frac{\hat{p}}{d} = \sigma \frac{p}{d*} \quad \Rightarrow \quad \frac{n^*}{d*}\frac{\hat{p}}{d} = \sigma \frac{p}{d*} + \frac{\xi}{d} \quad \Rightarrow \quad \boxed{n^* \hat{p} = \sigma d p + d^* \xi}$$

The quantities p, $\hat{p}$, π, ξ are polynomials having degree less than the degree of d.

$$\deg p, \ \deg \hat{p}, \ \deg \pi, \ \deg \xi < \deg d = \nu.$$

$\boxed{\hat{p} = \epsilon p^*, \ \epsilon = \pm 1}$ $\Rightarrow$ Schmidt pairs are $\left\{ \frac{p}{d*}, \epsilon \frac{p^*}{d} \right\}$, $\epsilon = \pm 1$, and $\deg p < \deg d$.

The fundamental equation

$\Rightarrow \sigma$ is a singular value if and only if

$$np = \lambda d^* p^* + d\pi, \quad \lambda = \epsilon\sigma$$

where $\deg \pi = \deg p$. This equation can be converted into a matrix equation.

The fundamental equation

$\Rightarrow$ σ is a singular value if and only if

$$np = \lambda d^* p^* + d\pi, \quad \lambda = \epsilon\sigma$$

where $\deg \pi = \deg p$. This equation can be converted into a matrix equation.

Let e_i be a basis of the space X_d of all polynomials of degree less than $\deg d$, and F_d be the shift in X_d (its characteristic polynomial is d). $\underline{p}$, $\underline{p^*}$ denote the vector representations of p, p^* respectively in the basis e_i. Then we obtain the matrix equation

$$n(F_d)\underline{p} = \lambda d^*(F_d)\underline{p^*}.$$

The fundamental equation

$\Rightarrow$ σ is a singular value if and only if

$$np = \lambda d^* p^* + d\pi, \quad \lambda = \epsilon\sigma$$

where $\deg \pi = \deg p$. This equation can be converted into a matrix equation.

Let e_i be a basis of the space X_d of all polynomials of degree less than $\deg d$, and F_d be the shift in X_d (its characteristic polynomial is d). $\underline{p}$, $\underline{p^*}$ denote the vector representations of p, p^* respectively in the basis e_i. Then we obtain the matrix equation

$$n(F_d)\underline{p} = \lambda d^*(F_d)\underline{p^*}.$$

Since d, d^* are coprime, there exist polynomials a, b such that $ad + bd^* = 1$. Thus the inverse of $d^*(F_d)$ is $b(F_d)$. This implies $b(F_d)n(F_d)\underline{p} = \lambda \underline{p^*}$. Finally let K be the map such that $Kp^* = p$; the matrix representation of K is a sign matrix. Hence, with $\nu = \deg d$, there holds

The fundamental equation

$\Rightarrow \sigma$ is a singular value if and only if

$$np = \lambda d^* p^* + d\pi, \quad \lambda = \epsilon\sigma$$

where $\deg \pi = \deg p$. This equation can be converted into a matrix equation.

Let e_i be a basis of the space X_d of all polynomials of degree less than $\deg d$, and F_d be the shift in X_d (its characteristic polynomial is d). $\underline{p}$, $\underline{p^*}$ denote the vector representations of p, p^* respectively in the basis e_i. Then we obtain the matrix equation

$$n(F_d)\underline{p} = \lambda d^*(F_d)\underline{p^*}.$$

Since d, d^* are coprime, there exist polynomials a, b such that $ad + bd^* = 1$. Thus the inverse of $d^*(F_d)$ is $b(F_d)$. This implies $b(F_d)n(F_d)\underline{p} = \lambda \underline{p^*}$. Finally let K be the map such that $K\underline{p^*} = \underline{p}$; the matrix representation of K is a sign matrix. Hence, with $\nu = \deg d$, there holds

$$M\underline{p} = \lambda \underline{p} \quad \text{where} \quad M = Kb(F_d)n(F_d) \in \mathbb{R}^{\nu\times\nu} \quad \text{and} \quad \underline{p} \in \mathbb{R}^{\nu}$$

This eigenvalue equation has to be solved in order to compute the singular values and singular vectors of the Hankel operator $\mathcal{H}_\phi$.

A consequence of the fundamental equation

$$np = \lambda d^* p^* + d\pi, \quad \lambda = \epsilon\sigma \; \Rightarrow \left\{ \begin{array}{l} \dfrac{n}{d} - \dfrac{\pi}{p} = \lambda \cdot \overbrace{\dfrac{d^*}{d} \cdot \dfrac{p^*}{p}}^{all-pass\ dilation} \\[2ex] \dfrac{n}{d^*} = \lambda \dfrac{p^*}{p} + \dfrac{\pi}{p} \cdot \dfrac{d}{d^*} \end{array} \right.$$

A consequence of the fundamental equation

$$np = \lambda d^* p^* + d\pi, \quad \lambda = \epsilon\sigma \;\Rightarrow\; \begin{cases} \dfrac{n}{d} - \dfrac{\pi}{p} = \lambda \cdot \overbrace{\dfrac{d^*}{d} \cdot \dfrac{p^*}{p}}^{all-pass\,dilation} \\[2ex] \dfrac{n}{d^*} = \lambda \dfrac{p^*}{p} + \dfrac{\pi}{p} \cdot \dfrac{d}{d^*} \end{cases}$$

$$\Downarrow$$

$$\left.\frac{n}{d^*}\right|_{s=\lambda_i} = \lambda \left.\frac{p^*}{p}\right|_{s=\lambda_i}$$

A consequence of the fundamental equation

$$np = \lambda d^* p^* + d\pi, \quad \lambda = \epsilon\sigma \;\Rightarrow\; \begin{cases} \dfrac{n}{d} - \dfrac{\pi}{p} = \lambda \cdot \overbrace{\dfrac{d^*}{d} \cdot \dfrac{p^*}{p}}^{all-pass\,dilation} \\ \dfrac{n}{d^*} = \lambda \dfrac{p^*}{p} + \dfrac{\pi}{p} \cdot \dfrac{d}{d^*} \end{cases}$$

$$\Downarrow$$

$$\left.\frac{n}{d^*}\right|_{s=\lambda_i} = \lambda \left.\frac{p^*}{p}\right|_{s=\lambda_i}$$

Thus the all-pass dilation of the original system
can be obtained by rational interpolation
at the mirror image of the poles.

The all-pass dilation: MIMO case – Auba and Funahashi

Given is the system with n states, $m = p \geq 1$ inputs and outputs:

$$G(s) = \left(\begin{array}{c|c} \Lambda & L^* \\ \hline W & 0 \end{array}\right), \ \Lambda = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_n), \ L = [\ell_1, \cdots, \ell_n], \ W = [w_1, \cdots, w_n],$$

where $\lambda_i \neq \lambda_j$, $i \neq j$. Let Γ_L be the reachability gramian

$$\Lambda \Gamma_L + \Gamma_L \Lambda^* + L^* L = 0.$$

The following system is all-pass (unitary) with poles the *mirror images* of the poles of the original system:

$$U(s) = \left(\begin{array}{c|c} -\Lambda^* & \Gamma_L^{-1} L^* \\ \hline -L & I \end{array}\right) \ \Rightarrow \ U(s)U^*(-s) = I_m$$

It follows that $G(s) = N(s)U^{-1}(s) = N(s)U^*(-s)$, where

$$N(s) = \left(\begin{array}{c|c} -\Lambda^* & \Gamma_L^{-1} \\ \hline W\Gamma_L & 0 \end{array}\right)$$

Notice that $\ell_i U(\lambda_i) = 0$. Define $V = L\Gamma_L^{-1}$. Then $U(\lambda_i)v_i = 0$. It turns out that

$$\boxed{w_i = N(\lambda_i)v_i,\ i = 1,\cdots,n}$$

The associated interpolation problem is a *right tangential interpolation problem*:

$$(\lambda_i;\ v_i,\ w_i),\ i = 1,\cdots,n,\ \ \text{where}\ \ \lambda_i \in \mathbb{C},\ v_i, w_i \in \mathbb{C}^m$$

Problem. Find an all-pass matrix rational function $\Xi(s)$ which satisfies $\Xi(\lambda_i)v_i = w_i$.

Observation. $\Xi(s)$ is all-pass $\Rightarrow v_j^* = w_j^*\Xi(-\lambda_j^*)$.

Consequence. Ξ is defined by means of left and right tangential interpolation conditions:

$$\boxed{(-\lambda_i^*;\ w_j^*,\ v_j^*)\ \ \text{and}\ \ (\lambda_i;\ v_i,\ w_i),\ i = 1,\cdots,n.}$$

Next, define the associated Pick (or Löwner) matrix

$$\Pi_\mu = \left[\frac{v_i^* v_j - \mu^{-2} w_i^* w_j}{\lambda_i^* + \lambda_j}\right]_{i,j}$$

and $\Phi = V - \mu^{-1}W$.

The following is an all-pass system

$$\Xi(s) = \left(\begin{array}{c|c} \Lambda - \Pi_\mu^{-1}\Phi^* V & \Pi_\mu^{-1}\Phi^* \\ \hline -\Phi & I \end{array}\right) \Rightarrow \boxed{\Xi(s)\Xi^*(-s) = I_m}$$

which satisfies the bi-tangential interpolation constraints.

The all-pass dilation of the given system is

$$G_e(s) = \Xi(s)U^*(-s) = \left[\begin{array}{c|c} \begin{pmatrix} \Lambda - \Pi_\mu^{-1}\Phi^* V & 0 \\ 0 & \Lambda \end{pmatrix} & \begin{pmatrix} \Pi_\mu^{-1}\Phi^* - \Gamma_V^{-1} V^* \\ \Gamma_V^{-1} \end{pmatrix} \\ \hline \begin{pmatrix} -\mu\Phi & W \end{pmatrix} & \mu I_m \end{array}\right]$$

where $\Lambda^*\Gamma_V + \Gamma_V\Lambda + V^*V = 0$. Thus

$$\boxed{\hat{G}(s) = P_-(G_e(s) - G(s)) = P_-\left(\left[\begin{array}{c|c} \Lambda - \Pi_\mu^{-1}\Phi^* V & \Pi_\mu^{-1}\Phi^* - \Gamma_V^{-1} V^* \\ \hline \mu\Phi & -\mu I_m \end{array}\right]\right)}$$

Summary

•	Positive real interpolation and model reduction	mirror image of spectral zeros
•	Optimal H_2 model reduction	mirror image of reduced system poles
•	Hankel norm model reduction	mirror image of system poles

Summary

•	Positive real interpolation and model reduction	mirror image of spectral zeros
•	Optimal H_2 model reduction	mirror image of reduced system poles
•	Hankel norm model reduction	mirror image of system poles

Question: What is the deeper meaning of **mirror image** points?

References

• Passive model reduction	• Antoulas, SCL (2005) • Sorensen, SCL (2005)
• Optimal $\mathcal{H}_2$ model reduction	• Gugercin, Antoulas, Beattie, Report (2005)
• Decay rates of HSV • MIMO Hankel norm approximation	• Antoulas, Sorensen, Zhou, SCL (2003) • Auba, Funahashi, IEEE CAS (1996)

THANKS PAUL!