Chair of Automatic Control

Technische Universität München

Interpolation-based $\mathcal{H}_2$ Pseudo-Optimal Model Reduction of Bilinear Systems

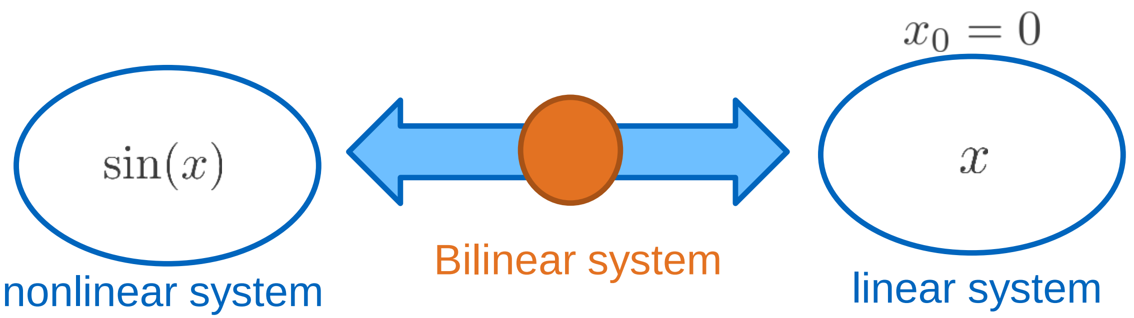
M. Cruz Varona, E. Olcay, B. Lohmann

GMA-Fachausschuss 1.30

Anif, 19.09.2016

- **Bilinear systems** are a special class of nonlinear systems (weakly nonlinear)

- Interface between fully nonlinear and linear systems



- The analogy between linear and bilinear systems allows us to **transfer** some of the **existing linear reduction techniques to the bilinear case**

Nonlinear state equation

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

$$\det(\mathbf{E}) \neq 0$$

Carleman bilinearization [Rugh '81]

Bilinear model

$$\boldsymbol{\Sigma}: \quad \mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \sum_{j=1}^{m} \mathbf{N}_j \mathbf{x}(t) u_j(t) + \mathbf{B}\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t), \;\; \mathbf{x}(0) = \mathbf{x}_0$$

$\mathbf{E}, \mathbf{A}, \mathbf{N}_j \in \mathbb{R}^{n \times n}$

$\mathbf{B} \in \mathbb{R}^{n \times m}; \;\; \mathbf{C} \in \mathbb{R}^{p \times n}$

Outline

I. Why bilinear systems?

- Motivation

II. Model Reduction for Bilinear Systems

- Projective MOR of bilinear systems
- Bilinear systems theory
- Interpolation-based model reduction via Krylov subspaces
- $\mathcal{H}_2$ optimal model reduction of bilinear systems
- $\mathcal{H}_2$ pseudo-optimal reduction

III. Numerical Examples

IV. Summary and Outlook

Projective Reduction of Bilinear Systems

Bilinear model

$$\boldsymbol{\Sigma}: \quad \begin{aligned} \mathbf{E}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \sum_{j=1}^{m} \mathbf{N}_j \mathbf{x}(t) u_j(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t), \;\; \mathbf{x}(0) = \mathbf{x}_0 \end{aligned}$$

$\mathbf{E}, \mathbf{A}, \mathbf{N}_j \in \mathbb{R}^{n \times n}$

$\mathbf{B} \in \mathbb{R}^{n \times m}; \;\; \mathbf{C} \in \mathbb{R}^{p \times n}$

$n \gg r$

Projection

$$\mathbf{V}, \mathbf{W} \in \mathbb{R}^{n \times r} \quad \Rightarrow \quad \mathbf{E}_r = \mathbf{W}^T \mathbf{E} \mathbf{V}, \quad \mathbf{A}_r = \mathbf{W}^T \mathbf{A} \mathbf{V}, \quad \mathbf{N}_{r,j} = \mathbf{W}^T \mathbf{N}_j \mathbf{V},$$

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

Reduced bilinear model

$$\boldsymbol{\Sigma}_r: \quad \begin{aligned} \mathbf{E}_r \dot{\mathbf{x}}_r(t) &= \mathbf{A}_r \mathbf{x}_r(t) + \sum_{j=1}^{m} \mathbf{N}_{r,j} \mathbf{x}_r(t) u_j(t) + \mathbf{B}_r \mathbf{u}(t) \\ \mathbf{y}_r(t) &= \mathbf{C}_r \mathbf{x}_r(t) \end{aligned}$$

$\mathbf{E}_r, \mathbf{A}_r, \mathbf{N}_{r,j} \in \mathbb{R}^{r \times r}$

$\mathbf{B}_r \in \mathbb{R}^{r \times m}; \;\; \mathbf{C}_r \in \mathbb{R}^{p \times r}$

I. Why bilinear systems?

- Motivation

II. Model Reduction for Bilinear Systems

- Projective MOR of bilinear systems
- Bilinear systems theory
- Interpolation-based model reduction via Krylov subspaces
- $\mathcal{H}_2$ optimal model reduction of bilinear systems
- $\mathcal{H}_2$ pseudo-optimal reduction

III. Numerical Examples

IV. Summary and Outlook

[Rugh '81]

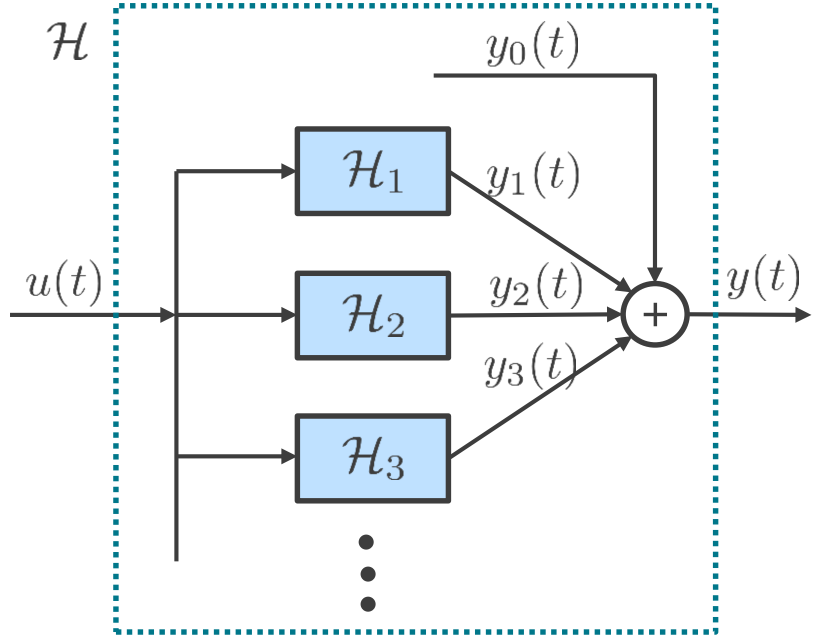
Output Response and Transfer Functions of Bilinear Systems

$$y(t) = \sum_{k=0}^{\infty} y_k(t) = \sum_{k=0}^{\infty} \mathcal{H}_k\left[u(t)\right]$$

$y_k(t)$: output of *k*-th homogenous subsystem

$\mathcal{H}_k$: *k*-th order Volterra operator

$y_0 = \mathcal{H}_0$: constant output



[Rugh '81]

- Within this framework, the input-output representation is given by

$$y(t) = \sum_{k=1}^{\infty} y_k(t_1, \ldots, t_k)$$
$$= \sum_{k=1}^{\infty} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{k-1}} g_k(t_1, \ldots, t_k)\, u(t-t_1) \ldots u(t-t_k) \mathrm{d}t_k \ldots \mathrm{d}t_1$$

- Definition by convolution integrals

Input-output representation

$$y(t) = \sum_{k=1}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \underbrace{\mathbf{c}^T e^{\mathbf{E}^{-1}\mathbf{A}\tau_k} \mathbf{E}^{-1}\mathbf{N} \cdots \mathbf{E}^{-1}\mathbf{N} e^{\mathbf{E}^{-1}\mathbf{A}\tau_2} \mathbf{E}^{-1}\mathbf{N} e^{\mathbf{E}^{-1}\mathbf{A}\tau_1} \mathbf{E}^{-1}\mathbf{b}}_{g_k(\tau_1, \ldots, \tau_k)}$$
$$\times u(t-\tau_k) \cdots u(t-\tau_k - \ldots - \tau_1)\, \mathrm{d}\tau_k \cdots \mathrm{d}\tau_1$$

***k*-th order transfer function of a bilinear system**

$$G_k(s_1, \ldots, s_k) = \mathbf{c}^T (s_k\mathbf{E} - \mathbf{A})^{-1}\mathbf{N} \cdots \mathbf{N}(s_2\mathbf{E} - \mathbf{A})^{-1}\mathbf{N}(s_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{b}$$

[Rugh '81], [Flagg '12]

- First three subsystems:

$$k = 1: \quad G_1(s_1) = \mathbf{c}^T(s_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{b}$$

$$k = 2: \quad G_2(s_1, s_2) = \mathbf{c}^T(s_2\mathbf{E} - \mathbf{A})^{-1}\mathbf{N}(s_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{b}$$

$$k = 3: \quad G_3(s_1, s_2, s_3) = \mathbf{c}^T(s_3\mathbf{E} - \mathbf{A})^{-1}\mathbf{N}(s_2\mathbf{E} - \mathbf{A})^{-1}\mathbf{N}(s_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{b}$$

$\mathcal{H}_2$ norm for bilinear systems

$$||\mathbf{\Sigma}||^2_{\mathcal{H}_2} := \sum_{k=1}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^k} G_k(j\omega_1, \ldots, j\omega_k) G_k^*(j\omega_1, \ldots, j\omega_k) \mathrm{d}\omega_1 \cdots \mathrm{d}\omega_k$$

- **P** and **Q** satisfy the following bilinear Lyapunov equations:

Bilinear Lyapunov equations

$$\mathbf{APE}^T + \mathbf{EPA}^T + \mathbf{NPN}^T + \mathbf{bb}^T = \mathbf{0},$$

$$\mathbf{A}^T\mathbf{QE} + \mathbf{E}^T\mathbf{QA} + \mathbf{N}^T\mathbf{QN} + \mathbf{cc}^T = \mathbf{0}$$

Outline

I. Why bilinear systems?

- Motivation

II. Model Reduction for Bilinear Systems

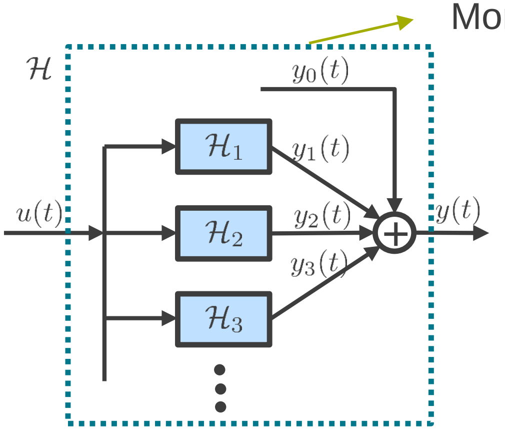
- Projective MOR of bilinear systems
- Bilinear systems theory
- Interpolation-based model reduction via Krylov subspaces
- $\mathcal{H}_2$ optimal model reduction of bilinear systems
- $\mathcal{H}_2$ pseudo-optimal reduction

III. Numerical Examples

IV. Summary and Outlook

Volterra series-based interpolation:

Enforcing multipoint interpolation of the underlying Volterra series



Volterra series interpolation [Flagg/Gugercin '15]

Set of interpolation points: $S = \{s_1, \ldots, s_r\}$

$$\sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \cdots \sum_{l_{k-1}}^{r} \eta_{l_1,\ldots,l_{k-1},j} G_k(s_{l_1}, \ldots, s_{l_{k-1}}, s_j) = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \cdots \sum_{l_{k-1}}^{r} \eta_{l_1,\ldots,l_{k-1},j} G_{k,r}(s_{l_1}, \ldots, s_{l_{k-1}}, s_j)$$

This approach interpolates the weighted series at the interpolation points $s_1, \ldots, s_r$

Weighting matrices: $\mathbf{U}_V = \{u_{i,j}\}, \mathbf{U}_W = \{\hat{u}_{i,j}\} \in \mathbb{R}^{r \times r}$

$\eta_{l_1,\ldots,l_{k-1},j} = u_{j,l_{k-1}} u_{l_{k-1},l_{k-2}} \ldots u_{l_2,l_1}$ for $k \geq 2$ and $\eta_{l_1} = 1$ for $l_1 = 1, \ldots, r$

Weights and shifts are defined by the user

Example: $\eta_{1,2,3} = u_{3,2} \cdot u_{2,1}$

Explicit computation of Volterra series-based interpolation [Flagg/Gugercin '15]

$$\mathbf{v}_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \cdots \sum_{l_{k-1}}^{r} \eta_{l_1,\ldots,l_{k-1},j} (\mathbf{A} - s_j \mathbf{E})^{-1} \mathbf{N} (\mathbf{A} - s_{l_{k-1}} \mathbf{E})^{-1} \mathbf{N} \ldots \mathbf{N} (\mathbf{A} - s_1 \mathbf{E})^{-1} \mathbf{b}$$

$$\mathbf{w}_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \cdots \sum_{l_{k-1}}^{r} \hat{\eta}_{l_1,\ldots,l_{k-1},j} (\mathbf{A}^T - \mu_j \mathbf{E}^T)^{-1} \mathbf{N}^T (\mathbf{A}^T - \mu_{l_{k-1}} \mathbf{E}^T)^{-1} \mathbf{N}^T \ldots \mathbf{N}^T (\mathbf{A}^T - \mu_1 \mathbf{E}^T)^{-1} \mathbf{c}$$

Link Krylov-Sylvester

$$\tilde{\mathbf{V}} = [\mathbf{v}_1, \ldots, \mathbf{v}_r] \in \mathbb{R}^{n \times r}$$

$$\tilde{\mathbf{W}} = [\mathbf{w}_1, \ldots, \mathbf{w}_r] \in \mathbb{R}^{n \times r}$$

Volterra Sylvester equations

$$\mathbf{A}\tilde{\mathbf{V}} - \mathbf{E}\tilde{\mathbf{V}}\mathbf{S}_V - \mathbf{N}\tilde{\mathbf{V}}\mathbf{U}_V^T = \mathbf{b}\mathbf{e}^T$$

$$\mathbf{A}^T\tilde{\mathbf{W}} - \mathbf{E}^T\tilde{\mathbf{W}}\mathbf{S}_W^T - \mathbf{N}^T\tilde{\mathbf{W}}\mathbf{U}_W^T = \mathbf{c}\mathbf{e}^T$$

Goal of $\mathcal{H}_2$ optimality

$$\|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_r\|_{\mathcal{H}_2} = \min_{\dim(\mathbf{H}_r)=r} \|\boldsymbol{\Sigma} - \mathbf{H}_r\|_{\mathcal{H}_2}$$

- minimizing the approximation error $\|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_r\|_{\mathcal{H}_2}$

Error system

$$\boldsymbol{\Sigma}_{err} := \boldsymbol{\Sigma} - \boldsymbol{\Sigma}_r$$

⇒

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

$$E^2 := \|\boldsymbol{\Sigma}^2_{err}\|^2_{\mathcal{H}_2} := \|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_r\|^2_{\mathcal{H}_2}$$

Necessary conditions for $\mathcal{H}_2$ optimality [Benner/Breiten '12]

First order necessary conditions:

I) $\dfrac{\partial E^2}{\partial \tilde{\mathbf{C}}_{ij}} = 0 \iff \mathbf{G}_r(-\overline{\lambda}_i)\tilde{\mathbf{B}}_i = \mathbf{G}(-\overline{\lambda}_i)\tilde{\mathbf{B}}_i$

II) $\dfrac{\partial E^2}{\partial \tilde{\mathbf{B}}_{ij}} = 0 \iff \tilde{\mathbf{C}}_i^T\mathbf{G}(-\overline{\lambda}_i) = \tilde{\mathbf{C}}_i^T\mathbf{G}_r(-\overline{\lambda}_i)$

III) $\dfrac{\partial E^2}{\partial \lambda_i} = 0 \iff \tilde{\mathbf{C}}_i^T\mathbf{G}'(-\overline{\lambda}_i)\tilde{\mathbf{B}}_i = \tilde{\mathbf{C}}_i^T\mathbf{G}_r'(-\overline{\lambda}_i)\tilde{\mathbf{B}}_i$

IV) $\dfrac{\partial E^2}{\partial \tilde{\mathbf{N}}_{ij}} = 0$

$\mathcal{H}_2$ optimality for bilinear systems

$$\|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_r\|_{\mathcal{H}_2} = \min_{\dim(\mathbf{H}_r)=r} \|\boldsymbol{\Sigma} - \mathbf{H}_r\|_{\mathcal{H}_2}$$

$\boldsymbol{\Sigma}_r$ satisfies

$$\frac{\partial E^2}{\partial \tilde{\mathbf{C}}_{ij}} = 0 \qquad \frac{\partial E^2}{\partial \tilde{\mathbf{N}}_{ij}} = 0 \qquad \frac{\partial E^2}{\partial \tilde{\mathbf{B}}_{ij}} = 0 \qquad \frac{\partial E^2}{\partial \lambda_i} = 0$$

$$\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}}^{r} \eta_{l_1,\ldots,l_{k-1},j} G_k(s_{l_1},\ldots,s_{l_{k-1}},s_j)$$
$$= \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}}^{r} \eta_{l_1,\ldots,l_{k-1},j} G_{k,r}(s_{l_1},\ldots,s_{l_{k-1}},s_j)$$

$\Leftrightarrow$

$\phi_{l_1,\cdots,l_k}$: reduced order residues

λ_{l_i} : reduced order poles

[Flagg/Gugercin '15]

$$\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r} \phi_{l_1,\cdots,l_k} G_k(-\overline{\lambda}_{l_1},\cdots,-\overline{\lambda}_k) = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r} \phi_{l_1,\cdots,l_k} G_{k,r}(-\overline{\lambda}_{l_1},\cdots,-\overline{\lambda}_k),$$

$$\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r} \phi_{l_1,\cdots,l_k} \left(\sum_{j=1}^{k} \frac{\partial}{\partial s_j} G_k(-\overline{\lambda}_{l_1},\cdots,-\overline{\lambda}_k)\right) = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r} \phi_{l_1,\cdots,l_k} \left(\sum_{j=1}^{k} \frac{\partial}{\partial s_j} G_{k,r}(-\overline{\lambda}_{l_1},\cdots,-\overline{\lambda}_k)\right)$$

$\mathcal{H}_2$ optimality vs. $\mathcal{H}_2$ pseudo-optimality of bilinear systems

$\mathcal{H}_2$ optimality

$$\|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_r\|_{\mathcal{H}_2} = \min_{\dim(\mathbf{H}_r)=r} \|\boldsymbol{\Sigma} - \mathbf{H}_r\|_{\mathcal{H}_2}$$

$\boldsymbol{\Sigma}_r$ satisfies

$$\frac{\partial E^2}{\partial \tilde{\mathbf{C}}_{ij}} = 0 \quad \frac{\partial E^2}{\partial \tilde{\mathbf{N}}_{ij}} = 0 \quad \frac{\partial E^2}{\partial \tilde{\mathbf{B}}_{ij}} = 0 \quad \frac{\partial E^2}{\partial \lambda_i} = 0 \quad \Leftrightarrow \quad \begin{aligned} \boldsymbol{\Sigma}(-\overline{\lambda}_i) &= \boldsymbol{\Sigma}_r(-\overline{\lambda}_i) \\ \boldsymbol{\Sigma}'(-\overline{\lambda}_i) &= \boldsymbol{\Sigma}_r'(-\overline{\lambda}_i) \end{aligned}$$

$\mathcal{H}_2$ pseudo-optimality

$\mathcal{L} = \{\lambda_1, \cdots, \lambda_n\}$: fixed reduced poles

$\mathcal{G}(\mathcal{L})$: Subset of reduced models

$\boldsymbol{\Sigma}_r$ satisfies

$$||\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_r||_{\mathcal{H}_2} = \min_{\mathbf{H}_r \in \mathcal{G}(\mathcal{L})} ||\boldsymbol{\Sigma} - \mathbf{H}_r||_{\mathcal{H}_2}$$

if

$$\boldsymbol{\Sigma}(-\overline{\lambda}_i) = \boldsymbol{\Sigma}_r(-\overline{\lambda}_i)$$

$$\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r} \phi_{l_1,\cdots,l_k} G_k(-\overline{\lambda}_{l_1},\cdots,-\overline{\lambda}_k) = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r} \phi_{l_1,\cdots,l_k} G_{k,r}(-\overline{\lambda}_{l_1},\cdots,-\overline{\lambda}_k),$$

$$\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r} \phi_{l_1,\cdots,l_k} \left(\sum_{j=1}^{k} \frac{\partial}{\partial s_j} G_k(-\overline{\lambda}_{l_1},\cdots,-\overline{\lambda}_k)\right) = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r} \phi_{l_1,\cdots,l_k} \left(\sum_{j=1}^{k} \frac{\partial}{\partial s_j} G_{k,r}(-\overline{\lambda}_{l_1},\cdots,-\overline{\lambda}_k)\right)$$

Notation

$$\mathbf{B}_\perp = \mathbf{B} - \mathbf{E}\mathbf{V}\mathbf{E}_r^{-1}\mathbf{B}_r$$

$$\mathbf{A}\mathbf{V} - \mathbf{E}\mathbf{V}\mathbf{S}_V - \mathbf{N}\mathbf{V}\mathbf{U}_V^T = \mathbf{B}\mathbf{R}_V$$

$$\mathbf{A}_r\mathbf{P}_r\mathbf{E}_r^T + \mathbf{E}_r\mathbf{P}_r\mathbf{A}_r^T + \mathbf{N}_{r,j}\mathbf{P}_r\mathbf{N}_{r,j}^T + \mathbf{B}_r\mathbf{B}_r^T = \mathbf{0}$$

$$\mathbf{A}\mathbf{X}\mathbf{E}_r^T + \mathbf{E}\mathbf{X}\mathbf{A}_r^T + \mathbf{N}\mathbf{X}\mathbf{N}_r^T + \mathbf{B}\mathbf{B}_r = \mathbf{0}$$

$\lambda_i(\mathbf{S}_V)$: Shifts
$\mathbf{U}_V$: Weights
$\mathbf{R}_V$: Right tangential directions

New conditions for pseudo-optimality for bilinear systems

i) $\mathbf{S}_V = -\mathbf{P}_r\mathbf{A}_r^T\mathbf{E}_r^{-T}\mathbf{P}_r^{-1}$

ii-1) $\mathbf{E}_r^{-1}\mathbf{B}_r + \mathbf{P}_r\mathbf{R}_V^T = \mathbf{0}$

ii-2) $\mathbf{P}_r^{-1}\mathbf{U}_V^T + \mathbf{N}_r^T\mathbf{E}_r^{-T}\mathbf{P}_r^{-1} = \mathbf{0}$

iii) $\mathbf{S}_V\mathbf{P}_r + \mathbf{P}_r\mathbf{S}_V^T - \mathbf{P}_r\mathbf{R}_V^T\mathbf{R}_V\mathbf{P}_r + \mathbf{P}_r\mathbf{U}_V\mathbf{N}_r^T\mathbf{E}_r^{-T} = \mathbf{0}$
$\Leftrightarrow \mathbf{P}_r^{-1}\mathbf{S}_V + \mathbf{S}_V^T\mathbf{P}_r^{-1} - \mathbf{U}_V\mathbf{P}_r^{-1}\mathbf{U}_V^T - \mathbf{R}_V^T\mathbf{R}_V = \mathbf{0}$

iv) $\mathbf{X} = \mathbf{V}\mathbf{P}_r$

v) $\mathbf{A}\hat{\mathbf{P}}\mathbf{E}^T + \mathbf{E}\hat{\mathbf{P}}\mathbf{A}^T + \mathbf{N}\hat{\mathbf{P}}\mathbf{N}^T + \mathbf{B}\mathbf{B}^T = \mathbf{B}_\perp\mathbf{B}_\perp^T$

vi) $\mathbf{P}_r^{-1} = \mathbf{E}_r^T\mathbf{Q}_f\mathbf{E}_r$

$\mathcal{H}_2$ pseudo-optimal reduction of bilinear systems

BIPORK: Bilinear pseudo-optimal rational Krylov

Algorithm 1 Bilinear pseudo-optimal rational Krylov (BIPORK)

Input: $\mathbf{V}$, $\mathbf{S}_V$, $\mathbf{U}_V$, $\mathbf{R}_V$, $\mathbf{C}$, such that $\mathbf{AV}-\mathbf{EVS}_V-\mathbf{NVU}_V^T=\mathbf{BR}_V$ is satisfied

Output: $\mathcal{H}_2$ pseudo-optimal reduced model $\mathbf{\Sigma}_r$

1: $\mathbf{P}_r^{-1}$: solution of condition iii): $\mathbf{P}_r^{-1}\mathbf{S}_V+\mathbf{S}_V^T\mathbf{P}_r^{-1}-\mathbf{U}_V\mathbf{P}_r^{-1}\mathbf{U}_V^T-\mathbf{R}_V^T\mathbf{R}_V=\mathbf{0}$

2: $\mathbf{N}_r=-(\mathbf{P}_r^{-1})^{-1}\mathbf{U}_V\mathbf{P}_r^{-1}$ condition ii-2)

3: $\mathbf{B}_r=-\left(\mathbf{P}_r^{-1}\right)^{-1}\mathbf{R}_V^T$ condition ii-1)

4: $\mathbf{A}_r=\mathbf{S}_V+\mathbf{B}_r\mathbf{R}_V+\mathbf{N}_r\mathbf{U}_V^T$, $\mathbf{E}_r=\mathbf{I}_r$, $\mathbf{C}_r=\mathbf{CV}$

Advantages and properties of BIPORK

- ROM is globally optimal within a subset: $||\mathbf{\Sigma}-\mathbf{\Sigma}_r||_{\mathcal{H}_2}=\min_{\mathbf{H}_r\in\mathcal{G}(\mathcal{L})}||\mathbf{\Sigma}-\mathbf{H}_r||_{\mathcal{H}_2}$
- Eigenvalues of ROM: $\Lambda(\mathbf{S}_V)=\Lambda(-\mathbf{E}_r^{-1}\mathbf{A}_r)$

  → choice of the shifts is twice as important
- Stability preservation in the ROM can be ensured (choice of shifts & weights)
- Low numerical effort required: solution of a bilinear Lyapunov equation and two linear systems of equations, both of reduced order

Outline

I. Why bilinear systems?

- Motivation

II. Model Reduction for Bilinear Systems

- Projective MOR of bilinear systems
- Bilinear systems theory
- Interpolation-based model reduction via Krylov subspaces
- $\mathcal{H}_2$ optimal model reduction of bilinear systems
- $\mathcal{H}_2$ pseudo-optimal reduction

III. Numerical Examples

IV. Summary and Outlook

Heat Transfer Model: Bilinear boundary controlled heat transfer system

Heat equation

$$x_t = \Delta x$$

$$\frac{\partial x}{\partial t} = \frac{\partial^2 x}{\partial z_1^2} + \frac{\partial^2 x}{\partial z_2^2} \quad \text{on unit square } \Omega = [0,1] \times [0,1]$$

Boundary conditions

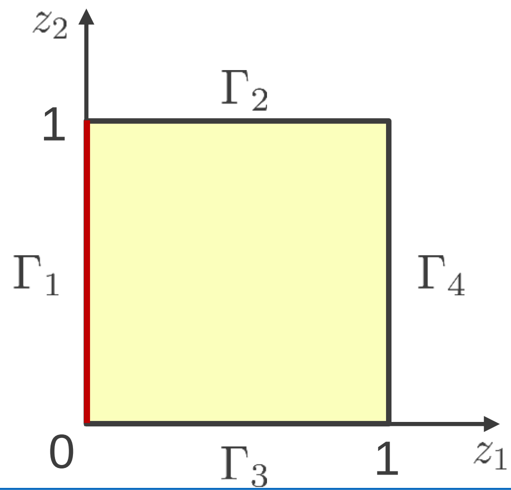
$$n \cdot \left(\frac{\partial x}{\partial z_1} + \frac{\partial x}{\partial z_2}\right) = (x-1)\,u \quad \text{on } \Gamma_1$$

$$x = 0 \quad \text{on } \Gamma_2, \Gamma_3, \Gamma_4$$

- Spatial discretization on an equidistant $k \times k$ grid together with the boundary conditions yields: [Benner/Breiten '12]

  $\Rightarrow \; \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{N}\,\mathbf{x}\,u + \mathbf{b}u$ of dimension $n = k^2$

- Output: $y = \mathbf{c}^T\mathbf{x} = \frac{1}{k^2}\begin{bmatrix}1 & \cdots & 1\end{bmatrix}\mathbf{x}$



$\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$: boundary domains

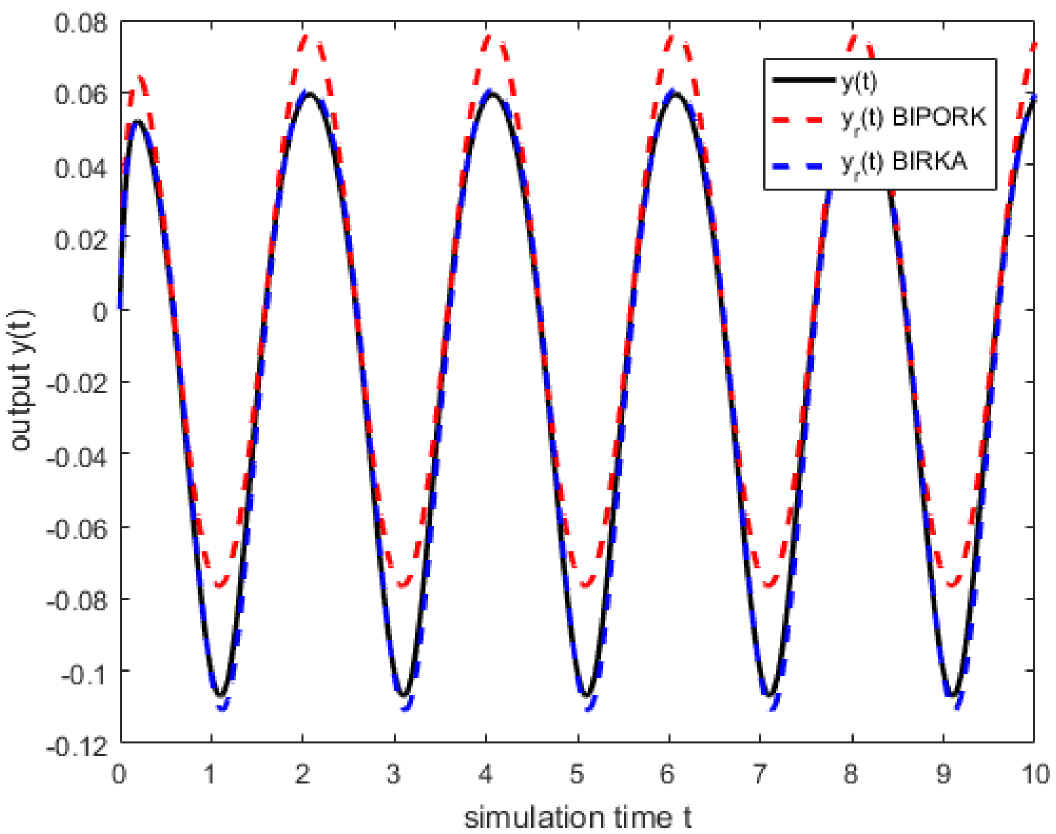
BIRKA: $\mathcal{H}_2$ optimality
No convergence after 50 iterations

BIPORK: $\mathcal{H}_2$ pseudo-optimality

$$s_0 = \begin{bmatrix} 10 & 20 & 30 & 40 & 50 & 60 \end{bmatrix}$$

$$\mathbf{U}_V = \mathrm{diag}([1e^{-10}\ 2e^{-10}\ 7e^{-10}\ 5e^{-10}\ 7e^{-10}\ 8e^{-10}])$$

Output response for $u(t) = \cos(\pi t)$:



$k = 50;\ n = 2500$
$r = 6$

I. Why bilinear systems?

- Motivation

II. Model Reduction for Bilinear Systems

- Projective MOR of bilinear systems
- Bilinear systems theory
- Interpolation-based model reduction via Krylov subspaces
- $\mathcal{H}_2$ optimal model reduction of bilinear systems
- $\mathcal{H}_2$ pseudo-optimal reduction

III. Numerical Examples

IV. Summary and Outlook

Summary and Outlook

Summary:

- **Goal:** Reduction of high dimensional nonlinear systems
- Approximation of a nonlinear system by a bilinear system using Volterra theory
- Systems theory and model reduction for bilinear systems (based on Krylov)
- $\mathcal{H}_2$ pseudo-optimal model reduction for bilinear systems
  - Derivation of new conditions for $\mathcal{H}_2$ pseudo-optimality for bilinear systems
  - Bilinear pseudo-optimal rational Krylov (BIPORK) – **conditions ii-1), ii-2), iii)**

Outlook:

- Solution of bilinear Lyapunov equations with BIPORK and the link with the alternating direction implicit (ADI) method: **conditions iv)-v)**
- Cumulative reduction (CuRe) for bilinear systems: **condition vi)**
- Quadratic-bilinear MOR
  - Stability-preserving two-sided rational Krylov for QBDAEs?
  - MIMO reduction for QBDAEs? Choice of optimal expansion points?

References

[Benner/Breiten '12]	Interpolation-based H2-model reduction of bilinear control systems. SIAM Journal on Matrix Analysis and Applications
[Flagg '12]	Interpolation Methods for the Model Reduction of Bilinear Systems, PhD thesis
[Flagg/Gugercin '15]	Multipoint Volterra series interpolation and H2 optimal model reduction of bilinear systems, SIAM Journal on …
[Rugh '81]	Nonlinear system theory. The Volterra/Wiener Approach
[Wolf '14]	H2 Pseudo-Optimal Model Order Reduction, PhD thesis

Thank you for your attention!

Backup slides

- **Duality:** Krylov subspaces with Sylvester equations

$$\text{span}\{\mathbf{V}\} = \mathcal{K}_r\left((\mathbf{A}-s_0\mathbf{E})^{-1}\mathbf{E}, (\mathbf{A}-s_0\mathbf{E})^{-1}\mathbf{B}\right)$$
$$\text{span}\{\mathbf{W}\} = \mathcal{K}_r\left((\mathbf{A}-s_0\mathbf{E})^{-T}\mathbf{E}^T, (\mathbf{A}-s_0\mathbf{E})^{-T}\mathbf{C}^T\right)$$

$\Longleftrightarrow$

$$\mathbf{AV} - \mathbf{EVS}_V = \mathbf{BR}_V$$
$$\mathbf{A}^T\mathbf{W} - \mathbf{E}^T\mathbf{WS}_W^T = \mathbf{C}^T\mathbf{L}_W$$

$\lambda_i(\mathbf{S}_V) = s_0$: shifts

$\mathbf{R}_V, \mathbf{L}_W$: tangential directions

- $\mathcal{H}_2$ optimality vs. $\mathcal{H}_2$ pseudo-optimality

$\mathcal{H}_2$ optimality

- Problem:

$$\|\mathbf{G}-\mathbf{G}_r\|_{\mathcal{H}_2} = \min_{\dim(\widetilde{\mathbf{G}}_r)=r} \left\|\mathbf{G}-\widetilde{\mathbf{G}}_r\right\|_{\mathcal{H}_2}$$

- Necessary conditions for local $\mathcal{H}_2$ optimality (SISO): (Meier-Luenberger)

$$G(-\overline{\lambda}_{r,i}) = G_r(-\overline{\lambda}_{r,i})$$
$$G'(-\overline{\lambda}_{r,i}) = G_r'(-\overline{\lambda}_{r,i})$$

- $\mathbf{G}_r$ minimizes the $\mathcal{H}_2$ error locally within the set of all ROMs of order r

$\mathcal{H}_2$ pseudo-optimality

- Problem: $\Lambda = \{\lambda_1, \ldots, \lambda_r\},\ \lambda_i \in \mathbb{C}^-$

$$\|\mathbf{G}-\mathbf{G}_r\|_{\mathcal{H}_2} = \min_{\widetilde{\mathbf{G}}_r \in \mathcal{G}(\Lambda)} \left\|\mathbf{G}-\widetilde{\mathbf{G}}_r\right\|_{\mathcal{H}_2}$$

- Necessary **and** sufficient condition for global $\mathcal{H}_2$ pseudo-optimality:

$$G(-\overline{\lambda}_{r,i}) = G_r(-\overline{\lambda}_{r,i})$$

- Pseudo-optimal means optimal in a certain subset
- $\mathbf{G}_r$ minimizes the $\mathcal{H}_2$ error globally within the subset of all ROMs of order r with poles Λ

Notation

Gramian	$\mathbf{A}_r\mathbf{P}_r\mathbf{E}_r^T + \mathbf{E}_r\mathbf{P}_r\mathbf{A}_r^T + \mathbf{B}_r\mathbf{B}_r^T = \mathbf{0}$	(known)
Scalar product	$\mathbf{A}\mathbf{X}\mathbf{E}_r^T + \mathbf{E}\mathbf{X}\mathbf{A}_r^T + \mathbf{B}\mathbf{B}_r^T = \mathbf{0}$	(unknown)
Krylov	$\mathbf{A}\mathbf{V} - \mathbf{E}\mathbf{V}\mathbf{S}_V = \mathbf{B}\mathbf{R}_V$	(known)
Projection	$\mathbf{B}_\perp = \mathbf{B} - \mathbf{E}\mathbf{V}\mathbf{E}_r^{-1}\mathbf{B}_r$	(known)

New conditions for pseudo-optimality for linear systems [Wolf '14]

Let $\mathbf{V}$ be a basis of a Krylov subspace. Let $\mathbf{G}_r(s)$ be the reduced model obtained by projection with $\mathbf{W}$. Then, the following conditions are equivalent:

i) $\mathbf{S}_V = -\mathbf{P}_r\mathbf{A}_r^T\mathbf{E}_r^{-T}\mathbf{P}_r^{-1}$

ii) $\mathbf{E}_r^{-1}\mathbf{B}_r + \mathbf{P}_r\mathbf{R}_V^T = \mathbf{0}$

iii) $\mathbf{S}_V\mathbf{P}_r + \mathbf{P}_r\mathbf{S}_V^T - \mathbf{P}_r\mathbf{R}_V^T\mathbf{R}_V\mathbf{P}_r = \mathbf{0} \Leftrightarrow \mathbf{P}_r^{-1}\mathbf{S}_V + \mathbf{S}_V^T\mathbf{P}_r^{-1} - \mathbf{R}_V^T\mathbf{R}_V = \mathbf{0}$

iv) $\mathbf{X} = \mathbf{V}\mathbf{P}_r$

v) $\mathbf{A}\widehat{\mathbf{P}}\mathbf{E}^T + \mathbf{E}\widehat{\mathbf{P}}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T = \mathbf{B}_\perp\mathbf{B}_\perp^T$

vi) $\mathbf{P}_r^{-1} = \mathbf{E}_r^T\widehat{\mathbf{Q}}_r\mathbf{E}_r$

PORK: Pseudo-optimal rational Krylov [Wolf '14]

Algorithm 1 Pseudo-optimal rational Krylov (PORK)

Input: $\mathbf{V}$, $\mathbf{S}_V$, $\mathbf{R}_V$, $\mathbf{C}$, such that $\mathbf{AV} - \mathbf{EVS}_V = \mathbf{BR}_V$ is satisfied

Output: $\mathcal{H}_2$ pseudo-optimal reduced model $\mathbf{G}_r(s) = \mathbf{C}_r\left(s\mathbf{E}_r - \mathbf{A}_r\right)^{-1}\mathbf{B}_r$

1: $\mathbf{P}_r^{-1} = \text{lyap}\left(-\mathbf{S}_V^T, \mathbf{R}_V^T\mathbf{R}_V\right)$ condition iii): $\mathbf{P}_r^{-1}\mathbf{S}_V + \mathbf{S}_V^T\mathbf{P}_r^{-1} - \mathbf{R}_V^T\mathbf{R}_V = \mathbf{0}$

2: $\mathbf{B}_r = -\left(\mathbf{P}_r^{-1}\right)^{-1}\mathbf{R}_V^T$ condition ii)

3: $\mathbf{A}_r = \mathbf{S}_V + \mathbf{B}_r\mathbf{R}_V$, $\mathbf{E}_r = \mathbf{I}$, $\mathbf{C}_r = \mathbf{CV}$

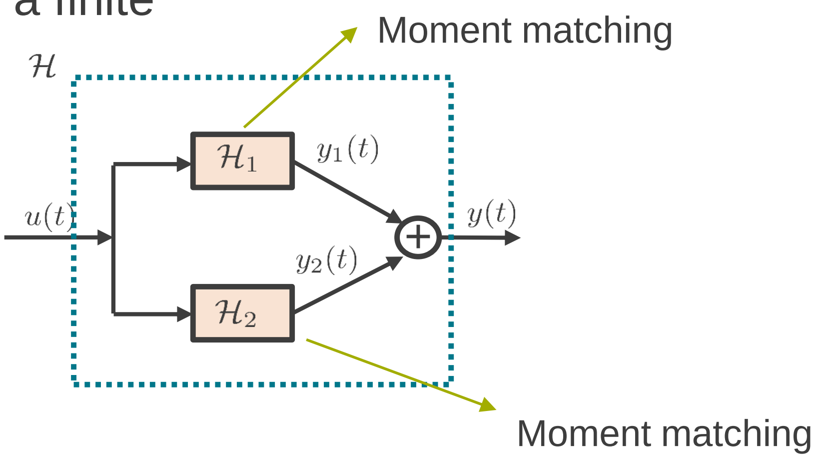
Advantages and properties of PORK

- ROM is globally optimal within a subset: $\|\mathbf{G} - \mathbf{G}_r\|_{\mathcal{H}_2} = \min_{\widetilde{\mathbf{G}}_r \in \mathcal{G}(\Lambda)} \left\|\mathbf{G} - \widetilde{\mathbf{G}}_r\right\|_{\mathcal{H}_2}$
- Eigenvalues of ROM: $\Lambda(\mathbf{S}) = \Lambda(-\mathbf{E}_r^{-1}\mathbf{A}_r)$
  → choice of the shifts is twice as important
- Stability preservation in the ROM can be ensured (choice of shifts)
- Low numerical effort required: solution of a Lyapunov equation and a linear system of equations, both of reduced order.

Subsystem interpolation:

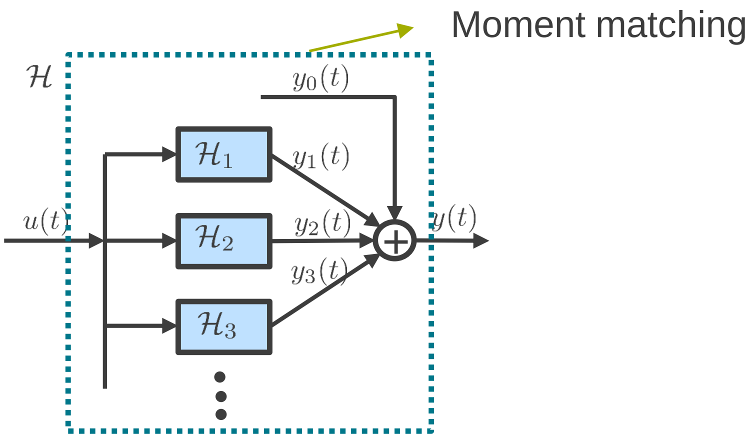
Interpolation is forced on some of the leading subsystem transfer functions.

The interpolation information is placed for a finite number of subsystems in the span of the projection basis.



Volterra series-based interpolation:

Enforcing multipoint interpolation of the underlying Volterra series



Summary and Outlook

Summary:

- ✓ Bilinear systems
- ✓ Mathematical background of several reduction methods
- ✓ Implementation
- ✓ $\mathcal{H}_2$ optimal model reduction for bilinear systems
- ✓ New conditions for $\mathcal{H}_2$ pseudo-optimality for bilinear systems

Conclusions:

- BIRKA: Solving another form of Sylvester equations – convergence problems as IRKA as the reduced dimension increases
- T-BIRKA: better results for high reduced orders - “nearly“ $\mathcal{H}_2$ optimal
- BIRKA initialization strategies
- T-BIRKA implementation: reducing the amount of iterations

Discussion:

- The backslash operator "\" in MATLAB for the matrix inversion may cause memory space exceedance – computation of large scale Sylvester equations in vectorized form
- $\mathcal{H}_2$ norm of the error system – Solution of bilinear Lyapunov equations
- Sensitivity analysis of weighting matrices