Householder Symposium XVIII, Tahoe City, California
June 13, 2011

Rational Krylov Subspaces for Nonlinear Model Reduction

Peter Benner

Max Planck Institute for Dynamics of Complex Technical Systems
Computational Methods in Systems and Control Theory
Magdeburg, Germany

NETWORK THEORY

MAX PLANCK INSTITUTE FOR DYNAMICS OF COMPLEX TECHNICAL SYSTEMS MAGDEBURG

Overview

Introduction

Model Order Reduction

Here, we consider large-scale nonlinear control systems of the form

$$\Sigma : \quad \begin{cases} \dot{x}(t) = f(x(t)) + g(x(t))u(t), \\ y(t) = Cx(t), \quad x(0) = x_0, \end{cases}$$

with $f, g : \mathbb{R}^n \to \mathbb{R}^n$, $C \in \mathbb{R}^{p \times n}$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$.

Introduction

Model Order Reduction

Here, we consider large-scale nonlinear control systems of the form

$$\Sigma : \quad \begin{cases} \dot{x}(t) = f(x(t)) + g(x(t))u(t), \\ y(t) = Cx(t), \quad x(0) = x_0, \end{cases}$$

with $f, g : \mathbb{R}^n \to \mathbb{R}^n$, $C \in \mathbb{R}^{p \times n}$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$.

$$\hat{\Sigma} : \quad \begin{cases} \dot{\hat{x}}(t) = \hat{f}(\hat{x}(t)) + \hat{g}(\hat{x}(t))u(t), \\ \hat{y}(t) = \hat{C}\hat{x}(t), \quad \hat{x}(0) = \hat{x}_0, \end{cases}$$

with $\hat{f}, \hat{g} : \mathbb{R}^{\hat{n}} \to \mathbb{R}^{\hat{n}}$, $C \in \mathbb{R}^{p \times \hat{n}}$, $x \in \mathbb{R}^{\hat{n}}$, $u \in \mathbb{R}^m$, $\hat{y} \in \mathbb{R}^p$, $\hat{n} \ll n$.

Goal

$\hat{y} \approx y$ for all admissible u.

Introduction

Linear System Norms

Let us start with linear systems, i.e. $f(x) = Ax$ and $g(x) = B$.

Two common system norms for measuring approximation quality:

- $\mathcal{H}_2$-norm, $||\Sigma||_{\mathcal{H}_2} = \left(\frac{1}{2\pi} \int_0^{2\pi} \operatorname{tr}\left(H^*(-i\omega) H(i\omega) \right) d\omega \right)^{\frac{1}{2}}$,

Introduction

Linear System Norms

Let us start with linear systems, i.e. $f(x) = Ax$ and $g(x) = B$.

Two common system norms for measuring approximation quality:

- $\mathcal{H}_2$-norm, $||\Sigma||_{\mathcal{H}_2} = \left(\frac{1}{2\pi} \int_0^{2\pi} \operatorname{tr}\left(H^*(-i\omega) H(i\omega) \right) d\omega \right)^{\frac{1}{2}}$,
- $\mathcal{H}_\infty$-norm, $||\Sigma||_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} \sigma_{max}\left(H(i\omega) \right)$,

where

$$H(s) = C\left(sI - A\right)^{-1} B$$

denotes the corresponding transfer function of the linear system.

Introduction

Linear System Norms

Let us start with linear systems, i.e. $f(x) = Ax$ and $g(x) = B$.

Two common system norms for measuring approximation quality:

- $\mathcal{H}_2$-norm, $||\Sigma||_{\mathcal{H}_2} = \left(\frac{1}{2\pi} \int_0^{2\pi} \operatorname{tr}\left(H^*(-i\omega) H(i\omega) \right) d\omega \right)^{\frac{1}{2}}$,
- $\mathcal{H}_\infty$-norm, $||\Sigma||_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} \sigma_{max}\left(H(i\omega) \right)$,

where

$$H(s) = C\left(sI - A\right)^{-1} B$$

denotes the corresponding transfer function of the linear system.

We focus on the first one $\rightsquigarrow$ interpolation-based model reduction approaches.

Introduction

Error system and $\mathcal{H}_2$-Optimality

In order to find an $\mathcal{H}_2$-optimal reduced system, define the error system:

$$A^{err} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}, \quad B^{err} = \begin{bmatrix} B \\ \hat{B} \end{bmatrix}, \quad C^{err} = \begin{bmatrix} C & -\hat{C} \end{bmatrix}.$$

Introduction

Error system and $\mathcal{H}_2$-Optimality

In order to find an $\mathcal{H}_2$-optimal reduced system, define the error system:

$$A^{err} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}, \quad B^{err} = \begin{bmatrix} B \\ \hat{B} \end{bmatrix}, \quad C^{err} = \begin{bmatrix} C & -\hat{C} \end{bmatrix}.$$

$\rightsquigarrow$ first-order necessary $\mathcal{H}_2$-optimality conditions (SISO)

$$H(-\lambda_i) = \hat{H}(-\lambda_i),$$
$$H'(-\lambda_i) = \hat{H}'(-\lambda_i),$$

where λ_i are the poles of the reduced system $\hat{\Sigma}$.

Introduction

Error system and $\mathcal{H}_2$-Optimality

In order to find an $\mathcal{H}_2$-optimal reduced system, define the error system:

$$A^{err} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}, \quad B^{err} = \begin{bmatrix} B \\ \hat{B} \end{bmatrix}, \quad C^{err} = \begin{bmatrix} C & -\hat{C} \end{bmatrix}.$$

$\rightsquigarrow$ first-order necessary $\mathcal{H}_2$-optimality conditions (MIMO)

$$\begin{aligned} H(-\lambda_i)\tilde{B}_i &= \hat{H}(-\lambda_i)\tilde{B}_i, && \text{for } i = 1, \dots, \hat{n}, \\ \tilde{C}_i^T H(-\lambda_i) &= \tilde{C}_i^T \hat{H}(-\lambda_i), && \text{for } i = 1, \dots, \hat{n}, \\ \tilde{C}_i^T H'(-\lambda_i)\tilde{B}_i &= \tilde{C}_i^T \hat{H}'(-\lambda_i)\tilde{B}_i && \text{for } i = 1, \dots, \hat{n}, \end{aligned}$$

where $\hat{A} = R\Lambda R^{-T}$ is the spectral decomposition of the reduced system and $\tilde{B} = \hat{B}^T R^{-T}$, $\tilde{C} = \hat{C}R$.

Introduction

Error system and $\mathcal{H}_2$-Optimality

In order to find an $\mathcal{H}_2$-optimal reduced system, define the error system:

$$A^{err} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}, \quad B^{err} = \begin{bmatrix} B \\ \hat{B} \end{bmatrix}, \quad C^{err} = \begin{bmatrix} C & -\hat{C} \end{bmatrix}.$$

$\rightsquigarrow$ first-order necessary $\mathcal{H}_2$-optimality conditions (MIMO)

$$\begin{aligned} H(-\lambda_i)\tilde{B}_i &= \hat{H}(-\lambda_i)\tilde{B}_i, && \text{for } i = 1, \ldots, \hat{n}, \\ \tilde{C}_i^T H(-\lambda_i) &= \tilde{C}_i^T \hat{H}(-\lambda_i), && \text{for } i = 1, \ldots, \hat{n}, \\ \tilde{C}_i^T H'(-\lambda_i)\tilde{B}_i &= \tilde{C}_i^T \hat{H}'(-\lambda_i)\tilde{B}_i && \text{for } i = 1, \ldots, \hat{n}, \end{aligned}$$

$$\begin{aligned} &\operatorname{vec}(I_p)^T \left(e_j e_i^T \otimes C\right) (-\Lambda \otimes I_n - I_{\hat{n}} \otimes A)^{-1} \left(\tilde{B}^T \otimes B\right) \operatorname{vec}(I_m) \\ &= \operatorname{vec}(I_p)^T \left(e_j e_i^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_{\hat{n}} - I_{\hat{n}} \otimes \hat{A}\right)^{-1} \left(\tilde{B}^T \otimes \hat{B}\right) \operatorname{vec}(I_m), \end{aligned}$$

for $i = 1, \ldots, \hat{n}$ and $j = 1, \ldots, p$.

Introduction

Interpolation of the Transfer Function [Grimme '97]

Construct reduced transfer function by Petrov-Galerkin projection $\mathcal{P} = VW^T$, i.e.

$$\hat{H}(s) = CV\left(sI - W^T AV\right)^{-1} W^T B,$$

Introduction

Interpolation of the Transfer Function [Grimme '97]

Construct reduced transfer function by Petrov-Galerkin projection $\mathcal{P} = VW^T$, i.e.

$$\hat{H}(s) = CV\left(sI - W^T AV\right)^{-1} W^T B,$$

where V and W are given as

$$V = \left[(\sigma_1 I - A)^{-1}B, \ldots, (\sigma_r I - A)^{-1}B\right],$$
$$W = \left[(\sigma_1 I - A^T)^{-1}C^T, \ldots, (\sigma_r I - A^T)^{-1}C^T\right].$$

Introduction

Interpolation of the Transfer Function [Grimme '97]

Construct reduced transfer function by Petrov-Galerkin projection $\mathcal{P} = VW^T$, i.e.

$$\hat{H}(s) = CV\left(sI - W^T AV\right)^{-1} W^T B,$$

where V and W are given as

$$V = \left[(\sigma_1 I - A)^{-1}B, \ldots, (\sigma_r I - A)^{-1}B\right],$$
$$W = \left[(\sigma_1 I - A^T)^{-1}C^T, \ldots, (\sigma_r I - A^T)^{-1}C^T\right].$$

Then

$$H(\sigma_i) = \hat{H}(\sigma_i) \quad \text{and} \quad H'(\sigma_i) = \hat{H}'(\sigma_i),$$

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

Introduction

Interpolation of the Transfer Function [Grimme '97]

Construct reduced transfer function by Petrov-Galerkin projection $\mathcal{P} = VW^T$, i.e.

$$\hat{H}(s) = CV\left(sI - W^TAV\right)^{-1} W^TB,$$

where V and W are given as

$$V = \left[(\sigma_1 I - A)^{-1}B, \ldots, (\sigma_r I - A)^{-1}B\right],$$
$$W = \left[(\sigma_1 I - A^T)^{-1}C^T, \ldots, (\sigma_r I - A^T)^{-1}C^T\right].$$

Then

$$H(\sigma_i) = \hat{H}(\sigma_i) \quad \text{and} \quad H'(\sigma_i) = \hat{H}'(\sigma_i),$$

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

$\rightsquigarrow$ iterative algorithms (IRKA/MIRIAm) that yield $\mathcal{H}_2$-optimal models.

[Gugercin et al. '08], [Bunse-Gerstner et al. '07], [Van Dooren et al. '08]

$\mathcal{H}_2$-Model Reduction for Bilinear Systems

Bilinear Control Systems

Let us now focus on the special case $f(x) = A$ and

$$g(x) = B + \left[N_1, \ldots, N_m\right](I_m \otimes x),$$

i.e. bilinear control systems:

$$\Sigma: \quad \begin{cases} \dot{x}(t) = Ax(t) + \sum_{i=1}^{m} N_i x(t) u_i(t) + Bu(t), \\ y(t) = Cx(t), \quad x(0) = x_0, \end{cases}$$

where $A, N_i \in \mathbb{R}^{n \times n},\ B \in \mathbb{R}^{n \times m},\ C \in \mathbb{R}^{p \times n}$.

$\mathcal{H}_2$-Model Reduction for Bilinear Systems

Bilinear Control Systems

Let us now focus on the special case $f(x) = A$ and

$$g(x) = B + \left[N_1, \ldots, N_m\right] (I_m \otimes x)\,,$$

i.e. bilinear control systems:

$$\Sigma : \quad \begin{cases} \dot{x}(t) = Ax(t) + \sum_{i=1}^{m} N_i x(t) u_i(t) + Bu(t), \\ y(t) = Cx(t), \quad x(0) = x_0, \end{cases}$$

where $A, N_i \in \mathbb{R}^{n \times n},\ B \in \mathbb{R}^{n \times m},\ C \in \mathbb{R}^{p \times n}$.

- Approximation of weakly nonlinear systems $\rightsquigarrow$ Carleman linearization.
- A lot of linear concepts can be extended, e.g. transfer functions, Gramians, Lyapunov equations, ...
- An equivalent structure arise for some stochastic control systems.

$\mathcal{H}_2$-Model Reduction for Bilinear Systems

Some Basic Facts

Output Characterization (SISO): Volterra series

$$y(t) = \sum_{k=1}^{\infty} \int_0^t \int_0^{t_1} \dots \int_0^{t_{k-1}} K(t_1, \dots, t_k) u(t - t_1 - \dots - t_k) \cdots u(t - t_k) dt_k \cdots dt_1,$$

with kernels $K(t_1, \dots, t_k) = Ce^{At_k} N_1 \cdots e^{At_2} N_1 e^{At_1} B$.

$\mathcal{H}_2$-Model Reduction for Bilinear Systems

Some Basic Facts

Output Characterization (SISO): Volterra series

$$y(t) = \sum_{k=1}^{\infty} \int_0^t \int_0^{t_1} \dots \int_0^{t_{k-1}} K(t_1,\dots,t_k) u(t-t_1-\dots-t_k)\cdots u(t-t_k) dt_k \cdots dt_1,$$

with kernels $K(t_1,\dots,t_k) = Ce^{At_k} N_1 \cdots e^{At_2} N_1 e^{At_1} B.$

Multivariate Laplace-transform (SISO):

$$H_k(s_1,\dots,s_k) = C(s_k I - A)^{-1} N_1 \cdots (s_2 I - A)^{-1} N_1 (s_1 I - A)^{-1} B.$$

$\mathcal{H}_2$-Model Reduction for Bilinear Systems

Some Basic Facts

Output Characterization (SISO): Volterra series

$$y(t) = \sum_{k=1}^{\infty} \int_0^t \int_0^{t_1} \dots \int_0^{t_{k-1}} K(t_1, \dots, t_k) u(t - t_1 - \dots - t_k) \cdots u(t - t_k) dt_k \cdots dt_1,$$

with kernels $K(t_1, \dots, t_k) = Ce^{At_k} N_1 \cdots e^{At_2} N_1 e^{At_1} B$.

Multivariate Laplace-transform (SISO):

$$H_k(s_1, \dots, s_k) = C(s_k I - A)^{-1} N_1 \cdots (s_2 I - A)^{-1} N_1 (s_1 I - A)^{-1} B.$$

Bilinear $\mathcal{H}_2$-norm (MIMO):

$$||\Sigma||_{\mathcal{H}_2} := \left(\operatorname{tr} \left(\sum_{k=1}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^k} \overline{H_k(i\omega_1, \dots, i\omega_k)} H_k^T(i\omega_1, \dots, i\omega_k) \right) \right)^{\frac{1}{2}}.$$

[Zhang/Lam. '02]

$\mathcal{H}_2$-Model Reduction for Bilinear Systems

$\mathcal{H}_2$-Norm Computation

Lemma [B./Breiten '11]

Let Σ denote a bilinear system. Then, the $\mathcal{H}_2$-norm is given as:

$$||\Sigma||^2_{\mathcal{H}_2} = (\text{vec}(I_p))^T \, (C \otimes C) \left(-A \otimes I - I \otimes A - \sum_{i=1}^{m} N_i \otimes N_i \right)^{-1} (B \otimes B) \, \text{vec}(I_m).$$

Error System

In order to find an $\mathcal{H}_2$-optimal reduced system, define the error system $\Sigma^{err} := \Sigma - \hat{\Sigma}$ as follows:

$$A^{err} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}, \quad N_i^{err} = \begin{bmatrix} N_i & 0 \\ 0 & \hat{N}_i \end{bmatrix}, \quad B^{err} = \begin{bmatrix} B \\ \hat{B} \end{bmatrix}, \quad C^{err} = \begin{bmatrix} C & -\hat{C} \end{bmatrix}.$$

$\mathcal{H}_2$-Model Reduction

$\mathcal{H}_2$-Optimality Conditions

Let us assume $\hat{\Sigma}$ is given by its eigenvalue decomposition:

$$\hat{A} = R\Lambda R^{-1}, \quad \tilde{N}_i = R^{-1}\hat{N}_i R, \quad \tilde{B} = R^{-1}\hat{B}, \quad \tilde{C} = \hat{C}R.$$

$\mathcal{H}_2$-Model Reduction

$\mathcal{H}_2$-Optimality Conditions

Let us assume $\hat{\Sigma}$ is given by its eigenvalue decomposition:

$$\hat{A} = R\Lambda R^{-1}, \quad \tilde{N}_i = R^{-1}\hat{N}_i R, \quad \tilde{B} = R^{-1}\hat{B}, \quad \tilde{C} = \hat{C}R.$$

Using Λ, $\tilde{N}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g.:

$\mathcal{H}_2$-Model Reduction

$\mathcal{H}_2$-Optimality Conditions

Let us assume $\hat{\Sigma}$ is given by its eigenvalue decomposition:

$$\hat{A} = R\Lambda R^{-1}, \quad \tilde{N}_i = R^{-1}\hat{N}_i R, \quad \tilde{B} = R^{-1}\hat{B}, \quad \tilde{C} = \hat{C}R.$$

Using Λ, $\tilde{N}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g.:

$$(\operatorname{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right)\left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes A - \sum_{i=1}^{m} \tilde{N}_i \otimes N_i\right)^{-1}\left(\tilde{B} \otimes B\right)\operatorname{vec}(I_m)$$

$$= (\operatorname{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right)\left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes \hat{A} - \sum_{i=1}^{m} \tilde{N}_i \otimes \hat{N}_i\right)^{-1}\left(\tilde{B} \otimes \hat{B}\right)\operatorname{vec}(I_m).$$

$\mathcal{H}_2$-Model Reduction

$\mathcal{H}_2$-Optimality Conditions

Let us assume $\hat{\Sigma}$ is given by its eigenvalue decomposition:

$$\hat{A} = R\Lambda R^{-1}, \quad \tilde{N}_i = R^{-1}\hat{N}_i R, \quad \tilde{B} = R^{-1}\hat{B}, \quad \tilde{C} = \hat{C}R.$$

Using Λ, $\tilde{N}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g.:

$$(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes A - \sum_{i=1}^{m} \tilde{N}_i \otimes N_i\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m)$$

$$= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes \hat{A} - \sum_{i=1}^{m} \tilde{N}_i \otimes \hat{N}_i\right)^{-1} \left(\tilde{B} \otimes \hat{B}\right) \mathrm{vec}(I_m).$$

Where is the connection to the interpolation of transfer functions?

$\mathcal{H}_2$-Model Reduction

$\mathcal{H}_2$-Optimality Conditions

Let us assume $\hat{\Sigma}$ is given by its eigenvalue decomposition:

$$\hat{A} = R\Lambda R^{-1}, \quad \tilde{N}_i = R^{-1}\hat{N}_i R, \quad \tilde{B} = R^{-1}\hat{B}, \quad \tilde{C} = \hat{C}R.$$

Using Λ, $\tilde{N}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g.:

$$(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes A - \sum_{i=1}^m \tilde{N}_i \otimes N_i\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m)$$
$$= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes \hat{A} - \sum_{i=1}^m \tilde{N}_i \otimes \hat{N}_i\right)^{-1} \left(\tilde{B} \otimes \hat{B}\right) \mathrm{vec}(I_m).$$

$$(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes A\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m)$$
$$= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes \hat{A}\right)^{-1} \left(\tilde{B} \otimes \hat{B}\right) \mathrm{vec}(I_m).$$

$\mathcal{H}_2$-Model Reduction

$\mathcal{H}_2$-Optimality Conditions

Let us assume $\hat{\Sigma}$ is given by its eigenvalue decomposition:

$$\hat{A} = R\Lambda R^{-1}, \quad \tilde{N}_i = R^{-1}\hat{N}_i R, \quad \tilde{B} = R^{-1}\hat{B}, \quad \tilde{C} = \hat{C}R.$$

Using Λ, $\tilde{N}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g.:

$$(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes A - \sum_{i=1}^{m} \tilde{N}_i \otimes N_i\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m)$$
$$= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes \hat{A} - \sum_{i=1}^{m} \tilde{N}_i \otimes \hat{N}_i\right)^{-1} \left(\tilde{B} \otimes \hat{B}\right) \mathrm{vec}(I_m).$$

$$(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes A\right)^{-1} \mathrm{vec}(B\tilde{B}^T)$$
$$= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes \hat{A}\right)^{-1} \mathrm{vec}(\hat{B}\tilde{B}^T).$$

$\mathcal{H}_2$-Model Reduction

$\mathcal{H}_2$-Optimality Conditions

Let us assume $\hat{\Sigma}$ is given by its eigenvalue decomposition:

$$\hat{A} = R\Lambda R^{-1}, \quad \tilde{N}_i = R^{-1}\hat{N}_i R, \quad \tilde{B} = R^{-1}\hat{B}, \quad \tilde{C} = \hat{C}R.$$

Using Λ, $\tilde{N}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g.:

$$(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes A - \sum_{i=1}^m \tilde{N}_i \otimes N_i\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m)$$

$$= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes \hat{A} - \sum_{i=1}^m \tilde{N}_i \otimes \hat{N}_i\right)^{-1} \left(\tilde{B} \otimes \hat{B}\right) \mathrm{vec}(I_m).$$

$$(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \begin{pmatrix} -\lambda_1 I - A & & \\ & \ddots & \\ & & -\lambda_{\hat{n}} I - A \end{pmatrix}^{-1} \begin{pmatrix} B\tilde{B}_1^T \\ \vdots \\ B\tilde{B}_{\hat{n}}^T \end{pmatrix}$$

$$= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \begin{pmatrix} -\lambda_1 I - \hat{A} & & \\ & \ddots & \\ & & -\lambda_{\hat{n}} I - \hat{A} \end{pmatrix}^{-1} \begin{pmatrix} \hat{B}\tilde{B}_1^T \\ \vdots \\ \hat{B}\tilde{B}_{\hat{n}}^T \end{pmatrix}.$$

$\mathcal{H}_2$-Model Reduction

$\mathcal{H}_2$-Optimality Conditions

Let us assume $\hat{\Sigma}$ is given by its eigenvalue decomposition:

$$\hat{A} = R\Lambda R^{-1}, \quad \tilde{N}_i = R^{-1}\hat{N}_i R, \quad \tilde{B} = R^{-1}\hat{B}, \quad \tilde{C} = \hat{C}R.$$

Using Λ, $\tilde{N}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g.:

$$(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes A - \sum_{i=1}^m \tilde{N}_i \otimes N_i\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m)$$

$$= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes \hat{A} - \sum_{i=1}^m \tilde{N}_i \otimes \hat{N}_i\right)^{-1} \left(\tilde{B} \otimes \hat{B}\right) \mathrm{vec}(I_m).$$

$$H(-\lambda_\ell)\tilde{B}_\ell^T = \hat{H}(-\lambda_\ell)\tilde{B}_\ell^T$$

$\rightsquigarrow$ tangential interpolation at mirror images of reduced system poles

A First Iterative Approach

Algorithm 1 Bilinear IRKA

Input: $A, N_i, B, C, \hat{A}, \hat{N}_i, \hat{B}, \hat{C}$

Output: $A^{opt}, N_i^{opt}, B^{opt}, C^{opt}$

1: **while** (change in $\Lambda > \epsilon$) **do**

2: $R\Lambda R^{-1} = \hat{A}$, $\tilde{B} = R^{-1}\hat{B}$, $\tilde{C} = \hat{C}R$, $\tilde{N}_i = R^{-1}\hat{N}_i R$

3: $\mathrm{vec}(V) = \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes A - \sum_{i=1}^{m} \tilde{N}_i \otimes N_i\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m)$

4: $\mathrm{vec}(W) = \left(-\Lambda \otimes I_n - I_{\hat{n}} \otimes A^T - \sum_{i=1}^{m} \tilde{A}_i^T \otimes N_i^T\right)^{-1} \left(\tilde{C}^T \otimes C^T\right) \mathrm{vec}(I_q)$

5: $V = \mathrm{orth}(V)$, $W = \mathrm{orth}(W)$

6: $\hat{A} = \left(W^T V\right)^{-1} W^T A V$, $\hat{N}_i = \left(W^T V\right)^{-1} W^T N_i V$,
$\hat{B} = \left(W^T V\right)^{-1} W^T B$, $\hat{C} = CV$

7: **end while**

8: $A^{opt} = \hat{A}$, $N_i^{opt} = \hat{N}_i$, $B^{opt} = \hat{B}$, $C^{opt} = \hat{C}$

$\mathcal{H}_2$-Model Reduction for Bilinear Systems

A Heat Transfer Model

- 2-dimensional heat distribution
  [Benner, Saak '05]
- Boundary control by spraying intensities of a cooling fluid

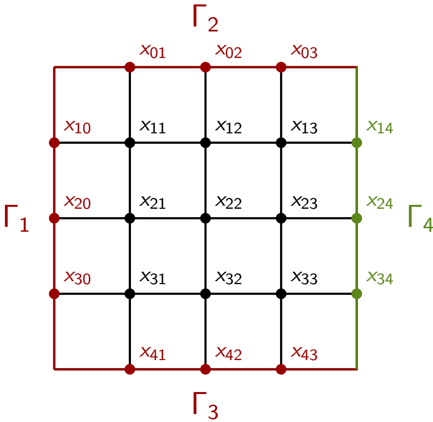
$$
\begin{aligned}
\Omega &= (0,1)\times(0,1), & \\
x_t &= \Delta x & \text{in } \Omega,\\
n\cdot\nabla x &= c\cdot u_{1,2,3}(x-1) & \text{on } \Gamma_1,\Gamma_2,\Gamma_3,\\
x &= u_4 & \text{on } \Gamma_4.
\end{aligned}
$$

- Spatial discretization $k\times k$-grid

$$\Rightarrow \dot{x} \approx A_1 x + \sum_{i=1}^{3} N_i x u_i + Bu$$

$$\Rightarrow A_2 = 0.$$

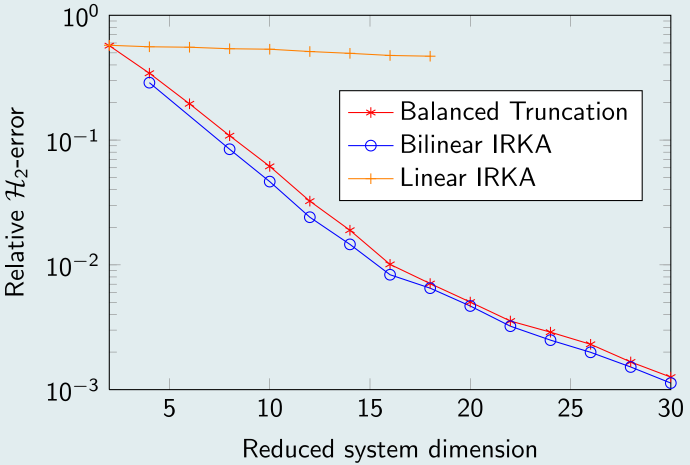
- Output: $y = \frac{1}{k^2}\begin{bmatrix}1 & \dots & 1\end{bmatrix}.$



$\mathcal{H}_2$-Model Reduction for Bilinear Systems

A Heat Transfer Model

Comparison of relative $\mathcal{H}_2$-error for $n = 10.000$



$\mathcal{H}_2$-Model Reduction for Bilinear Systems

Fokker-Planck Equation

As a second example, we consider a dragged Brownian particle whose one-dimensional motion is given by

$$dX_t = -\nabla V(X_t, t)dt + \sqrt{2\sigma}dW_t,$$

with $\sigma = \frac{2}{3}$ and $V(x, u) = W(x, t) + \Phi(x, u_t) = (x^2 - 1)^2 - xu - x$.
Alternatively, one can consider ([HARTMANN ET AL. '10]) ,

$$\rho(x, t)dx = \mathbf{P}\left[X_t \in [x, x + dx)\right]$$

which is described by the Fokker-Planck equation

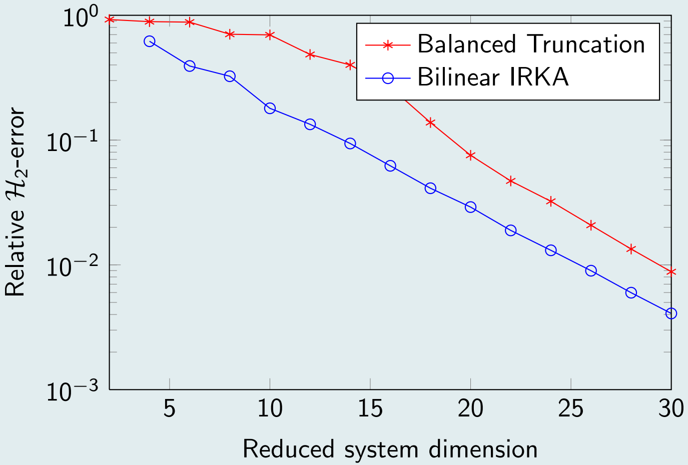
$$\begin{aligned}
\frac{\partial \rho}{\partial t} &= \sigma\Delta\rho + \nabla \cdot (\rho\nabla V), && (x, t) \in (-2, 2) \times (0, T],\\
0 &= \sigma\nabla\rho + \rho\nabla B, && (x, t) \in \{-2, 2\} \times [0, T],\\
\rho_0 &= \rho, && (x, t) \in (-2, 2) \times 0.
\end{aligned}$$

Output C discrete characteristic function of the interval $[0.95, 1.05]$.

$\mathcal{H}_2$-Model Reduction for Bilinear Systems

Fokker-Planck Equation

Comparison of relative $\mathcal{H}_2$-error for $n = 500$



Nonlinear Model Reduction

Quadratic-Bilinear Differential Algebraic Equations (QBDAEs)

Finally, we come back to the more general case with $f(x)$ nonlinear and $g(x) = B$. Here, the class of quadratic-bilinear differential algebraic equations

$$\Sigma : \quad \begin{cases} E\dot{x}(t) = A_1 x(t) + A_2 x(t) \otimes x(t) + N x(t) u(t) + B u(t), \\ \quad y(t) = C x(t), \quad x(0) = x_0, \end{cases}$$

where $E, A_1, N \in \mathbb{R}^{n \times n}, A_2 \in \mathbb{R}^{n \times n^2}$ (Hessian tensor), $B, C^T \in \mathbb{R}^n$ are quite helpful.

- A large class of smooth nonlinear control-affine systems can be transformed into the above type of control system.
- The transformation is exact, but a slight increase of the state dimension has to be accepted.
- Input-output behavior can be characterized by generalized transfer functions $\rightsquigarrow$ enables us to use Krylov-based reduction techniques.

Nonlinear Model Reduction

Transformation via McCormick Relaxation

Theorem [Gu'09]

Assume that the state equation of a nonlinear system Σ is given by

$$\dot{x} = a_0 x + a_1 g_1(x) + \ldots + a_k g_k(x) + Bu,$$

where $g_i(x) : \mathbb{R}^n \to \mathbb{R}^n$ are compositions of uni-variable rational, exponential, logarithmic, trigonometric or root functions, respectively. Then, by iteratively taking derivatives and adding algebraic equations, respectively, Σ can be transformed into a system of QBDAEs.

Nonlinear Model Reduction

Transformation via McCormick Relaxation

Theorem [Gu'09]

Assume that the state equation of a nonlinear system Σ is given by

$$\dot{x} = a_0 x + a_1 g_1(x) + \ldots + a_k g_k(x) + Bu,$$

where $g_i(x) : \mathbb{R}^n \to \mathbb{R}^n$ are compositions of uni-variable rational, exponential, logarithmic, trigonometric or root functions, respectively. Then, by iteratively taking derivatives and adding algebraic equations, respectively, Σ can be transformed into a system of QBDAEs.

Example

- $\dot{x}_1 = \exp(-x_2) \cdot \sqrt{x_1^2 + 1}, \quad \dot{x}_2 = -x_2 + u.$

Nonlinear Model Reduction

Transformation via McCormick Relaxation

Theorem [Gu'09]

Assume that the state equation of a nonlinear system Σ is given by

$$\dot{x} = a_0 x + a_1 g_1(x) + \ldots + a_k g_k(x) + Bu,$$

where $g_i(x) : \mathbb{R}^n \to \mathbb{R}^n$ are compositions of uni-variable rational, exponential, logarithmic, trigonometric or root functions, respectively. Then, by iteratively taking derivatives and adding algebraic equations, respectively, Σ can be transformed into a system of QBDAEs.

Example

- $\dot{x}_1 = \exp(-x_2) \cdot \sqrt{x_1^2 + 1}, \quad \dot{x}_2 = -x_2 + u.$
- $z_1 := \exp(-x_2),$

Nonlinear Model Reduction

Transformation via McCormick Relaxation

Theorem [Gu'09]

Assume that the state equation of a nonlinear system Σ is given by

$$\dot{x} = a_0 x + a_1 g_1(x) + \ldots + a_k g_k(x) + Bu,$$

where $g_i(x) : \mathbb{R}^n \to \mathbb{R}^n$ are compositions of uni-variable rational, exponential, logarithmic, trigonometric or root functions, respectively. Then, by iteratively taking derivatives and adding algebraic equations, respectively, Σ can be transformed into a system of QBDAEs.

Example

- $\dot{x}_1 = \exp(-x_2) \cdot \sqrt{x_1^2 + 1}, \quad \dot{x}_2 = -x_2 + u.$
- $z_1 := \exp(-x_2), \quad z_2 := \sqrt{x_1^2 + 1}.$

Nonlinear Model Reduction

Transformation via McCormick Relaxation

Theorem [Gu'09]

Assume that the state equation of a nonlinear system Σ is given by

$$\dot{x} = a_0 x + a_1 g_1(x) + \ldots + a_k g_k(x) + Bu,$$

where $g_i(x) : \mathbb{R}^n \to \mathbb{R}^n$ are compositions of uni-variable rational, exponential, logarithmic, trigonometric or root functions, respectively. Then, by iteratively taking derivatives and adding algebraic equations, respectively, Σ can be transformed into a system of QBDAEs.

Example

- $\dot{x}_1 = \exp(-x_2) \cdot \sqrt{x_1^2 + 1}, \quad \dot{x}_2 = -x_2 + u.$
- $z_1 := \exp(-x_2), \quad z_2 := \sqrt{x_1^2 + 1}.$
- $\dot{x}_1 = z_1 \cdot z_2,$

Nonlinear Model Reduction

Transformation via McCormick Relaxation

Theorem [Gu'09]

Assume that the state equation of a nonlinear system Σ is given by

$$\dot{x} = a_0 x + a_1 g_1(x) + \ldots + a_k g_k(x) + Bu,$$

where $g_i(x) : \mathbb{R}^n \to \mathbb{R}^n$ are compositions of uni-variable rational, exponential, logarithmic, trigonometric or root functions, respectively. Then, by iteratively taking derivatives and adding algebraic equations, respectively, Σ can be transformed into a system of QBDAEs.

Example

- $\dot{x}_1 = \exp(-x_2) \cdot \sqrt{x_1^2 + 1}, \quad \dot{x}_2 = -x_2 + u.$
- $z_1 := \exp(-x_2), \quad z_2 := \sqrt{x_1^2 + 1}.$
- $\dot{x}_1 = z_1 \cdot z_2, \quad \dot{x}_2 = -x_2 + u,$

Nonlinear Model Reduction

Transformation via McCormick Relaxation

Theorem [Gu'09]

Assume that the state equation of a nonlinear system Σ is given by

$$\dot{x} = a_0 x + a_1 g_1(x) + \ldots + a_k g_k(x) + Bu,$$

where $g_i(x) : \mathbb{R}^n \to \mathbb{R}^n$ are compositions of uni-variable rational, exponential, logarithmic, trigonometric or root functions, respectively. Then, by iteratively taking derivatives and adding algebraic equations, respectively, Σ can be transformed into a system of QBDAEs.

Example

- $\dot{x}_1 = \exp(-x_2) \cdot \sqrt{x_1^2 + 1}, \quad \dot{x}_2 = -x_2 + u.$
- $z_1 := \exp(-x_2), \quad z_2 := \sqrt{x_1^2 + 1}.$
- $\dot{x}_1 = z_1 \cdot z_2, \quad \dot{x}_2 = -x_2 + u, \quad \dot{z}_1 = -z_1 \cdot (-x_2 + u),$

Nonlinear Model Reduction

Transformation via McCormick Relaxation

Theorem [Gu'09]

Assume that the state equation of a nonlinear system Σ is given by

$$\dot{x} = a_0 x + a_1 g_1(x) + \ldots + a_k g_k(x) + Bu,$$

where $g_i(x) : \mathbb{R}^n \to \mathbb{R}^n$ are compositions of uni-variable rational, exponential, logarithmic, trigonometric or root functions, respectively. Then, by iteratively taking derivatives and adding algebraic equations, respectively, Σ can be transformed into a system of QBDAEs.

Example

- $\dot{x}_1 = \exp(-x_2) \cdot \sqrt{x_1^2 + 1}, \quad \dot{x}_2 = -x_2 + u.$
- $z_1 := \exp(-x_2), \quad z_2 := \sqrt{x_1^2 + 1}.$
- $\dot{x}_1 = z_1 \cdot z_2, \quad \dot{x}_2 = -x_2 + u, \quad \dot{z}_1 = -z_1 \cdot (-x_2 + u),$
  $\dot{z}_2 = \frac{2 \cdot x_1 \cdot z_1 \cdot z_2}{2 \cdot z_2} = x_1 \cdot z_1.$

Nonlinear Model Reduction

Variational Analysis and Linear Subsystems

Analysis of nonlinear systems by variational equation approach:

Nonlinear Model Reduction

Variational Analysis and Linear Subsystems

Analysis of nonlinear systems by variational equation approach:

- consider input of the form $\alpha u(t)$,

Nonlinear Model Reduction

Variational Analysis and Linear Subsystems

Analysis of nonlinear systems by variational equation approach:

- consider input of the form $\alpha u(t)$,
- nonlinear system is assumed to be a series of homogeneous nonlinear subsystems, i.e. response should be of the form

$$x(t) = \alpha x_1(t) + \alpha^2 x_2(t) + \alpha^3 x_3(t) + \dots .$$

Nonlinear Model Reduction

Variational Analysis and Linear Subsystems

Analysis of nonlinear systems by variational equation approach:

- consider input of the form $\alpha u(t)$,
- nonlinear system is assumed to be a series of homogeneous nonlinear subsystems, i.e. response should be of the form

$$x(t) = \alpha x_1(t) + \alpha^2 x_2(t) + \alpha^3 x_3(t) + \dots .$$

- comparison of terms $\alpha^i, i = 1, 2, \dots$ leads to series of systems

$$\begin{aligned} E\dot{x}_1 &= A_1 x_1 + Bu, \\ E\dot{x}_2 &= A_1 x_2 + A_2 x_1 \otimes x_1 + N x_1 u, \\ E\dot{x}_3 &= A_1 x_3 + A_2 (x_1 \otimes x_2 + x_2 \otimes x_1) + N x_2 u \\ &\ \ \vdots \end{aligned}$$

Nonlinear Model Reduction

Variational Analysis and Linear Subsystems

Analysis of nonlinear systems by variational equation approach:

- consider input of the form $\alpha u(t)$,
- nonlinear system is assumed to be a series of homogeneous nonlinear subsystems, i.e. response should be of the form

$$x(t) = \alpha x_1(t) + \alpha^2 x_2(t) + \alpha^3 x_3(t) + \dots .$$

- comparison of terms $\alpha^i, i = 1, 2, \dots$ leads to series of systems

$$\begin{aligned} E\dot{x}_1 &= A_1 x_1 + Bu, \\ E\dot{x}_2 &= A_1 x_2 + A_2 x_1 \otimes x_1 + N x_1 u, \\ E\dot{x}_3 &= A_1 x_3 + A_2 \left(x_1 \otimes x_2 + x_2 \otimes x_1 \right) + N x_2 u \\ &\vdots \end{aligned}$$

- although i-th subsystem is coupled nonlinearly to preceding systems, linear systems are obtained if terms x_j, $j < i$, are interpreted as pseudo-inputs.

Nonlinear Model Reduction

Generalized Transfer Functions

In a similar way, a series of generalized symmetric transfer functions can be obtained via the growing exponential approach:

Nonlinear Model Reduction

Generalized Transfer Functions

In a similar way, a series of generalized symmetric transfer functions can be obtained via the growing exponential approach:

$$H_1(s_1) = C \underbrace{(s_1 E - A_1)^{-1} B}_{G_1(s_1)},$$

Nonlinear Model Reduction

Generalized Transfer Functions

In a similar way, a series of generalized symmetric transfer functions can be obtained via the growing exponential approach:

$$H_1(s_1) = C\underbrace{(s_1 E - A_1)^{-1} B}_{G_1(s_1)},$$

$$H_2(s_1, s_2) = \frac{1}{2!} C\left((s_1 + s_2)E - A_1\right)^{-1} \left[N\left(G_1(s_1) + G_1(s_2)\right) + A_2\left(G_1(s_1) \otimes G_1(s_2) + G_1(s_2) \otimes G_1(s_1)\right)\right],$$

Nonlinear Model Reduction

Generalized Transfer Functions

In a similar way, a series of generalized symmetric transfer functions can be obtained via the growing exponential approach:

$$H_1(s_1) = C\underbrace{(s_1E - A_1)^{-1}B}_{G_1(s_1)},$$

$$H_2(s_1, s_2) = \frac{1}{2!}C\left((s_1+s_2)E - A_1\right)^{-1}\left[N\left(G_1(s_1) + G_1(s_2)\right) \right.$$
$$\left. +A_2\left(G_1(s_1)\otimes G_1(s_2) + G_1(s_2)\otimes G_1(s_1)\right)\right],$$

$$H_3(s_1, s_2, s_3) = \frac{1}{3!}C\left((s_1+s_2+s_3)E - A_1\right)^{-1}$$
$$\Big[N(G_2(s_1,s_2) + G_2(s_2,s_3) + G_2(s_1,s_3))$$
$$+ A_2\big(G_1(s_1)\otimes G_2(s_2,s_3) + G_1(s_2)\otimes G_2(s_1,s_3)$$
$$+ G_1(s_3)\otimes G_2(s_1,s_3) + G_2(s_2,s_3)\otimes G_1(s_1)$$
$$+ G_2(s_1,s_3)\otimes G_1(s_2) + G_2(s_1,s_2)\otimes G_1(s_3)\big)\Big].$$

Nonlinear Model Reduction

Characterization via Multimoments

For simplicity, focus on the first two transfer functions. For $H_1(s_1)$, choosing σ and making use of the Neumann lemma leads to

$$H_1(s_1) = \sum_{i=0}^{\infty} C \underbrace{\left((A_1 - \sigma E)^{-1} E\right)^i (A_1 - \sigma E)^{-1} B \, (s_1 - \sigma)^i}_{m^i_{s_1,\sigma}}.$$

Nonlinear Model Reduction

Characterization via Multimoments

For simplicity, focus on the first two transfer functions. For $H_1(s_1)$, choosing σ and making use of the Neumann lemma leads to

$$H_1(s_1) = \sum_{i=0}^{\infty} C \underbrace{\left((A_1 - \sigma E)^{-1} E\right)^i (A_1 - \sigma E)^{-1} B \, (s_1 - \sigma)^i}_{m^i_{s_1,\sigma}}.$$

Similarly, specifying an expansion point (τ, ξ) yields

$$H_2(s_1, s_2) = \frac{1}{2} \sum_{i=0}^{\infty} C \left((A_1 - (\tau + \xi)E)^{-1} E\right)^i (A_1 - (\tau + \xi)E)^{-1} \, (s_1 + s_2 - \tau - \xi)^i \cdot$$
$$\left[A_2 \left(\sum_{j=0}^{\infty} m^j_{s_1,\tau} \otimes \sum_{k=0}^{\infty} m^k_{s_2,\xi} + \sum_{k=0}^{\infty} m^k_{s_2,\xi} \otimes \sum_{j=0}^{\infty} m^j_{s_1,\tau}\right) + N \left(\sum_{p=0}^{\infty} m^p_{s_1,\tau} + \sum_{p=0}^{\infty} m^q_{s_2,\xi}\right)\right]$$

Nonlinear Model Reduction

Constructing the Projection Matrix

For derivatives around $\sigma = \tau = \xi$ up to order $q - 1$, construct the Krylov spaces:

Nonlinear Model Reduction

Constructing the Projection Matrix

For derivatives around $\sigma = \tau = \xi$ up to order $q-1$, construct the Krylov spaces:

$$U = \mathcal{K}_q\left((A_1 - \sigma E)^{-1}E, (A_1 - \sigma E)^{-1}B\right)$$

Nonlinear Model Reduction

Constructing the Projection Matrix

For derivatives around $\sigma = \tau = \xi$ up to order $q-1$, construct the Krylov spaces:

$$
\begin{aligned}
U =& \mathcal{K}_q\left((A_1 - \sigma E)^{-1} E, (A_1 - \sigma E)^{-1} B\right) \\
&\textbf{for } i = 1 : q \\
&\quad W_i = \mathcal{K}_{q-i+1}\left((A_1 - 2\sigma E)^{-1} E, (A_1 - 2\sigma E)^{-1} N U_i\right),
\end{aligned}
$$

Nonlinear Model Reduction

Constructing the Projection Matrix

For derivatives around $\sigma = \tau = \xi$ up to order $q-1$, construct the Krylov spaces:

$$
\begin{aligned}
U =& \mathcal{K}_q \left((A_1 - \sigma E)^{-1} E, (A_1 - \sigma E)^{-1} B\right) \\
&\textbf{for } i = 1 : q \\
&\quad W_i = \mathcal{K}_{q-i+1} \left((A_1 - 2\sigma E)^{-1} E, (A_1 - 2\sigma E)^{-1} N U_i\right), \\
&\qquad \textbf{for } j = 1 : \min(q-i+1, i) \\
&\qquad\quad Z_i = \mathcal{K}_{q-i-j+2} \left((A_1 - 2\sigma E)^{-1} E, (A_1 - 2\sigma E)^{-1} A_2 U_i \otimes U_j\right),
\end{aligned}
$$

U_i denoting the i-*th* column of U.

Nonlinear Model Reduction

Constructing the Projection Matrix

For derivatives around $\sigma = \tau = \xi$ up to order $q-1$, construct the Krylov spaces:

$$
\begin{aligned}
U =& \mathcal{K}_q\left((A_1-\sigma E)^{-1}E, (A_1-\sigma E)^{-1}B\right) \\
&\textbf{for } i = 1:q \\
&\quad W_i = \mathcal{K}_{q-i+1}\left((A_1-2\sigma E)^{-1}E, (A_1-2\sigma E)^{-1}NU_i\right), \\
&\qquad \textbf{for } j = 1:\min(q-i+1, i) \\
&\qquad\quad Z_i = \mathcal{K}_{q-i-j+2}\left((A_1-2\sigma E)^{-1}E, (A_1-2\sigma E)^{-1}A_2U_i\otimes U_j\right),
\end{aligned}
$$

U_i denoting the i-*th* column of U. Set $V = \mathrm{orth}([U, W, Z])$ and construct $\hat{\Sigma}$ by the Galerkin-Projection $\mathcal{P} = VV^T$:

$$
\hat{A}_1 = V^TA_1V \in \mathbb{R}^{\hat{n}\times\hat{n}}, \quad \hat{A}_2 = V^TA_2V\otimes V \in \mathbb{R}^{\hat{n}\times\hat{n}^2},
$$

$$
\hat{N} = V^TNV \in \mathbb{R}^{\hat{n}\times\hat{n}}, \quad \hat{B} = V^TB \in \mathbb{R}^{\hat{n}}, \quad \hat{C}^T = V^TC \in \mathbb{R}^{\hat{n}}.
$$

Nonlinear Model Reduction

Two-Sided Projection Methods

Similarly to the linear case, one can exploit duality concepts, in order to construct two-sided projection methods.

Nonlinear Model Reduction

Two-Sided Projection Methods

Similarly to the linear case, one can exploit duality concepts, in order to construct two-sided projection methods.

Interpreting A_2 as the matricization of a 3-tensor from $\mathbb{R}^{n\times n\times n}$. one can show that the dual Krylov spaces have to be constructed as follows

$$\tilde{U} = \mathcal{K}_q\left((A_1-\sigma E)^{-T}E^T, (A_1-\sigma E)^{-T}C^T\right)$$

for $i = 1 : q$

$$\tilde{W}_i = \mathcal{K}_{q-i+1}\left((A_1-2\sigma E)^{-T}E^T, (A_1-2\sigma E)^{-T}N^T\tilde{U}_i\right),$$

for $j = 1 : \min(q-i+1, i)$

$$\tilde{Z}_i = \mathcal{K}_{q-i-j+2}\left((A_1-2\sigma E)^{-T}E^T, (A_1-2\sigma E)^{-T}\tilde{A}_2 U_i \otimes \tilde{U}_j\right),$$

where $\tilde{A}_2$ now is another matricization of the Hessian tensor.

Nonlinear Model Reduction

Two-Sided Projection Methods

Similarly to the linear case, one can exploit duality concepts, in order to construct two-sided projection methods.

Interpreting A_2 as the matricization of a 3-tensor from $\mathbb{R}^{n\times n\times n}$. one can show that the dual Krylov spaces have to be constructed as follows

$$
\begin{aligned}
&\tilde{U} = \mathcal{K}_q\left((A_1-\sigma E)^{-T}E^T,(A_1-\sigma E)^{-T}C^T\right)\\
&\textbf{for } i=1:q\\
&\quad \tilde{W}_i = \mathcal{K}_{q-i+1}\left((A_1-2\sigma E)^{-T}E^T,(A_1-2\sigma E)^{-T}N^T\tilde{U}_i\right),\\
&\qquad \textbf{for } j=1:\min(q-i+1,i)\\
&\qquad\quad \tilde{Z}_i = \mathcal{K}_{q-i-j+2}\left((A_1-2\sigma E)^{-T}E^T,(A_1-2\sigma E)^{-T}\tilde{A}_2 U_i\otimes\tilde{U}_j\right),
\end{aligned}
$$

where $\tilde{A}_2$ now is another matricization of the Hessian tensor.
Note: If one uses the third matricization, then $U_i \otimes \tilde{U}_j$ has to be replaced by $\tilde{U}_i \otimes U_j$. For matricizations, see e.g. [Kressner/Tobler '10].

Nonlinear Model Reduction

The FitzHugh-Nagumo System

- FitzHugh-Nagumo system modeling a neuron

  [Chaturantabut, Sorensen '09]

$$
\begin{aligned}
\epsilon v_t(x,t) &= \epsilon^2 v_{xx}(x,t) + f(v(x,t)) - w(x,t) + g, \\
w_t(x,t) &= hv(x,t) - \gamma w(x,t) + g,
\end{aligned}
$$

  with $f(v) = v(v-0.1)(1-v)$ and initial and boundary conditions

$$
\begin{aligned}
v(x,0) &= 0, & w(x,0) &= 0, & x &\in [0,1], \\
v_x(0,t) &= -i_0(t), & v_x(1,t) &= 0, & t &\geq 0,
\end{aligned}
$$

  where $\epsilon = 0.015,\ h = 0.5,\ \gamma = 2,\ g = 0.05,\ i_0(t) = 5 \cdot 10^4 t^3 \exp(-15t)$
- original state dimension $n = 2 \cdot 400$, reduced state dimension $\hat{n} = 26$, chosen interpolation point $\sigma = 1$
- 3D phase space

Nonlinear Model Reduction

Two-Dimensional Burgers Equation

- 2D-Burgers Equation on $\underbrace{(0,1)\times(0,1)}_{:=\Omega}\times[0,T]$

$$u_t = -\left(u\cdot\nabla\right)u + \nu\Delta u$$

with $u(x,y,t)\in\mathbb{R}^2$ describing the motion of a compressible fluid.

Nonlinear Model Reduction

Two-Dimensional Burgers Equation

- 2D-Burgers Equation on $\underbrace{(0,1)\times(0,1)}_{:=\Omega}\times[0,T]$

$$u_t = -\left(u\cdot\nabla\right)u + \nu\Delta u$$

with $u(x,y,t)\in\mathbb{R}^2$ describing the motion of a compressible fluid.

- Consider initial and boundary conditions

$$\begin{aligned}
&u_x(x,y,0)=\frac{\sqrt{2}}{2}, && u_y(x,y,0)=\frac{\sqrt{2}}{2}, && \text{for } x,y\in\Omega_1:=(0,0.5],\\
&u_x(x,y,0)=0, && u_y(x,y,0)=0, && \text{for } x,y,\in\Omega\backslash\Omega_1,\\
&u_x=0, && u_y=0, && \text{for } x,y,\in\partial\Omega.
\end{aligned}$$

Nonlinear Model Reduction

Two-Dimensional Burgers Equation

- 2D-Burgers Equation on $\underbrace{(0,1)\times(0,1)}_{:=\Omega}\times[0,T]$

$$u_t = -\left(u\cdot\nabla\right)u + \nu\Delta u$$

  with $u(x,y,t)\in\mathbb{R}^2$ describing the motion of a compressible fluid.
- Consider initial and boundary conditions

$$\begin{aligned}
u_x(x,y,0) &= \frac{\sqrt{2}}{2}, & u_y(x,y,0) &= \frac{\sqrt{2}}{2}, && \text{for } x,y\in\Omega_1 := (0,0.5],\\
u_x(x,y,0) &= 0, & u_y(x,y,0) &= 0, && \text{for } x,y,\in\Omega\backslash\Omega_1,\\
u_x &= 0, & u_y &= 0, && \text{for } x,y,\in\partial\Omega.
\end{aligned}$$

- Spatial discretization $\rightsquigarrow$ QBDAE system with nonzero i.c. and $N=0$. $\rightsquigarrow$ reformulate as system with zero i.c. and constant input.

Nonlinear Model Reduction

Two-Dimensional Burgers Equation

- 2D-Burgers Equation on $\underbrace{(0,1)\times(0,1)}_{:=\Omega}\times[0,T]$

$$u_t = -\left(u\cdot\nabla\right)u + \nu\Delta u$$

  with $u(x,y,t)\in\mathbb{R}^2$ describing the motion of a compressible fluid.
- Consider initial and boundary conditions

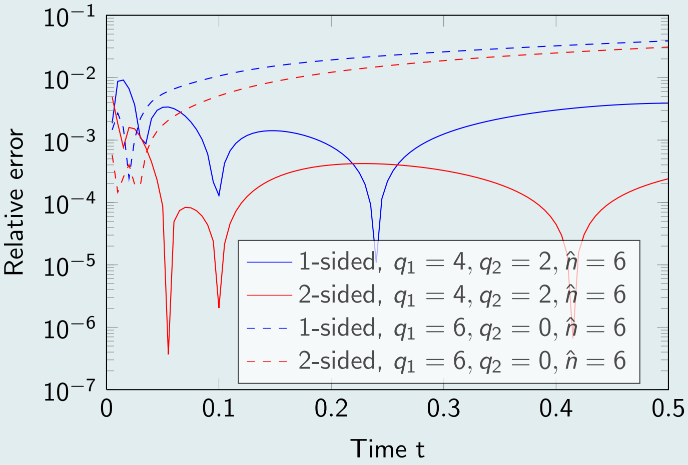
$$\begin{aligned} u_x(x,y,0) &= \frac{\sqrt{2}}{2}, & u_y(x,y,0) &= \frac{\sqrt{2}}{2}, & &\text{for } x,y\in\Omega_1 := (0,0.5], \\ u_x(x,y,0) &= 0, & u_y(x,y,0) &= 0, & &\text{for } x,y,\in\Omega\backslash\Omega_1, \\ u_x &= 0, & u_y &= 0, & &\text{for } x,y,\in\partial\Omega. \end{aligned}$$

- Spatial discretization $\rightsquigarrow$ QBDAE system with nonzero i.c. and $N=0$. $\rightsquigarrow$ reformulate as system with zero i.c. and constant input.
- Output C chosen to be average x-velocity.

Nonlinear Model Reduction

Two-Dimensional Burgers Equation

Comparison of relative time-domain error for $n = 1600$



Nonlinear Model Reduction

Two-Dimensional Burgers Equation

- 2D-Burgers Equation on $\underbrace{(0,1)\times(0,1)}_{:=\Omega}\times[0,T]$

$$u_t = -(u\cdot\nabla)\,u + \nu\Delta u$$

  with $u(x,y,t)\in\mathbb{R}^2$ describing the motion of a compressible fluid.

- Now consider initial and boundary conditions

$$\begin{aligned}
&u_x(x,y,0)=0, && u_y(x,y,0)=0, && \text{for } x,y\in\Omega,\\
&u_x=\cos(\pi t), && u_y=\cos(2\pi t), && \text{for } (x,y)\in\{0,1\}\times(0,1),\\
&u_x=\sin(\pi t), && u_y=\sin(2\pi t), && \text{for } (x,y)\in(0,1)\times\{0,1\}.
\end{aligned}$$

Nonlinear Model Reduction

Two-Dimensional Burgers Equation

- 2D-Burgers Equation on $\underbrace{(0,1)\times(0,1)}_{:=\Omega}\times[0,T]$

$$u_t = -(u\cdot\nabla)\,u + \nu\Delta u$$

with $u(x,y,t)\in\mathbb{R}^2$ describing the motion of a compressible fluid.

- Now consider initial and boundary conditions

$$\begin{aligned}
&u_x(x,y,0)=0, && u_y(x,y,0)=0, && \text{for } x,y\in\Omega,\\
&u_x=\cos(\pi t), && u_y=\cos(2\pi t), && \text{for } (x,y)\in\{0,1\}\times(0,1),\\
&u_x=\sin(\pi t), && u_y=\sin(2\pi t), && \text{for } (x,y)\in(0,1)\times\{0,1\}.
\end{aligned}$$

- Spatial discretization $\rightsquigarrow$ QBDAE system with zero i.c. and 4 inputs $B\in\mathbb{R}^{n\times 4}$, N_1,N_2,N_3,N_4, ROM with $q_1=5, q_2=2, \sigma=0, \hat{n}=52$.

Nonlinear Model Reduction

Two-Dimensional Burgers Equation

- 2D-Burgers Equation on $\underbrace{(0,1)\times(0,1)}_{:=\Omega}\times[0,T]$

$$u_t = -\left(u\cdot\nabla\right)u + \nu\Delta u$$

  with $u(x,y,t)\in\mathbb{R}^2$ describing the motion of a compressible fluid.

- Now consider initial and boundary conditions

$$\begin{aligned}
u_x(x,y,0) &= 0, & u_y(x,y,0) &= 0, && \text{for } x,y\in\Omega,\\
u_x &= \cos(\pi t), & u_y &= \cos(2\pi t), && \text{for } (x,y)\in\{0,1\}\times(0,1),\\
u_x &= \sin(\pi t), & u_y &= \sin(2\pi t), && \text{for } (x,y)\in(0,1)\times\{0,1\}.
\end{aligned}$$

- Spatial discretization $\rightsquigarrow$ QBDAE system with zero i.c. and 4 inputs $B\in\mathbb{R}^{n\times 4}$, N_1,N_2,N_3,N_4, ROM with $q_1=5, q_2=2, \sigma=0, \hat{n}=52$.
- State reconstruction by reduced model $x\approx V\hat{x}$, max. rel. err $< 3\%$.

Conclusions

- A lot of linear reduction techniques can be transferred to the nonlinear case.
- We have shown a generalized interpolation theory for bilinear control systems $\rightsquigarrow$ $\mathcal{H}_2$-optimal model reduction.
- Many nonlinear dynamics can be expressed by a system of quadratic-bilinear differential algebraic equations.
- There exist Krylov subspace methods that extend the concept of moment-matching $\rightsquigarrow$ using basic tools from tensor theory allow for better approximations.

Conclusions

- A lot of linear reduction techniques can be transferred to the nonlinear case.
- We have shown a generalized interpolation theory for bilinear control systems $\rightsquigarrow$ $\mathcal{H}_2$-optimal model reduction.
- Many nonlinear dynamics can be expressed by a system of quadratic-bilinear differential algebraic equations.
- There exist Krylov subspace methods that extend the concept of moment-matching $\rightsquigarrow$ using basic tools from tensor theory allow for better approximations.

Thank you for your attention. . .