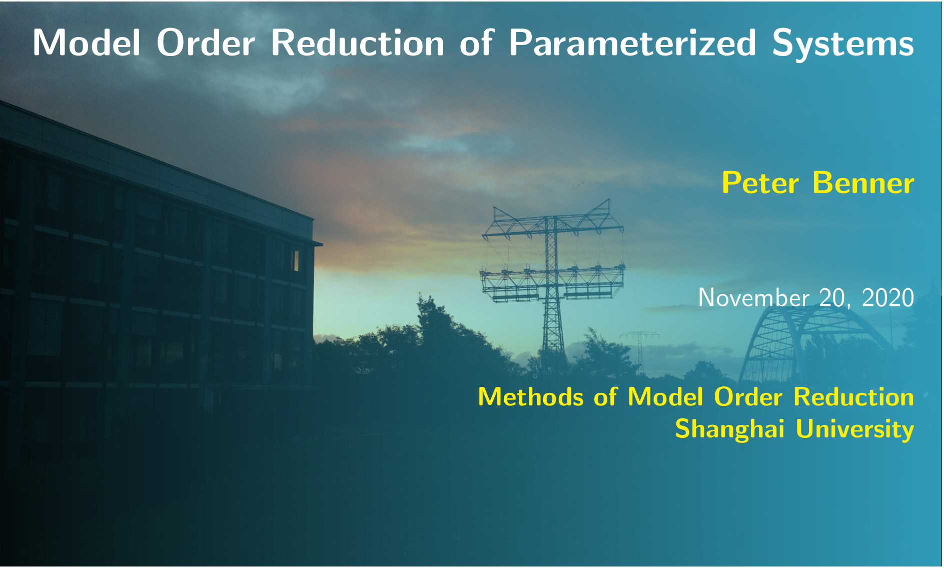
MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

CSC COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY



Model Order Reduction of Parameterized Systems

Peter Benner

November 20, 2020

Methods of Model Order Reduction
Shanghai University

Main Literature for this Lecture

P. Benner, S. Gugercin, and K. Willcox.
A Survey of Model Reduction Methods for Parametric Systems.
SIAM Review 57(4):483–531, 2015.

Outline

Outline

Parametric Dynamical Systems

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) &= f(t, x(t;p), u(t), p), \quad x(t_0) = x_0, & \text{(a)} \\ y(t;p) &= g(t, x(t;p), u(t), p) & \text{(b)} \end{cases}$$

with

- (generalized) states $x(t;p) \in \mathbb{R}^n$ ($E(p) \in \mathbb{R}^{n \times n}$),
- inputs (controls) $u(t) \in \mathbb{R}^m$,
- outputs (measurements, quantities of interest) $y(t;p) \in \mathbb{R}^q$,
  (b) is called output equation,
- $p \in \Omega \subset \mathbb{R}^d$ is a parameter vector, Ω is bounded.

Parametric Dynamical Systems

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) = f(t, x(t;p), u(t), p), \quad x(t_0) = x_0, & \text{(a)} \\ y(t;p) = g(t, x(t;p), u(t), p) & \text{(b)} \end{cases}$$

with

- (generalized) states $x(t;p) \in \mathbb{R}^n$ ($E(p) \in \mathbb{R}^{n \times n}$),
- inputs (controls) $u(t) \in \mathbb{R}^m$,
- outputs (measurements, quantities of interest) $y(t;p) \in \mathbb{R}^q$,
  (b) is called output equation,
- $p \in \Omega \subset \mathbb{R}^d$ is a parameter vector, Ω is bounded.

$E(p)$ singular $\Rightarrow$ (a) is system of differential-algebraic equations (DAEs)
otherwise $\Rightarrow$ (a) is system of ordinary differential equations (ODEs)

Parametric Dynamical Systems

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) &= f(t, x(t;p), u(t), p), \quad x(t_0) = x_0, & \text{(a)} \\ y(t;p) &= g(t, x(t;p), u(t), p) & \text{(b)} \end{cases}$$

with

- (generalized) states $x(t;p) \in \mathbb{R}^n$ ($E(p) \in \mathbb{R}^{n \times n}$),
- inputs (controls) $u(t) \in \mathbb{R}^m$,
- outputs (measurements, quantities of interest) $y(t;p) \in \mathbb{R}^q$, (b) is called output equation,
- $p \in \Omega \subset \mathbb{R}^d$ is a parameter vector, Ω is bounded.

Applications:

- Repeated simulation for varying material or geometry parameters, boundary conditions,
- control, optimization and design,
- of models, often generated by FE software (e.g., ANSYS, NASTRAN,...) or automatic tools (e.g., Modelica).

Parametric Dynamical Systems

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) &= f(t, x(t;p), u(t), p), \quad x(t_0) = x_0, & \text{(a)} \\ y(t;p) &= g(t, x(t;p), u(t), p) & \text{(b)} \end{cases}$$

with

- (generalized) states $x(t;p) \in \mathbb{R}^n$ ($E(p) \in \mathbb{R}^{n \times n}$),
- inputs (controls) $u(t) \in \mathbb{R}^m$,
- outputs (measurements, quantities of interest) $y(t;p) \in \mathbb{R}^q$,
  (b) is called output equation,
- $p \in \Omega \subset \mathbb{R}^d$ is a parameter vector, Ω is bounded.

Underlying PDE and boundary conditions often not accessible!

Parametric Dynamical Systems

$$\Sigma(p):\begin{cases} E(p)\dot{x}(t;p) &= f(t,x(t;p),u(t),p), \quad x(t_0)=x_0, & \text{(a)}\\ y(t;p) &= g(t,x(t;p),u(t),p) & \text{(b)}\end{cases}$$

with

- (generalized) states $x(t;p)\in\mathbb{R}^n$ ($E(p)\in\mathbb{R}^{n\times n}$),
- inputs (controls) $u(t)\in\mathbb{R}^m$,
- outputs (measurements, quantities of interest) $y(t;p)\in\mathbb{R}^q$,
  (b) is called output equation,
- $p\in\Omega\subset\mathbb{R}^d$ is a parameter vector, Ω is bounded.

Underlying PDE and boundary conditions often not accessible!

Parametric discretized model often not available,

but matrices for certain parameter values can be extracted

(or output data for given u and p can be generated!)

$$
\begin{aligned}
E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \quad && A(p),\ E(p) \in \mathbb{R}^{n\times n},\\
y(t;p) &= C(p)x(t;p), && B(p) \in \mathbb{R}^{n\times m},\ C(p) \in \mathbb{R}^{q\times n}.
\end{aligned}
$$

$$
\begin{aligned}
E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \quad & A(p),\ E(p) \in \mathbb{R}^{n\times n},\\
y(t;p) &= C(p)x(t;p), & B(p) \in \mathbb{R}^{n\times m},\ C(p) \in \mathbb{R}^{q\times n}.
\end{aligned}
$$

Laplace Transformation / Frequency Domain

Application of Laplace transformation

$$x(t;p) \mapsto x(s;p), \quad \dot{x}(t;p) \mapsto sx(s;p)$$

to linear system with $x(0;p) \equiv 0$:

$$sE(p)x(s;p) = A(p)x(s;p) + B(p)u(s), \quad y(s;p) = C(p)x(s;p),$$

yields I/O-relation in frequency domain:

$$y(s;p) = \Big(\underbrace{C(p)(sE(p) - A(p))^{-1}B(p)}_{=:G(s,p)} \Big) u(s).$$

$G(s,p)$ is the parameter-dependent transfer function of $\Sigma(p)$.

$$\begin{aligned} E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \quad & A(p),\ E(p) \in \mathbb{R}^{n\times n}, \\ y(t;p) &= C(p)x(t;p), & B(p) \in \mathbb{R}^{n\times m},\ C(p) \in \mathbb{R}^{q\times n}. \end{aligned}$$

Laplace Transformation / Frequency Domain

Application of Laplace transformation

$$x(t;p) \mapsto x(s;p), \quad \dot{x}(t;p) \mapsto sx(s;p)$$

to linear system with $x(0;p) \equiv 0$:

$$sE(p)x(s;p) = A(p)x(s;p) + B(p)u(s), \quad y(s;p) = C(p)x(s;p),$$

yields I/O-relation in frequency domain:

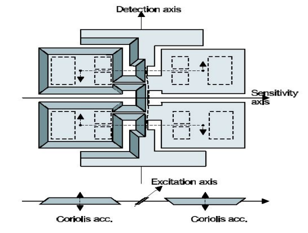
$$y(s;p) = \Bigg(\underbrace{C(p)(sE(p) - A(p))^{-1}B(p)}_{=:G(s,p)} \Bigg) u(s).$$

$G(s,p)$ is the parameter-dependent transfer function of $\Sigma(p)$.

Goal: Fast evaluation of mapping $(u,p) \;\rightarrow\; y(s;p)$.

Microgyroscope (butterfly gyro)

- Applications:
  - inertial navigation,
  - electronic stability control (ESP).

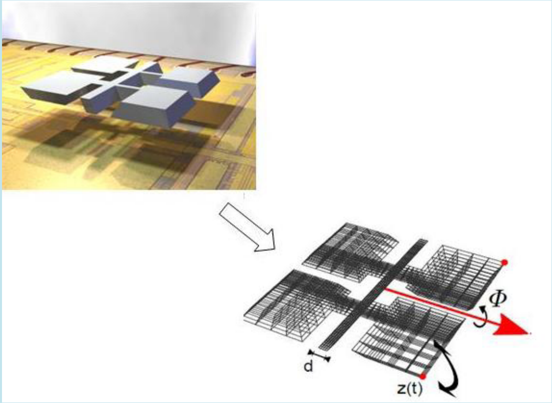


- Voltage applied to electrodes induces vibration of wings, resulting rotation due to Coriolis force yields sensor data.
- FE model of second order:
  $N = 17.361 \rightsquigarrow n = 34.722,\ m = 1,\ q = 12.$
- Sensor for position control based on acceleration and rotation.

Source: MOR Wiki: http://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Gyroscope

Microgyroscope (butterfly gyro)

Parametric FE model: $M(d)\ddot{x}(t) + D(\theta, d, \alpha, \beta)\dot{x}(t) + T(d)x(t) = Bu(t).$



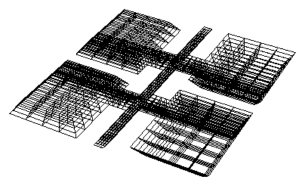
Microgyroscope (butterfly gyro)

Parametric FE model:

$$M(d)\ddot{x}(t) + D(\theta, d, \alpha, \beta)\dot{x}(t) + T(d)x(t) = Bu(t),$$

where

$$\begin{aligned} M(d) &= M_1 + dM_2, \\ D(\theta, d, \alpha, \beta) &= \theta(D_1 + dD_2) + \alpha M(d) + \beta T(d), \\ T(d) &= T_1 + \frac{1}{d} T_2 + dT_3, \end{aligned}$$

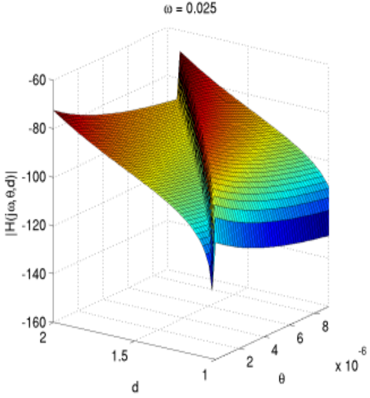


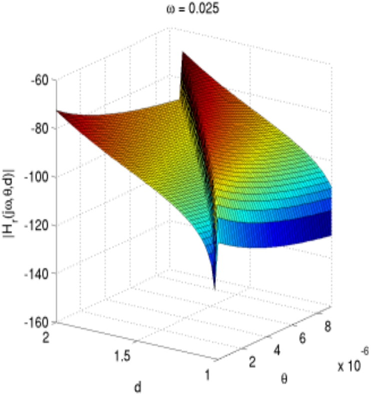
with the parameters

- width of bearing: d,
- angular velocity: θ,
- Rayleigh damping parameters: α, β.

Microgyroscope (butterfly gyro)

Original. . . and reduced-order model.





Parametric System

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) & = & A(p)x(t;p) + B(p)u(t), \\ y(t;p) & = & C(p)x(t;p). \end{cases}$$

Parametric System

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) & = & A(p)x(t;p) + B(p)u(t), \\ y(t;p) & = & C(p)x(t;p). \end{cases}$$

Desirable: parameter-affine representation:

$$\begin{array}{rcl} E(p) & = & E_0 + e_1(p)E_1 + \ldots + e_{q_E}(p)E_{q_E}, \\ A(p) & = & A_0 + a_1(p)A_1 + \ldots + a_{q_A}(p)A_{q_A}, \\ B(p) & = & B_0 + b_1(p)B_1 + \ldots + b_{q_B}(p)B_{q_B}, \\ C(p) & = & C_0 + c_1(p)C_1 + \ldots + c_{q_C}(p)C_{q_C}, \end{array}$$

allows easy parameter preservation for projection based model reduction.

Parametric Modeling

Parametric System

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \\ y(t;p) &= C(p)x(t;p). \end{cases}$$

Desirable: parameter-affine representation:

$$A(p) = A_0 + a_1(p)A_1 + \ldots + a_{q_A}(p)A_{q_A}, \quad \ldots$$

allows easy parameter preservation for projection based model reduction.

W.l.o.g. may assume this affine representation:

- Any system can be written in this affine form for some $q_M \leq n^2$, but for efficiency, need $q_M \ll n!$ $(M \in \{E, A, B, C\})$

Parametric Modeling

Parametric System

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) & = & A(p)x(t;p) + B(p)u(t), \\ y(t;p) & = & C(p)x(t;p). \end{cases}$$

Desirable: parameter-affine representation:

$$A(p) \quad = \quad A_0 + a_1(p)A_1 + \ldots + a_{q_A}(p)A_{q_A},$$

allows easy parameter preservation for projection based model reduction.

W.l.o.g. may assume this affine representation:

- Any system can be written in this affine form for some $q_M \leq n^2$, but for efficiency, need $q_M \ll n!$ $(M \in \{E, A, B, C\})$
- Empirical (operator) interpolation [Barrault/Maday/Nguyen/Patera 2004] yields this structure for "smooth enough" parameter dependencies; variants (Q)DEIM et al.

Parametric Modeling

Parametric model often not easily available when using commercial (FE) software.

Parametric Modeling

Parametric model often not easily available when using commercial (FE) software.

Exporting discrete system, i.e., matrices for linear systems, for given parameter configuration often possible.

Parametric Modeling

Parametric model often not easily available when using commercial (FE) software.

Exporting discrete system, i.e., matrices for linear systems, for given parameter configuration often possible.

Obtaining a linear parametrization

Assumption: the parametric matrix $M(p)$ ($M \in \{A, B, C, E\}$) can be parameterized linearly:

$$M(p) = p_1 M_1 + \ldots + p_d M_d, \qquad p \in \Omega \subset \mathbb{R}^d.$$

Parametric Modeling

Parametric model often not easily available when using commercial (FE) software.

Exporting discrete system, i.e., matrices for linear systems, for given parameter configuration often possible.

Obtaining a linear parametrization

Assumption: the parametric matrix $M(p)$ ($M \in \{A, B, C, E\}$) can be parameterized linearly:

$$M(p) = p_1 M_1 + \ldots + p_d M_d, \qquad p \in \Omega \subset \mathbb{R}^d.$$

Given: $M(p^{(j)})$ for $j = 1, \ldots, d$. Then

$$p_1^{(j)} M_1 + p_2^{(j)} M_2 + \quad \ldots \quad + p_d^{(j)} M_d = M(p^{(j)}), \qquad j = 1, \ldots, d,$$

is a linear system in the unknown matrices M_j, $j = 1, \ldots, d$.

Parametric Modeling

Parametric model often not easily available when using commercial (FE) software.

Exporting discrete system, i.e., matrices for linear systems, for given parameter configuration often possible.

Obtaining a linear parametrization

Assumption: the parametric matrix $M(p)$ ($M \in \{A, B, C, E\}$) can be parameterized linearly:

$$M(p) = p_1 M_1 + \ldots + p_d M_d, \qquad p \in \Omega \subset \mathbb{R}^d.$$

Given: $M(p^{(j)})$ for $j = 1, \ldots, d$. Then

$$p_1^{(j)} M_1 \; + \; p_2^{(j)} M_2 \; + \quad \ldots \quad + \, p_d^{(j)} M_d \quad = \quad M(p^{(j)}), \qquad j = 1, \ldots, d,$$

is a linear system in the unknown matrices M_j, $j = 1, \ldots, d$.

Choice $p^{(j)} = e_j \in \mathbb{R}^d \rightsquigarrow M_j = M(p^{(j)})$, $j = 1, \ldots, d$.

But e_j may not be feasible, i.e., $e_j \notin \Omega$, so that different physics would be modeled!

Parametric Modeling

Parametric model often not easily available when using commercial (FE) software.

Exporting discrete system, i.e., matrices for linear systems, for given parameter configuration often possible.

Obtaining a linear parametrization

Assumption: the parametric matrix $M(p)$ ($M \in \{A, B, C, E\}$) can be parameterized linearly:

$$M(p) = p_1 M_1 + \ldots + p_d M_d, \qquad p \in \Omega \subset \mathbb{R}^d.$$

Given: $M(p^{(j)})$ for $j = 1, \ldots, d$. Then

$$p_1^{(j)} M_1 \ + \ p_2^{(j)} M_2 \ + \quad \ldots \quad + p_d^{(j)} M_d \quad = \quad M(p^{(j)}), \qquad j = 1, \ldots, d,$$

is a linear system in the unknown matrices M_j, $j = 1, \ldots, d$.

With $P := \left[p_i^{(j)} \right]_{i,j=1}^{d} \in \mathbb{R}^{d \times d}$, rewrite this as $\quad (P \otimes I) \begin{bmatrix} M_1 \\ \vdots \\ M_d \end{bmatrix} = \begin{bmatrix} M(p^{(1)}) \\ \vdots \\ M(p^{(d)}) \end{bmatrix}.$

Parametric model often not easily available when using commercial (FE) software.

Exporting discrete system, i.e., matrices for linear systems, for given parameter configuration often possible.

Obtaining a linear parametrization

Assumption: the parametric matrix $M(p)$ ($M \in \{A, B, C, E\}$) can be parameterized linearly:

$$M(p) = p_1 M_1 + \ldots + p_d M_d, \qquad p \in \Omega \subset \mathbb{R}^d.$$

With $P := \left[p_i^{(j)} \right]_{i,j=1}^{d} \in \mathbb{R}^{d \times d}$, rewrite this as $(P \otimes I) \begin{bmatrix} M_1 \\ \vdots \\ M_d \end{bmatrix} = \begin{bmatrix} M(p^{(1)}) \\ \vdots \\ M(p^{(d)}) \end{bmatrix}$.

Now obtain unknown matrices as

$$\begin{bmatrix} M_1 \\ \vdots \\ M_d \end{bmatrix} = (P^{-1} \otimes I) \begin{bmatrix} M(p^{(1)}) \\ \vdots \\ M(p^{(d)}) \end{bmatrix}.$$

Obtaining a linear parametrization — Example

Assumption: the parametric matrix $M(p)$ ($M \in \{A, B, C, E\}$) can be parameterized linearly:

$$M(p) = p_1 M_1 + \ldots + p_d M_d, \qquad p \in \Omega \subset \mathbb{R}^d.$$

Parabolic system — linear diffusion-convection-reaction equation:

$$\partial_t \mathrm{x} - \kappa \Delta \mathrm{x} + [\,\nu_1,\, \nu_2\,]\ .\nabla \mathrm{x} + \mu \mathrm{x} = \mathrm{b}(\xi) u(t), \quad \xi \in \mathcal{D} \subset \mathbb{R}^2,\ t > 0$$

with initial/boundary conditions

$$\begin{aligned} \alpha(\xi)\mathrm{x} + \beta(\xi)\partial_\eta \mathrm{x} &= 0, & \xi \in \partial\mathcal{D}, \quad t \in [0, T], \\ \mathrm{x}(0, \xi) &= \mathrm{x}_0(\xi), & \xi \in \mathcal{D}, \end{aligned}$$

Obtaining a linear parametrization — Example

Assumption: the parametric matrix $M(p)$ ($M \in \{A, B, C, E\}$) can be parameterized linearly:

$$M(p) = p_1 M_1 + \ldots + p_d M_d, \qquad p \in \Omega \subset \mathbb{R}^d.$$

Parabolic system — linear diffusion-convection-reaction equation:

$$\partial_t \mathrm{x} - \kappa \Delta \mathrm{x} + [\,\nu_1,\, \nu_2\,]\, .\nabla \mathrm{x} + \mu \mathrm{x} = \mathrm{b}(\xi) u(t), \quad \xi \in \mathcal{D} \subset \mathbb{R}^2,\ t > 0$$

with initial/boundary conditions

$$\begin{aligned} \alpha(\xi)\mathrm{x} + \beta(\xi)\partial_\eta \mathrm{x} \quad &= \quad 0, \qquad \xi \in \partial\mathcal{D}, \quad t \in [0, T], \\ \mathrm{x}(0, \xi) \quad &= \quad \mathrm{x}_0(\xi), \quad \xi \in \mathcal{D}, \end{aligned}$$

FEM/FVM/FDM $\rightsquigarrow$

$$E\dot{x}(t) \;=\; -\left(\kappa K + \nu_1 C_x + \nu_2 C_y + \mu R\right) x(t) + Bu(t),$$

Obtaining a linear parametrization — Example

Assumption: the parametric matrix $M(p)$ ($M \in \{A, B, C, E\}$) can be parameterized linearly:

$$M(p) = p_1 M_1 + \ldots + p_d M_d, \qquad p \in \Omega \subset \mathbb{R}^d.$$

Parabolic system — linear diffusion-convection-reaction equation:

$$\partial_t \mathrm{x} - \kappa \Delta \mathrm{x} + [\, \nu_1,\, \nu_2\,] \,.\nabla \mathrm{x} + \mu \mathrm{x} = \mathrm{b}(\xi) u(t), \quad \xi \in \mathcal{D} \subset \mathbb{R}^2,\ t > 0$$

with initial/boundary conditions

$$\begin{aligned} \alpha(\xi)\mathrm{x} + \beta(\xi)\partial_\eta \mathrm{x} &= 0, && \xi \in \partial\mathcal{D}, \quad t \in [0, T], \\ \mathrm{x}(0, \xi) &= \mathrm{x}_0(\xi), && \xi \in \mathcal{D}, \end{aligned}$$

FEM/FVM/FDM $\rightsquigarrow$

$$E\dot{x}(t) \;=\; -\left(\kappa K + \nu_1 C_x + \nu_2 C_y + \mu R\right) x(t) + Bu(t),$$

Hence, we obtain a linear parametrization, only the "A" matrix is parameterized here:

$$\begin{aligned} A(p) &= p_1 A_1 + p_2 A_2 + p_3 A_3 + p_4 A_4, \qquad \text{with} \\ p &= [\kappa, \nu_1, \nu_2, \mu] \quad \text{and} \quad A_1 = -K,\ A_2 = -C_x,\ A_3 = -C_y,\ A_4 = -R. \end{aligned}$$

Note: selecting, e.g., $p^{(1)} = e_1 \rightsquigarrow$ heat equation w/o convection/reaction $\rightsquigarrow$ different physics!

Obtaining an affine parametrization using interpolation

If $M(p^{(j)})$ (for $j = 1, \ldots, \ell$ and $M \in \{A, B, C, E\}$) are available (e.g., can be exported from FEM software), and parametrization is unknown anyway, using standard interpolation techniques to build parametric model:

Set $M_j := M(p^{(j)})$ and build for $M \in \{A, B, C, E\}$

$$M(p) = \phi_1(p)M_1 + \phi_2(p)M_2 + \ldots + \phi_\ell(p)M_\ell, \qquad p \in \Omega \subset \mathbb{R}^d,$$

where $\phi_k(p^{(j)}) = \delta_{kj}$, via

- polynomial interpolation (e.g., barycentric Lagrangian interpolation),
- rational (Loewner) interpolation,
- sparse grid interpolation, ...

The Parametric Model Order Reduction (PMOR) Problem

Problem

Approximate the dynamical system

$$\begin{array}{rcll} E(p)\dot{x} &=& A(p)x + B(p)u, & E(p),\ A(p) \in \mathbb{R}^{n\times n}, \\ y &=& C(p)x, & B(p) \in \mathbb{R}^{n\times m},\ C(p) \in \mathbb{R}^{q\times n}, \end{array}$$

by reduced-order system

$$\begin{array}{rcll} \hat{E}(p)\dot{\hat{x}} &=& \hat{A}(p)\hat{x} + \hat{B}(p)u, & \hat{E}(p),\ \hat{A}(p) \in \mathbb{R}^{r\times r}, \\ \hat{y} &=& \hat{C}(p)\hat{x}, & \hat{B}(p) \in \mathbb{R}^{r\times m},\ \hat{C}(p) \in \mathbb{R}^{q\times r}, \end{array}$$

of order $r \ll n$, such that

$$\|y - \hat{y}\| = \left\| Gu - \hat{G}u \right\| \leq \left\| G - \hat{G} \right\| \cdot \|u\| < \text{tolerance} \cdot \|u\| \quad \forall\ p \in \Omega.$$

The Parametric Model Order Reduction (PMOR) Problem

Problem

Approximate the dynamical system

$$\begin{array}{rcll} E(p)\dot{x} & = & A(p)x + B(p)u, & E(p),\ A(p) \in \mathbb{R}^{n\times n}, \\ y & = & C(p)x, & B(p) \in \mathbb{R}^{n\times m},\ C(p) \in \mathbb{R}^{q\times n}, \end{array}$$

by reduced-order system

$$\begin{array}{rcll} \hat{E}(p)\dot{\hat{x}} & = & \hat{A}(p)\hat{x} + \hat{B}(p)u, & \hat{E}(p),\ \hat{A}(p) \in \mathbb{R}^{r\times r}, \\ \hat{y} & = & \hat{C}(p)\hat{x}, & \hat{B}(p) \in \mathbb{R}^{r\times m},\ \hat{C}(p) \in \mathbb{R}^{q\times r}, \end{array}$$

of order $r \ll n$, such that

$$\|y - \hat{y}\| = \left\|Gu - \hat{G}u\right\| \leq \left\|G - \hat{G}\right\| \cdot \|u\| < \text{tolerance} \cdot \|u\| \quad \forall\ p \in \Omega.$$

$\Longrightarrow$ Approximation problem: $\displaystyle\min_{\text{order}\,(\hat{G})\leq r} \left\|G - \hat{G}\right\|.$

PMOR $\longleftrightarrow$ Multivariate Function Approximation

- Approximate (for fast evaluation) function G, defined on $\mathbb{C} \times \Omega$.

PMOR $\longleftrightarrow$ Multivariate Function Approximation

- Approximate (for fast evaluation) function G, defined on $\mathbb{C} \times \Omega$.
- But:

$$\begin{aligned} G : \mathbb{C} \times \Omega &\to \mathbb{C}^{q \times m}, \quad \Omega = [\alpha_1, \beta_1] \times \ldots \times [\alpha_d, \beta_d], \\ G(s; p_1, \ldots, p_d) &\in \mathbb{C}^{q \times m}. \end{aligned}$$

$\leadsto$ Variables s and p_j have different "meaning" for G.
Dynamical system is in the background!

$\leadsto$ Matrix-valued function, require matrix- not entry-wise approximation!

PMOR $\longleftrightarrow$ Multivariate Function Approximation

- Approximate (for fast evaluation) function G, defined on $\mathbb{C} \times \Omega$.
- But:

$$\begin{aligned} G : \mathbb{C} \times \Omega \quad &\rightarrow \quad \mathbb{C}^{q \times m}, \quad \Omega = [\alpha_1, \beta_1] \times \ldots \times [\alpha_d, \beta_d], \\ G(s; p_1, \ldots, p_d) \quad &\in \quad \mathbb{C}^{q \times m}. \end{aligned}$$

$\rightsquigarrow$ Variables s and p_j have different "meaning" for G.
Dynamical system is in the background!

$\rightsquigarrow$ Matrix-valued function, require matrix- not entry-wise approximation!

- G is rational in s, $n \sim$ degree of denominator polynomial.
  $\rightsquigarrow$ Require approximation to be rational in s.

PMOR $\longleftrightarrow$ Multivariate Function Approximation

- Approximate (for fast evaluation) function G, defined on $\mathbb{C} \times \Omega$.
- But:

$$\begin{aligned} G : \mathbb{C} \times \Omega \quad &\to \quad \mathbb{C}^{q \times m}, \quad \Omega = [\alpha_1, \beta_1] \times \ldots \times [\alpha_d, \beta_d], \\ G(s; p_1, \ldots, p_d) \quad &\in \quad \mathbb{C}^{q \times m}. \end{aligned}$$

$\rightsquigarrow$ Variables s and p_j have different "meaning" for G.
Dynamical system is in the background!

$\rightsquigarrow$ Matrix-valued function, require matrix- not entry-wise approximation!

- G is rational in s, $n \sim$ degree of denominator polynomial.
  $\rightsquigarrow$ Require approximation to be rational in s.
- Require structure-preserving approximation, e.g., for control design.
  $\rightsquigarrow$ Need realization as linear parametric system!

PMOR $\longleftrightarrow$ Multivariate Function Approximation

- Approximate (for fast evaluation) function G, defined on $\mathbb{C} \times \Omega$.
- But:

$$\begin{aligned} G : \mathbb{C} \times \Omega &\rightarrow \mathbb{C}^{q \times m}, \quad \Omega = [\alpha_1, \beta_1] \times \ldots \times [\alpha_d, \beta_d], \\ G(s; p_1, \ldots, p_d) &\in \mathbb{C}^{q \times m}. \end{aligned}$$

$\rightsquigarrow$ Variables s and p_j have different "meaning" for G.
Dynamical system is in the background!

$\rightsquigarrow$ Matrix-valued function, require matrix- not entry-wise approximation!

- G is rational in s, $n \sim$ degree of denominator polynomial.
  $\rightsquigarrow$ Require approximation to be rational in s.
- Require structure-preserving approximation, e.g., for control design.
  $\rightsquigarrow$ Need realization as linear parametric system!
- Also would like to be able to reproduce system dynamics (stability, passivity).

PMOR $\longleftrightarrow$ Multivariate Function Approximation

- Approximate (for fast evaluation) function G, defined on $\mathbb{C} \times \Omega$.
- But:

$$\begin{aligned} G : \mathbb{C} \times \Omega &\to \mathbb{C}^{q \times m}, \quad \Omega = [\alpha_1, \beta_1] \times \ldots \times [\alpha_d, \beta_d], \\ G(s; p_1, \ldots, p_d) &\in \mathbb{C}^{q \times m}. \end{aligned}$$

$\rightsquigarrow$ Variables s and p_j have different "meaning" for G.
Dynamical system is in the background!

$\rightsquigarrow$ Matrix-valued function, require matrix- not entry-wise approximation!

- G is rational in s, $n \sim$ degree of denominator polynomial.
  $\rightsquigarrow$ Require approximation to be rational in s.
- Require structure-preserving approximation, e.g., for control design.
  $\rightsquigarrow$ Need realization as linear parametric system!
- Also would like to be able to reproduce system dynamics (stability, passivity).

Hence, try to preserve the structure of the original model!

Parametric System

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \\ y(t;p) &= C(p)x(t;p). \end{cases}$$

Parametric System

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \\ y(t;p) &= C(p)x(t;p). \end{cases}$$

Parametric model reduction goal:

preserve parameters as *symbolic quantities* in reduced-order model:

$$\widehat{\Sigma}(p) : \begin{cases} \hat{E}(p)\dot{\hat{x}}(t;p) &= \hat{A}(p)\hat{x}(t;p) + \hat{B}(p)u(t), \\ \hat{y}(t;p) &= \hat{C}(p)\hat{x}(t;p) \end{cases}$$

with states $\hat{x}(t;p) \in \mathbb{R}^r$ and $r \ll n$.

Parametric System

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \\ y(t;p) &= C(p)x(t;p). \end{cases}$$

Assuming parameter-affine representation:

$$\begin{aligned} E(p) &= E_0 + e_1(p)E_1 + \ldots + e_{q_E}(p)E_{q_E}, \\ A(p) &= A_0 + a_1(p)A_1 + \ldots + a_{q_A}(p)A_{q_A}, \\ B(p) &= B_0 + b_1(p)B_1 + \ldots + b_{q_B}(p)B_{q_B}, \\ C(p) &= C_0 + c_1(p)C_1 + \ldots + c_{q_C}(p)C_{q_C}, \end{aligned}$$

allows easy parameter preservation for projection based model reduction.

Petrov-Galerkin-type projection

For given projection matrices $V, W \in \mathbb{R}^{n \times r}$ with $W^T V = I_r$
($\rightsquigarrow (VW^T)^2 = VW^T$ is projector), compute

$$
\begin{aligned}
\hat{E}(p) &= W^T E_0 V + e_1(p) W^T E_1 V + \ldots + e_{q_E}(p) W^T E_{q_E} V \\
\hat{A}(p) &= W^T A_0 V + a_1(p) W^T A_1 V + \ldots + a_{q_A}(p) W^T A_{q_A} V \\
\hat{B}(p) &= W^T B_0 + b_1(p) W^T B_1 + \ldots + b_{q_B}(p) W^T B_{q_B} \\
\hat{C}(p) &= C_0 V + c_1(p) C_1 V + \ldots + c_{q_C}(p) C_{q_C} V
\end{aligned}
$$

Petrov-Galerkin-type projection

For given projection matrices $V, W \in \mathbb{R}^{n\times r}$ with $W^T V = I_r$
($\rightsquigarrow (VW^T)^2 = VW^T$ is projector), compute

$$
\begin{aligned}
\hat{E}(p) &= W^T E_0 V + e_1(p) W^T E_1 V + \ldots + e_{q_E}(p) W^T E_{q_E} V \\
&= \hat{E}_0 + e_1(p)\hat{E}_1 + \ldots + e_{q_E}(p)\hat{E}_{q_E} \\
\hat{A}(p) &= W^T A_0 V + a_1(p) W^T A_1 V + \ldots + a_{q_A}(p) W^T A_{q_A} V \\
&= \hat{A}_0 + a_1(p)\hat{A}_1 + \ldots + a_{q_A}(p)\hat{A}_{q_A} \\
\hat{B}(p) &= W^T B_0 + b_1(p) W^T B_1 + \ldots + b_{q_B}(p) W^T B_{q_B} \\
&= \hat{B}_0 + b_1(p)\hat{B}_1 + \ldots + b_{q_B}(p)\hat{B}_{q_B} \\
\hat{C}(p) &= C_0 V + c_1(p) C_1 V + \ldots + c_{q_C}(p) C_{q_C} V \\
&= \hat{C}_0 + c_1(p)\hat{C}_1 + \ldots + c_{q_C}(p)\hat{C}_{q_C}
\end{aligned}
$$

Local Bases

Obtain $V_k, W_k \in \mathbb{R}^{n \times r_k}$ using any non-parametric linear MOR method for a number of full-order models $\Sigma(p^{(k)})$, $k = 1, \ldots, \ell$. Then compute reduced-order model by

Local Bases

Obtain $V_k, W_k \in \mathbb{R}^{n \times r_k}$ using any non-parametric linear MOR method for a number of full-order models $\Sigma(p^{(k)})$, $k = 1, \ldots, \ell$. Then compute reduced-order model by

1. **manifold interpolation** [Amsallam/Farhat 2008, Brunsch 2017]

Local Bases

Obtain $V_k, W_k \in \mathbb{R}^{n \times r_k}$ using any non-parametric linear MOR method for a number of full-order models $\Sigma(p^{(k)})$, $k = 1, \ldots, \ell$. Then compute reduced-order model by

1. **manifold interpolation** [Amsallam/Farhat 2008, Brunsch 2017]
2. **transfer function interpolation** (= interpolate $y(s, .)$ in frequency domain) [B./Baur 2008/09]

Local Bases

Obtain $V_k, W_k \in \mathbb{R}^{n \times r_k}$ using any non-parametric linear MOR method for a number of full-order models $\Sigma(p^{(k)})$, $k = 1, \ldots, \ell$. Then compute reduced-order model by

1. **manifold interpolation** [Amsallam/Farhat 2008, Brunsch 2017]
2. **transfer function interpolation** (= interpolate $y(s, .)$ in frequency domain) [B./Baur 2008/09]
3. **matrix interpolation** [Panzer/Mohring/Eid/Lohmann 2010, Amsallam/Farhat 2011]

Local Bases

Obtain $V_k, W_k \in \mathbb{R}^{n \times r_k}$ using any non-parametric linear MOR method for a number of full-order models $\Sigma(p^{(k)})$, $k = 1, \ldots, \ell$. Then compute reduced-order model by

1. **manifold interpolation** [Amsallam/Farhat 2008, Brunsch 2017]
2. **transfer function interpolation** (= interpolate $y(s, .)$ in frequency domain) [B./Baur 2008/09]
3. **matrix interpolation** [Panzer/Mohring/Eid/Lohmann 2010, Amsallam/Farhat 2011]

Advantage:

no need for affine parametrization, requires only system matrices $A(p^{(k)}), B(p^{(k)}), \ldots$.

Disadvantages:

Local Bases

Obtain $V_k, W_k \in \mathbb{R}^{n \times r_k}$ using any non-parametric linear MOR method for a number of full-order models $\Sigma(p^{(k)})$, $k = 1, \ldots, \ell$. Then compute reduced-order model by

1. **manifold interpolation** [Amsallam/Farhat 2008, Brunsch 2017]
2. **transfer function interpolation** (= interpolate $y(s,.)$ in frequency domain) [B./Baur 2008/09]
3. **matrix interpolation** [Panzer/Mohring/Eid/Lohmann 2010, Amsallam/Farhat 2011]

Advantage:
no need for affine parametrization, requires only system matrices $A(p^{(k)}), B(p^{(k)}), \ldots$.

Disadvantages:

1. **manifold interpolation:** originally, requires $\mathcal{O}(nr)$ operations in "online" phase. [Brunsch 2017, B./Cao/Schilders 2020] overcome this problem, but only for negative definite matrix pencils $(A(p), E(p))$. Some more recent work by [Jiang et al., Sato/Sato].

Local Bases

Obtain $V_k, W_k \in \mathbb{R}^{n \times r_k}$ using any non-parametric linear MOR method for a number of full-order models $\Sigma(p^{(k)})$, $k = 1, \ldots, \ell$. Then compute reduced-order model by

1. **manifold interpolation** [Amsallam/Farhat 2008, Brunsch 2017]
2. **transfer function interpolation** (= interpolate $y(s, .)$ in frequency domain) [B./Baur 2008/09]
3. **matrix interpolation** [Panzer/Mohring/Eid/Lohmann 2010, Amsallam/Farhat 2011]

Advantage:

no need for affine parametrization, requires only system matrices $A(p^{(k)}), B(p^{(k)}), \ldots$.

Disadvantages:

1. **manifold interpolation:** originally, requires $\mathcal{O}(nr)$ operations in "online" phase. [Brunsch 2017, B./Cao/Schilders 2020] overcome this problem, but only for negative definite matrix pencils $(A(p), E(p))$. Some more recent work by [Jiang et al., Sato/Sato].
2. **transfer function interpolation:** spurious poles of the parametric transfer function.

Local Bases

Obtain $V_k, W_k \in \mathbb{R}^{n\times r_k}$ using any non-parametric linear MOR method for a number of full-order models $\Sigma(p^{(k)})$, $k = 1, \ldots, \ell$. Then compute reduced-order model by

1. **manifold interpolation** [Amsallam/Farhat 2008, Brunsch 2017]
2. **transfer function interpolation** (= interpolate $y(s, .)$ in frequency domain) [B./Baur 2008/09]
3. **matrix interpolation** [Panzer/Mohring/Eid/Lohmann 2010, Amsallam/Farhat 2011]

Advantage:

no need for affine parametrization, requires only system matrices $A(p^{(k)}), B(p^{(k)}), \ldots$.

Disadvantages:

1. **manifold interpolation:** originally, requires $\mathcal{O}(nr)$ operations in "online" phase. [Brunsch 2017, B./Cao/Schilders 2020] overcome this problem, but only for negative definite matrix pencils $(A(p), E(p))$. Some more recent work by [Jiang et al., Sato/Sato].
2. **transfer function interpolation:** spurious poles of the parametric transfer function.
3. **matrix interpolation:** different models obtained in different coordinate systems $\rightsquigarrow$ Procrustes problem $\rightsquigarrow$ potential loss of accuracy; efficiency in "online" phase suffers from evaluating the interpolation operator.

Global Basis

Obtain $V, W \in \mathbb{R}^{n \times r_k}$ such that $V^T W = I_r$ and perform structure preserving (Petrov-) Galerkin projection, exploiting affine parametrization of the linear parametric system.

Global Basis

Obtain $V, W \in \mathbb{R}^{n\times r_k}$ such that $V^T W = I_r$ and perform structure preserving (Petrov-) Galerkin projection, exploiting affine parametrization of the linear parametric system.

Obtain global basis from

1. concatenation of local basis matrices:

$$V := [\, V_1, \ldots, V_\ell \,], \qquad W := [\, W_1, \ldots, W_\ell \,]$$

   and orthogonalization (truncation), using, e.g., SVD;

Global Basis

Obtain $V, W \in \mathbb{R}^{n \times r_k}$ such that $V^T W = I_r$ and perform structure preserving (Petrov-) Galerkin projection, exploiting affine parametrization of the linear parametric system.

Obtain global basis from

1. concatenation of local basis matrices:

$$V := [\, V_1, \ldots, V_\ell \,], \qquad W := [\, W_1, \ldots, W_\ell \,]$$

   and orthogonalization (truncation), using, e.g., SVD;

2. bilinearization and using bilinear MOR techniques;

Global Basis

Obtain $V, W \in \mathbb{R}^{n \times r_k}$ such that $V^T W = I_r$ and perform structure preserving (Petrov-) Galerkin projection, exploiting affine parametrization of the linear parametric system.

Obtain global basis from

1. concatenation of local basis matrices:

$$V := [\, V_1, \ldots, V_\ell \,], \qquad W := [\, W_1, \ldots, W_\ell \,]$$

   and orthogonalization (truncation), using, e.g., SVD;
2. bilinearization and using bilinear MOR techniques;
3. parametric balanced truncation [Son/Stykel 2017].

Global Basis

Obtain $V, W \in \mathbb{R}^{n \times r_k}$ such that $V^T W = I_r$ and perform structure preserving (Petrov-) Galerkin projection, exploiting affine parametrization of the linear parametric system.

Obtain global basis from

1. concatenation of local basis matrices:

$$V := [\, V_1, \ldots, V_\ell \,], \qquad W := [\, W_1, \ldots, W_\ell \,]$$

   and orthogonalization (truncation), using, e.g., SVD;

2. bilinearization and using bilinear MOR techniques;
3. parametric balanced truncation [Son/Stykel 2017].

Avoids most of the problems encountered with local bases, but requires parameter-affine representation of system.

Using a Global Basis w/o Affine Parametrization

Empirical Matrix Interpolation Method [B./Gugercin/Willcox 2015]

Given $V, W \in \mathbb{R}^{n \times r}$ and suppose only that $M(p) \in \mathbb{R}^{n \times t}$ can be evaluated at specific parameter values.

Empirical Matrix Interpolation Method [B./Gugercin/Willcox 2015]

Given $V, W \in \mathbb{R}^{n\times r}$ and suppose only that $M(p) \in \mathbb{R}^{n\times t}$ can be evaluated at specific parameter values.

- Let $m(p) := \operatorname{vec}(M(p)) \in \mathbb{R}^{nt}$.

Empirical Matrix Interpolation Method

[B./Gugercin/Willcox 2015]

Given $V, W \in \mathbb{R}^{n\times r}$ and suppose only that $M(p) \in \mathbb{R}^{n\times t}$ can be evaluated at specific parameter values.

- Let $m(p) := \mathrm{vec}\,(M(p)) \in \mathbb{R}^{nt}$.
- Goal: approximate $m(p) \approx \tilde{m}(p) = \Psi\alpha(p)$, where $\Psi \in \mathbb{R}^{nt\times \ell}$ and $\alpha(p) \in \mathbb{R}^{\ell}$ with $\ell \ll n$.

Empirical Matrix Interpolation Method [B./Gugercin/Willcox 2015]

Given $V, W \in \mathbb{R}^{n\times r}$ and suppose only that $M(p) \in \mathbb{R}^{n\times t}$ can be evaluated at specific parameter values.

- Let $m(p) := \operatorname{vec}(M(p)) \in \mathbb{R}^{nt}$.
- Goal: approximate $m(p) \approx \tilde{m}(p) = \Psi\alpha(p)$, where $\Psi \in \mathbb{R}^{nt\times\ell}$ and $\alpha(p) \in \mathbb{R}^{\ell}$ with $\ell \ll n$.
- Then $\hat{m}(p) = \operatorname{vec}(\hat{M}(p)) \in \mathbb{R}^{rt}$ (or $\mathbb{R}^{r^2}$ if $t = n$) can be computed cheaply and independent of n as

$$
\begin{aligned}
\hat{m}(p) &= \operatorname{vec}\left(W^T M(p) V\right) \\
&= (V^T \otimes W^T) m(p) \approx (V^T \otimes W^T)\tilde{m}(p) = (V^T \otimes W^T)\Psi\alpha(p) = \tilde{\hat{m}}(p).
\end{aligned}
$$

Using a Global Basis w/o Affine Parametrization

Empirical Matrix Interpolation Method [B./Gugercin/Willcox 2015]

Given $V, W \in \mathbb{R}^{n \times r}$ and suppose only that $M(p) \in \mathbb{R}^{n \times t}$ can be evaluated at specific parameter values.

- Let $m(p) := \operatorname{vec}(M(p)) \in \mathbb{R}^{nt}$.
- Goal: approximate $m(p) \approx \tilde{m}(p) = \Psi\alpha(p)$, where $\Psi \in \mathbb{R}^{nt \times \ell}$ and $\alpha(p) \in \mathbb{R}^{\ell}$ with $\ell \ll n$.
- Then $\hat{m}(p) = \operatorname{vec}(\hat{M}(p)) \in \mathbb{R}^{rt}$ (or $\mathbb{R}^{r^2}$ if $t = n$) can be computed cheaply and independent of n as

$$\begin{aligned}\hat{m}(p) &= \operatorname{vec}\left(W^T M(p) V\right)\\ &= (V^T \otimes W^T)m(p) \approx (V^T \otimes W^T)\tilde{m}(p) = (V^T \otimes W^T)\Psi\alpha(p) = \tilde{\hat{m}}(p).\end{aligned}$$

- This is achieved by sampling $M(p)$ at $p = p^{(j)}$, $j = 1, \ldots, \ell$, yielding

$$\psi_j = \operatorname{vec}(M(p^{(j)})) \qquad \text{and} \qquad \Psi = [\psi_1, \ldots, \psi_\ell].$$

  Then apply (Q,D)EIM (or alike) to determine $\alpha(p)$ s.t. selected entries of $\tilde{m}(p)$ interpolate those entries of $m(p)$.

Using a Global Basis w/o Affine Parametrization

Empirical Matrix Interpolation Method [B./Gugercin/Willcox 2015]

Given $V, W \in \mathbb{R}^{n \times r}$ and suppose only that $M(p) \in \mathbb{R}^{n \times t}$ can be evaluated at specific parameter values.

Goal: approximate $m(p) \approx \tilde{m}(p) = \Psi\alpha(p)$, where $\Psi \in \mathbb{R}^{nt \times \ell}$ and $\alpha(p) \in \mathbb{R}^{\ell}$ with $\ell \ll n$, and where Ψ is the sampling matrix built by $\operatorname{vec}\left(M(p^{(j)})\right)$.

Using a Global Basis w/o Affine Parametrization

Empirical Matrix Interpolation Method

[B./Gugercin/Willcox 2015]

Given $V, W \in \mathbb{R}^{n\times r}$ and suppose only that $M(p) \in \mathbb{R}^{n\times t}$ can be evaluated at specific parameter values.

Goal: approximate $m(p) \approx \tilde{m}(p) = \Psi\alpha(p)$, where $\Psi \in \mathbb{R}^{nt\times\ell}$ and $\alpha(p) \in \mathbb{R}^{\ell}$ with $\ell \ll n$, and where Ψ is the sampling matrix built by $\operatorname{vec}(M(p^{(j)}))$.

- Apply (Q,D)EIM (or alike) to determine $\alpha(p)$ s.t. selected entries of $\tilde{m}(p)$ interpolate those entries of $m(p)$.

  Let $z_1, z_2, \ldots, z_\ell$ be the selected indices to be exactly matched, and $Z := [e_{z_1}, \ldots, e_{z_\ell}]$. Then, forcing interpolation at the selected rows implies

$$Z^T m(p) = Z^T \Psi\alpha(p) \quad \Longrightarrow \quad \alpha(p) = (Z^T\Psi)^{-1} Z^T m(p).$$

Empirical Matrix Interpolation Method [B./Gugercin/Willcox 2015]

Given $V, W \in \mathbb{R}^{n\times r}$ and suppose only that $M(p) \in \mathbb{R}^{n\times t}$ can be evaluated at specific parameter values.

Goal: approximate $m(p) \approx \tilde{m}(p) = \Psi\alpha(p)$, where $\Psi \in \mathbb{R}^{nt\times\ell}$ and $\alpha(p) \in \mathbb{R}^{\ell}$ with $\ell \ll n$, and where Ψ is the sampling matrix built by $\operatorname{vec}\left(M(p^{(j)})\right)$.

- Apply (Q,D)EIM (or alike) to determine $\alpha(p)$ s.t. selected entries of $\tilde{m}(p)$ interpolate those entries of $m(p)$.

  Let $z_1, z_2, \ldots, z_\ell$ be the selected indices to be exactly matched, and $Z := [e_{z_1}, \ldots, e_{z_\ell}]$. Then, forcing interpolation at the selected rows implies

  $$Z^T m(p) = Z^T \Psi\alpha(p) \quad \Longrightarrow \quad \alpha(p) = (Z^T\Psi)^{-1} Z^T m(p).$$

- Hence, the approximation is given by $\tilde{m}(p) = \Psi(Z^T\Psi)^{-1}Z^T m(p)$.

Using a Global Basis w/o Affine Parametrization

Empirical Matrix Interpolation Method [B./Gugercin/Willcox 2015]

Given $V, W \in \mathbb{R}^{n\times r}$ and suppose only that $M(p) \in \mathbb{R}^{n\times t}$ can be evaluated at specific parameter values.

Goal: approximate $m(p) \approx \tilde{m}(p) = \Psi\alpha(p)$, where $\Psi \in \mathbb{R}^{nt\times\ell}$ and $\alpha(p) \in \mathbb{R}^{\ell}$ with $\ell \ll n$, and where Ψ is the sampling matrix built by $\operatorname{vec}(M(p^{(j)}))$.

- Apply (Q,D)EIM (or alike) to determine $\alpha(p)$ s.t. selected entries of $\tilde{m}(p)$ interpolate those entries of $m(p)$.

  Let $z_1, z_2, \ldots, z_\ell$ be the selected indices to be exactly matched, and $Z := [e_{z_1}, \ldots, e_{z_\ell}]$. Then, forcing interpolation at the selected rows implies

$$Z^T m(p) = Z^T \Psi\alpha(p) \quad \Longrightarrow \quad \alpha(p) = (Z^T\Psi)^{-1} Z^T m(p).$$

- Hence, the approximation is given by $\tilde{m}(p) = \Psi(Z^T\Psi)^{-1}Z^T m(p)$.
- Undoing the vectorization yields the reduced model matrix

$$\hat{M}(p) := \operatorname{vec}^{-1}\left(\tilde{\hat{m}}(p)\right) = \operatorname{vec}^{-1}\left((V^T \otimes W^T)\Psi\alpha(p)\right) = \sum_{j=1}^{\ell} \alpha_j(p)\,\underbrace{W^T M(p^{(j)}) V}_{\text{precomputable!}}$$

Using a Global Basis w/o Affine Parametrization

Empirical Matrix Interpolation Method [B./Gugercin/Willcox 2015]

Given $V, W \in \mathbb{R}^{n \times r}$ and suppose only that $M(p) \in \mathbb{R}^{n \times t}$ can be evaluated at specific parameter values.

Goal: approximate $m(p) \approx \tilde{m}(p) = \Psi\alpha(p)$, where $\Psi \in \mathbb{R}^{nt \times \ell}$ and $\alpha(p) \in \mathbb{R}^{\ell}$ with $\ell \ll n$, and where Ψ is the sampling matrix built by $\operatorname{vec}\left(M(p^{(j)})\right)$.

- Apply (Q,D)EIM (or alike) to determine $\alpha(p)$ s.t. selected entries of $\tilde{m}(p)$ interpolate those entries of $m(p)$.

  Let $z_1, z_2, \ldots, z_\ell$ be the selected indices to be exactly matched, and $Z := [e_{z_1}, \ldots, e_{z_\ell}]$. Then, forcing interpolation at the selected rows implies

  $$Z^T m(p) = Z^T \Psi\alpha(p) \quad \Longrightarrow \quad \alpha(p) = (Z^T\Psi)^{-1} Z^T m(p).$$

- Hence, the approximation is given by $\tilde{m}(p) = \Psi(Z^T\Psi)^{-1}Z^T m(p)$.
- Undoing the vectorization yields the reduced model matrix

$$\hat{M}(p) := \operatorname{vec}^{-1}\Big((V^T \otimes W^T)\Psi\alpha(p)\Big) = \sum_{j=1}^{\ell} \alpha_j(p) \underbrace{W^T M(p^{(j)}) V}_{\text{precomputable!}} =: \sum_{j=1}^{\ell} \alpha_j(p) \hat{M}_j.$$

Error Measures

Parametric Systems Norms

Mean-square error norm:

$$\|G - \hat{G}\|^2_{\mathcal{H}_2 \times L_2(\Omega)} := \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{\Omega} \left\| G(\jmath\omega, p) - \hat{G}(\jmath\omega, p) \right\|_F^2 \mathrm{d}p_1 \ldots \mathrm{d}p_d \, \mathrm{d}\omega,$$

where $\| \, . \, \|_F$ denotes the Frobenius norm.

Worst-case error norm:

$$\|G - \hat{G}\|_{\mathcal{H}_\infty \times L_\infty(\Omega)} := \sup_{\omega \in \mathbb{R}, \; p \in \Omega} \left\| G(\jmath\omega, p) - \hat{G}(\jmath\omega, p) \right\|_2 .$$

Outline

Recall Interpolatory Model Reduction

Computation of reduced-order model by projection

Given a linear (descriptor) system $E\dot{x} = Ax + Bu,\ y = Cx$ with transfer function $G(s) = C(sE - A)^{-1}B,$ a reduced-order model is obtained using truncation matrices $V, W \in \mathbb{R}^{n\times r}$ with $W^TV = I_r$ ($\rightsquigarrow$ $(VW^T)^2 = VW^T$ is projector) by computing

$$\hat{E} = W^TEV,\ \hat{A} = W^TAV,\ \hat{B} = W^TB,\ \hat{C} = CV.$$

Petrov-Galerkin-type (two-sided) projection: $W \neq V$,

Galerkin-type (one-sided) projection: $W = V$.

Recall Interpolatory Model Reduction

Computation of reduced-order model by projection

Given a linear (descriptor) system $E\dot{x} = Ax + Bu,\ y = Cx$ with transfer function $G(s) = C(sE - A)^{-1}B$, a reduced-order model is obtained using truncation matrices $V, W \in \mathbb{R}^{n\times r}$ with $W^TV = I_r$ ($\rightsquigarrow$ $(VW^T)^2 = VW^T$ is projector) by computing

$$\hat{E} = W^TEV,\ \hat{A} = W^TAV,\ \hat{B} = W^TB,\ \hat{C} = CV.$$

Petrov-Galerkin-type (two-sided) projection: $W \neq V$,

Galerkin-type (one-sided) projection: $W = V$.

Rational Interpolation/Moment-Matching

Choose V, W such that

$$G(s_j) = \hat{G}(s_j), \quad j = 1, \ldots, k,$$

and

$$\frac{d^i}{ds^i}G(s_j) = \frac{d^i}{ds^i}\hat{G}(s_j), \quad i = 1, \ldots, K_j, \quad j = 1, \ldots, k.$$

Recall Interpolatory Model Reduction

Theorem (simplified) [Grimme 1997, Villemagne/Skelton 1987]

If

$$\begin{aligned} \mathrm{span}\left\{(s_1 E - A)^{-1}B, \ldots, (s_k E - A)^{-1}B\right\} &\subset \mathrm{range}(V), \\ \mathrm{span}\left\{(s_1 E - A)^{-T}C^T, \ldots, (s_k E - A)^{-T}C^T\right\} &\subset \mathrm{range}(W), \end{aligned}$$

then

$$G(s_j) = \hat{G}(s_j), \quad \frac{d}{ds}G(s_j) = \frac{d}{ds}\hat{G}(s_j), \quad \text{for } j = 1, \ldots, k.$$

Recall Interpolatory Model Reduction

Theorem (simplified) [Grimme 1997, Villemagne/Skelton 1987]

If

$$\begin{aligned} \mathrm{span}\left\{(s_1E-A)^{-1}B,\ldots,(s_kE-A)^{-1}B\right\} &\subset \mathrm{range}(V), \\ \mathrm{span}\left\{(s_1E-A)^{-T}C^T,\ldots,(s_kE-A)^{-T}C^T\right\} &\subset \mathrm{range}(W), \end{aligned}$$

then

$$G(s_j)=\hat{G}(s_j), \quad \frac{d}{ds}G(s_j)=\frac{d}{ds}\hat{G}(s_j), \quad \text{for } j=1,\ldots,k.$$

Remarks:

computation of V, W from rational Krylov subspaces, e.g.,

- dual rational Arnoldi/Lanczos [Grimme 1997],
- Iter. Rational Krylov-Alg. (IRKA) [Antoulas/Beattie/Gugercin 2006/08].

Recall Interpolatory Model Reduction

Theorem (simplified) [Grimme 1997, Villemagne/Skelton 1987]

If

$$\begin{aligned}\operatorname{span}\left\{(s_1 E - A)^{-1}B, \ldots, (s_k E - A)^{-1}B\right\} &\subset \operatorname{range}(V),\\ \operatorname{span}\left\{(s_1 E - A)^{-T}C^T, \ldots, (s_k E - A)^{-T}C^T\right\} &\subset \operatorname{range}(W),\end{aligned}$$

then

$$G(s_j) = \hat{G}(s_j), \quad \frac{d}{ds}G(s_j) = \frac{d}{ds}\hat{G}(s_j), \quad \text{for } j = 1, \ldots, k.$$

Remarks:

using Galerkin/one-sided projection ($W \equiv V$) yields $G(s_j) = \hat{G}(s_j)$, but in general

$$\frac{d}{ds}G(s_j) \neq \frac{d}{ds}\hat{G}(s_j).$$

Recall Interpolatory Model Reduction

Theorem (simplified) [Grimme 1997, Villemagne/Skelton 1987]

If

$$\begin{aligned}\operatorname{span}\left\{(s_1E-A)^{-1}B,\ldots,(s_kE-A)^{-1}B\right\} &\subset \operatorname{range}(V),\\ \operatorname{span}\left\{(s_1E-A)^{-T}C^T,\ldots,(s_kE-A)^{-T}C^T\right\} &\subset \operatorname{range}(W),\end{aligned}$$

then

$$G(s_j)=\hat{G}(s_j),\quad \frac{d}{ds}G(s_j)=\frac{d}{ds}\hat{G}(s_j),\quad \text{for } j=1,\ldots,k.$$

Remarks:

$k=1$, standard Krylov subspace(s) of dimension K:

$$\operatorname{range}(V)=\mathcal{K}_K((s_1E-A)^{-1},(s_1E-A)^{-1}B).$$

$\rightsquigarrow$ **moment-matching methods**/Padé approximation,

$$\frac{d^i}{ds^i}G(s_1)=\frac{d^i}{ds^i}\hat{G}(s_1),\quad i=0,\ldots,K-1(+K).$$

- System in time domain:

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t). \end{aligned}$$

- System in frequency domain:

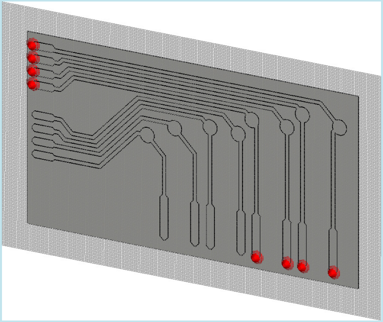
$$\begin{aligned} sEx(s) &= Ax(s) + Bu(s), \\ y(s) &= Cx(s). \end{aligned}$$

- Reduced basis method in a nutshell: consider s as a parameter, and use the system in frequency domain to compute

$$\mathrm{range}(V) = span\{x(s_1), \ldots, x(s_r)\}$$

by a greedy selection of $s_1, \ldots, s_r$. The ROM is obtained by Galerkin projection with V.

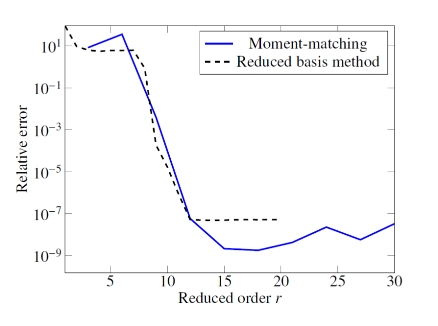
Printed circuit board

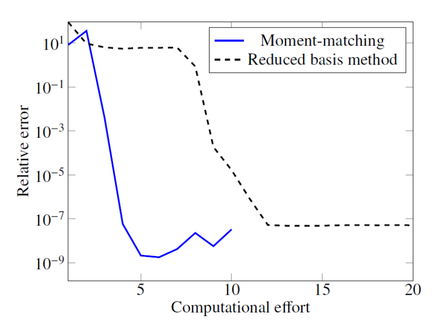


$n = 233,060, \; m = q = 1.$

Courtesy of TEMF, TU Darmstadt.

Moment-matching vs. reduced basis method





Consider stable (i.e. $\Lambda(A) \subset \mathbb{C}^-$) linear systems Σ,

$$\dot{x}(t) = Ax(t) + Bu(t),\ y(t) = Cx(t) \qquad \simeq \qquad y(s) = \underbrace{C(sI - A)^{-1}B}_{=:G(s)}\,u(s)$$

System norms

Recall: two common system norms for measuring approximation quality are

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

where

$$G(s) = C\,(sI - A)^{-1}\,B.$$

In order to find an $\mathcal{H}_2$-optimal reduced system, consider the error system $G(s) - \hat{G}(s)$ which can be realized by

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

In order to find an $\mathcal{H}_2$-optimal reduced system, consider the error system $G(s) - \hat{G}(s)$ which can be realized by

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

Assuming a coordinate system in which $\hat{A}$ is diagonal and taking derivatives of

$$\|G(\,.\,) - \hat{G}(\,.\,)\|^2_{\mathcal{H}_2}$$

with respect to free parameters in $\Lambda(\hat{A}), \hat{B}, \hat{C} \rightsquigarrow$ first-order necessary $\mathcal{H}_2$-optimality conditions (SISO)

$$G(-\hat{\lambda}_i) = \hat{G}(-\hat{\lambda}_i),$$
$$G'(-\hat{\lambda}_i) = \hat{G}'(-\hat{\lambda}_i),$$

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

In order to find an $\mathcal{H}_2$-optimal reduced system, consider the error system $G(s) - \hat{G}(s)$ which can be realized by

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

First-order necessary $\mathcal{H}_2$-optimality conditions (MIMO):

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

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

In order to find an $\mathcal{H}_2$-optimal reduced system, consider the error system $G(s) - \hat{G}(s)$ which can be realized by

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

First-order necessary $\mathcal{H}_2$-optimality conditions (MIMO):

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

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

for $i = 1, \ldots, r$ and $j = 1, \ldots, q$.

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

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

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

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

where V and W are given as

$$V = \left[(-\mu_1 I - A)^{-1}B, \ldots, (-\mu_r I - A)^{-1}B\right],$$
$$W = \left[(-\mu_1 I - A^T)^{-1}C^T, \ldots, (-\mu_r I - A^T)^{-1}C^T\right].$$

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

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

where V and W are given as

$$V = \left[(-\mu_1 I - A)^{-1}B, \ldots, (-\mu_r I - A)^{-1}B\right],$$
$$W = \left[(-\mu_1 I - A^T)^{-1}C^T, \ldots, (-\mu_r I - A^T)^{-1}C^T\right].$$

Then

$$G(-\mu_i) = \hat{G}(-\mu_i) \quad \text{and} \quad G'(-\mu_i) = \hat{G}'(-\mu_i),$$

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

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

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

where V and W are given as

$$V = \left[(-\mu_1 I - A)^{-1}B, \ldots, (-\mu_r I - A)^{-1}B\right],$$
$$W = \left[(-\mu_1 I - A^T)^{-1}C^T, \ldots, (-\mu_r I - A^T)^{-1}C^T\right].$$

Then

$$G(-\mu_i) = \hat{G}(-\mu_i) \quad \text{and} \quad G'(-\mu_i) = \hat{G}'(-\mu_i),$$

for $i = 1, \ldots, r$.
Starting with an initial guess for $\hat{\Lambda}$ and setting $\mu_i \equiv \hat{\lambda}_i \rightsquigarrow$ iterative algorithms (IRKA/MIRIAm) that yield $\mathcal{H}_2$-optimal models.

[Gugercin et al. 2006/08], [Bunse-Gerstner et al. 2007],
[Van Dooren et al. 2008]

Algorithm 1 IRKA (MIMO version/MIRIAm)

Input: A stable, B, C, $\hat{A}$ stable, $\hat{B}$, $\hat{C}$, $\delta > 0$.
Output: A^{opt}, B^{opt}, C^{opt}

while $(\max_{j=1,\ldots,r} \left\{ \frac{|\mu_j - \mu_j^{\text{old}}|}{|\mu_j|} \right\} > \delta)$ **do**
  $\mathrm{diag}\left(\mu_1, \ldots, \mu_r\right) := R^{-1}\hat{A}R =$ spectral decomposition.
  $\tilde{B} = \hat{B}^H R^{-T}$, $\tilde{C} = \hat{C}R$.
  $V = \left[(-\mu_1 I - A)^{-1} B\tilde{b}_1, \ldots, (-\mu_r I - A)^{-1} B\tilde{b}_r\right]$
  $W = \left[(-\mu_1 I - A^T)^{-1} C^T\tilde{c}_1, \ldots, (-\mu_r I - A^T)^{-1} C^T\tilde{c}_r\right]$
  $V = \mathrm{orth}(V)$, $W = \mathrm{orth}(W)$, $W = W(V^H W)^{-1}$
  $\hat{A} = W^H AV$, $\hat{B} = W^H B$, $\hat{C} = CV$.
end while
$A^{opt} = \hat{A}$, $B^{opt} = \hat{B}$, $C^{opt} = \hat{C}$.

Idea: choose appropriate frequency parameter $\hat{s}$ and parameter vector $\hat{p}$, expand into multivariate power series about $(\hat{s}, \hat{p})$ and compute reduced-order model, so that

$$G(s,p) = \hat{G}(s,p) + \mathcal{O}\left(|s-\hat{s}|^K + \|p-\hat{p}\|^L + |s-\hat{s}|^k \|p-\hat{p}\|^l \right),$$

i.e., first $K, L, k+l$ (mostly: $K = L = k+l$) coefficients (multi-moments) of Taylor/Laurent series coincide.

Idea: choose appropriate frequency parameter $\hat{s}$ and parameter vector $\hat{p}$, expand into multivariate power series about $(\hat{s}, \hat{p})$ and compute reduced-order model, so that

$$G(s,p) = \hat{G}(s,p) + \mathcal{O}\left(|s-\hat{s}|^K + \|p-\hat{p}\|^L + |s-\hat{s}|^k \|p-\hat{p}\|^l \right),$$

i.e., first $K, L, k+l$ (mostly: $K = L = k+l$) coefficients (multi-moments) of Taylor/Laurent series coincide.

Algorithms:

[1] [Daniel et al. 2004]: explicit computation of moments, numerically unstable.

[2] [Farle et al. 2006/07]: Krylov subspace approach, only polynomial param.-dependance, numerical properties not clear, but appears to be robust.

[3] [Weile et al. 1999, Feng/B. 2007/14]: Arnoldi-MGS method, employ recursive dependance of multi-moments, numerically robust, r often larger as for [2].

[4] **New:** employ dual-weighted residual error bound and greedy procedure to define interpolation points and # of multi-moments matched

[Antoulas/B./Feng 2014/17, B./Feng 2019/20].

Parametric System

Again, consider linear parametric system

$$\Sigma(p) : \begin{cases} E(p)\dot{x}(t;p) & = & A(p)x(t;p) + B(p)u(t), \\ y(t;p) & = & C(p)x(t;p) \end{cases}$$

together with its transfer function $G(s,p)$.

Parametric System

Again, consider linear parametric system

$$\Sigma(p):\begin{cases} E(p)\dot{x}(t;p) &= A(p)x(t;p)+B(p)u(t),\\ y(t;p) &= C(p)x(t;p)\end{cases}$$

together with its transfer function $G(s,p)$.

For simplicity, assume $B(\mu)\equiv B$, and re-parameterize — $\mu := [\,s, p^T\,]^T \in \mathbb{C}^\ell$ such that with

$$\begin{aligned} &G(\mu)\equiv G(s,p),\quad x(\mu)\equiv x(s,p),\quad y(\mu)\equiv y(s,p),\ldots\\ &\mathcal{A}(\mu):=sE(p)-A(p),\end{aligned}$$

we obtain linear-affine structure of $\mathcal{A}(\mu)$ (potentially after combining parameters):

$$\mathcal{A}(\mu)=\mathcal{A}_0+\mu_1\mathcal{A}_1+\ldots+\mu_\ell\mathcal{A}_\ell.$$

Parametric System

Again, consider linear parametric system

$$\Sigma(p): \begin{cases} E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \\ y(t;p) &= C(p)x(t;p) \end{cases}$$

together with its transfer function $G(s,p)$.

For simplicity, assume $B(\mu) \equiv B$, and re-parameterize — $\mu := [\, s, p^T \,]^T \in \mathbb{C}^\ell$ such that with

$$G(\mu) \equiv G(s,p), \quad x(\mu) \equiv x(s,p), \quad y(\mu) \equiv y(s,p), \ldots$$
$$\mathcal{A}(\mu) := sE(p) - A(p),$$

we obtain linear-affine structure of $\mathcal{A}(\mu)$ (potentially after combining parameters):

$$\mathcal{A}(\mu) = \mathcal{A}_0 + \mu_1\mathcal{A}_1 + \ldots + \mu_\ell\mathcal{A}_\ell.$$

In frequency domain, we may then re-write the parametric system as

$$\mathcal{A}(\mu)x(\mu) = Bu(s), \qquad y(\mu) = C(\mu)x(\mu).$$

Choose an expansion point $\mu^{(0)}$, and write

$$\begin{aligned}\mathcal{A}(\mu) &= \mathcal{A}_0 + \mu_1\mathcal{A}_1 + \ldots + \mu_\ell\mathcal{A}_m \\ &= \underbrace{(\mathcal{A}_0 + \mu_1^{(0)}\mathcal{A}_1 + \ldots + \mu_\ell^{(0)}\mathcal{A}_m)}_{:=\mathcal{M}_0} + \Big((\mu_1 - \mu_1^{(0)})\mathcal{A}_1 + \ldots + (\mu_\ell - \mu_\ell^{(0)})\mathcal{A}_\ell\Big)\end{aligned}$$

Choose an expansion point $\mu^{(0)}$, and write

$$
\begin{aligned}
\mathcal{A}(\mu) &= \mathcal{A}_0 + \mu_1 \mathcal{A}_1 + \ldots + \mu_\ell \mathcal{A}_m \\
&= \underbrace{(\mathcal{A}_0 + \mu_1^{(0)} \mathcal{A}_1 + \ldots + \mu_\ell^{(0)} \mathcal{A}_m)}_{:=\mathcal{M}_0} + \left((\mu_1 - \mu_1^{(0)})\mathcal{A}_1 + \ldots + (\mu_\ell - \mu_\ell^{(0)})\mathcal{A}_\ell\right) \\
&= \mathcal{M}_0 \left(I + (\mu_1 - \mu_1^{(0)})\mathcal{M}_0^{-1}\mathcal{A}_1 + \ldots + (\mu_\ell - \mu_\ell^{(0)})\mathcal{M}_0^{-1}\mathcal{A}_\ell\right)
\end{aligned}
$$

Choose an expansion point $\mu^{(0)}$, and write

$$\begin{aligned}\mathcal{A}(\mu) &= \mathcal{A}_0 + \mu_1\mathcal{A}_1 + \ldots + \mu_\ell\mathcal{A}_m \\ &= \underbrace{(\mathcal{A}_0 + \mu_1^{(0)}\mathcal{A}_1 + \ldots + \mu_\ell^{(0)}\mathcal{A}_m)}_{:=\mathcal{M}_0} + \left((\mu_1 - \mu_1^{(0)})\mathcal{A}_1 + \ldots + (\mu_\ell - \mu_\ell^{(0)})\mathcal{A}_\ell\right) \\ &= \mathcal{M}_0\left(I + (\mu_1 - \mu_1^{(0)})\mathcal{M}_0^{-1}\mathcal{A}_1 + \ldots + (\mu_\ell - \mu_\ell^{(0)})\mathcal{M}_0^{-1}\mathcal{A}_\ell\right)\end{aligned}$$

Using the Neumann lemma ($(I-F)^{-1} = \sum_{j=0}^{\infty} F^j$ if $\|F\| < 1$), we obtain

$$\begin{aligned}\mathcal{A}(\mu)^{-1} &= \sum_{j=0}^{\infty}(-1)^j\left((\mu_1 - \mu_1^{(0)})\mathcal{M}_0^{-1}\mathcal{A}_1 + \ldots + (\mu_\ell - \mu_\ell^{(0)})\mathcal{M}_0^{-1}\mathcal{A}_\ell\right)^j \mathcal{M}_0^{-1} \\ &= \sum_{j=0}^{\infty}(\sigma_1\mathcal{M}_1 + \ldots \sigma_\ell\mathcal{M}_\ell)^j\mathcal{M}_0^{-1},\end{aligned}$$

where $\sigma_i = \mu_i - \mu_i^{(0)}$ and $\mathcal{M}_i = -\mathcal{M}_0^{-1}\mathcal{A}_i$ for $i = 1, \ldots, \ell$.

We have

$$\mathcal{A}(\mu)x(\mu) = Bu(s).$$

and

$$\mathcal{A}(\mu)^{-1} = \sum_{j=0}^{\infty}(\sigma_1\mathcal{M}_1 + \ldots \sigma_\ell\mathcal{M}_\ell)^j\mathcal{M}_0^{-1},$$

where $\sigma_i = \mu_i - \mu_i^{(0)}$, $\mathcal{M}_0 = \mathcal{A}(\mu^{(0)})$ and $\mathcal{M}_i = -\mathcal{M}_0^{-1}\mathcal{A}_i$ for $i = 0, \ldots, \ell$.

We have

$$\mathcal{A}(\mu)x(\mu) = Bu(s).$$

and

$$\mathcal{A}(\mu)^{-1} = \sum_{j=0}^{\infty} (\sigma_1 \mathcal{M}_1 + \ldots \sigma_\ell \mathcal{M}_\ell)^j \mathcal{M}_0^{-1},$$

where $\sigma_i = \mu_i - \mu_i^{(0)}$, $\mathcal{M}_0 = \mathcal{A}(\mu^{(0)})$ and $\mathcal{M}_i = -\mathcal{M}_0^{-1}\mathcal{A}_i$ for $i = 0, \ldots, \ell$.

Hence,

$$\begin{aligned} x(\mu) = \mathcal{A}(\mu)^{-1}Bu(s) &= \sum_{j=0}^{\infty} (\sigma_1 \mathcal{M}_1 + \ldots + \sigma_\ell \mathcal{M}_\ell)^j \underbrace{\mathcal{M}_0^{-1}B}_{=:\mathcal{B}} u(s) \\ &\approx \sum_{j=0}^{k} (\sigma_1 \mathcal{M}_1 + \ldots + \sigma_\ell \mathcal{M}_\ell)^j \mathcal{B}u(s) =: \tilde{x}(\mu). \end{aligned}$$

We have

$$\mathcal{A}(\mu)x(\mu) = Bu(s).$$

and

$$\mathcal{A}(\mu)^{-1} = \sum_{j=0}^{\infty}(\sigma_1\mathcal{M}_1 + \ldots \sigma_\ell\mathcal{M}_\ell)^j\mathcal{M}_0^{-1},$$

where $\sigma_i = \mu_i - \mu_i^{(0)}$, $\mathcal{M}_0 = \mathcal{A}(\mu^{(0)})$ and $\mathcal{M}_i = -\mathcal{M}_0^{-1}\mathcal{A}_i$ for $i = 0, \ldots, \ell$.

Hence,

$$\begin{aligned} x(\mu) = \mathcal{A}(\mu)^{-1}Bu(s) &= \sum_{j=0}^{\infty}(\sigma_1\mathcal{M}_1 + \ldots + \sigma_\ell\mathcal{M}_\ell)^j \underbrace{\mathcal{M}_0^{-1}B}_{=:\mathcal{B}} u(s) \\ &\approx \sum_{j=0}^{k}(\sigma_1\mathcal{M}_1 + \ldots + \sigma_\ell\mathcal{M}_\ell)^j\mathcal{B}u(s) =: \tilde{x}(\mu). \end{aligned}$$

Thus, $x(\mu)$ is (approximately, locally) contained in the Krylov subspace $\mathcal{K}_{k+1}((\sigma_1\mathcal{M}_1 + \ldots + \sigma_\ell\mathcal{M}_\ell), \mathcal{B})$.

PMOR based on Multi-Moment Matching

$x(\mu)$ is (approximately, locally) contained in the Krylov subspace $\mathcal{K}_{k+1}((\sigma_1\mathcal{M}_1 + \ldots \sigma_\ell\mathcal{M}_\ell), \mathcal{B}) \Longrightarrow$

- Project the state-space onto this subspace.

PMOR based on Multi-Moment Matching

$x(\mu)$ is (approximately, locally) contained in the Krylov subspace $\mathcal{K}_{k+1}((\sigma_1\mathcal{M}_1 + \ldots \sigma_\ell\mathcal{M}_\ell), \mathcal{B}) \Longrightarrow$

- Project the state-space onto this subspace.
- Obtain an orthogonal basis using block-Arnoldi-MGS [B./Feng 2007/14], or TOAR [Bai/Su 2008].

PMOR based on Multi-Moment Matching

$x(\mu)$ is (approximately, locally) contained in the Krylov subspace $\mathcal{K}_{k+1}((\sigma_1\mathcal{M}_1 + \ldots \sigma_\ell\mathcal{M}_\ell), \mathcal{B}) \Longrightarrow$

- Project the state-space onto this subspace.
- Obtain an orthogonal basis using block-Arnoldi-MGS [B./Feng 2007/14], or TOAR [Bai/Su 2008].
- The ROM is obtained by Galerkin projection.

PMOR based on Multi-Moment Matching

$x(\mu)$ is (approximately, locally) contained in the Krylov subspace $\mathcal{K}_{k+1}((\sigma_1\mathcal{M}_1 + \ldots \sigma_\ell\mathcal{M}_\ell), \mathcal{B}) \Longrightarrow$

- Project the state-space onto this subspace.
- Obtain an orthogonal basis using block-Arnoldi-MGS [B./Feng 2007/14], or TOAR [Bai/Su 2008].
- The ROM is obtained by Galerkin projection.
- Petrov-Galerkin projection possible using the "dual" Krylov subspace obtained from using $\mathcal{A}^T$ and C^T [Ahmad/B./Feng 2017].

$x(\mu)$ is (approximately, locally) contained in the Krylov subspace $\mathcal{K}_{k+1}((\sigma_1 \mathcal{M}_1 + \ldots \sigma_\ell \mathcal{M}_\ell), \mathcal{B}) \Longrightarrow$

- Project the state-space onto this subspace.
- Obtain an orthogonal basis using block-Arnoldi-MGS [B./Feng 2007/14], or TOAR [Bai/Su 2008].
- The ROM is obtained by Galerkin projection.
- Petrov-Galerkin projection possible using the "dual" Krylov subspace obtained from using $\mathcal{A}^T$ and C^T [Ahmad/B./Feng 2017].
- First terms in the multivarite Taylor expansion match, i.e., we achieve matrix interpolation for partial derivatives up to order ℓ, or more in the Petrov-Galerkin case.

$x(\mu)$ is (approximately, locally) contained in the Krylov subspace $\mathcal{K}_{k+1}((\sigma_1\mathcal{M}_1 + \ldots \sigma_\ell\mathcal{M}_\ell), \mathcal{B}) \Longrightarrow$

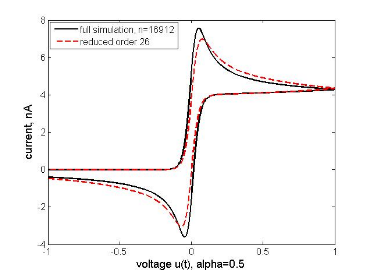
- Project the state-space onto this subspace.
- Obtain an orthogonal basis using block-Arnoldi-MGS [B./Feng 2007/14], or TOAR [Bai/Su 2008].
- The ROM is obtained by Galerkin projection.
- Petrov-Galerkin projection possible using the "dual" Krylov subspace obtained from using $\mathcal{A}^T$ and C^T [Ahmad/B./Feng 2017].
- First terms in the multivarite Taylor expansion match, i.e., we achieve matrix interpolation for partial derivatives up to order ℓ, or more in the Petrov-Galerkin case.
- Approximation is only valid locally (convergence radius of Neumann series!) $\rightsquigarrow$ use several expansion points $\mu^{(0)}, \ldots, \mu^{(h)}$, and concatenate (and truncate) the local bases to obtain a global basis.

Compute cyclic voltammogram based on FE model

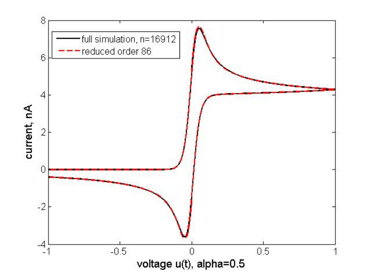
$$E\dot{x}(t) = (A_0 + p_1 A_1 + p_2 A_2)x(t) + Bu(t), \quad y(t) = c^T x(t),$$

where $n = 16,912$, $m = 3$, A_1, A_2 diagonal.

$K = L = k + \ell = 4 \Rightarrow r = 26$



$K = L = k + \ell = 9 \Rightarrow r = 86$



Source: MOR Wiki: http://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Scanning_Electrochemical_Microscopy

Open question

How to adaptively choose $\mu^{(i)}$?

Open question

How to adaptively choose $\mu^{(i)}$?

And how many partial derivatives to be matched at each interpolation point?

Open question

How to adaptively choose $\mu^{(i)}$?

And how many partial derivatives to be matched at each interpolation point?

Possible approach: adopt ideas from Reduced Basis Methods, i.e., let

$$\|G(\mu) - \hat{G}(\mu)\| \leq \Delta(\mu) \quad \text{or} \quad \|y(\mu) - \hat{y}(\mu)\| \leq \Delta_o(\mu)$$

guide the selection of $\mu^{(i)}$ for computable *a posteriori* error bounds for the state or the output.

Error Bound for Automatic ROM Construction

Theorem (SISO case) [Feng/Antoulas/B. 2014/17]

Assume that $\sigma_{\min}(G(s,p)) =: \beta(s,p) > 0 \quad \forall\, \mathrm{Re}(s) \geq 0, \forall\, p \in \Omega$, then

$$|G(s,p) - \hat{G}(s,p)| \leq \tilde{\Delta}(s,p) + |(\hat{x}^{du})^H \mathrm{r}^{pr}(s,p)| =: \Delta(s,p),$$

where

$$\tilde{\Delta}(s,p) = \frac{\|r^{du}(s,p)\|_2 \|r^{pr}(s,p)\|_2}{\beta(s,p)},$$

with the primal and dual residuals r^{pr}, r^{du} and the reduced "dual state" $\hat{x}^{du}$:

$$\begin{aligned}
r^{pr}(s,p) &= \|\,(B - (sE(p) - A(p)))\left(V(s\hat{E}(p) - \hat{A}(p))^{-1}\hat{B}\right)\|,\\
r^{du}(s,p) &= \|\left(C^T - (\bar{s}E(p) - A(p))^T\right)\hat{x}^{du}\|,\\
\hat{x}^{du} &= -V^{du}(\bar{s}\hat{E}^{du}(p) - \hat{A}^{du}(p))^{-T}\hat{C}^{du}.
\end{aligned}$$

The dual reduced-order system is computed using Galerkin projection with V^{du} obtained by applying multi-moment matching algorithm to "dual" system $(\bar{s}E(p)^T - A(p)^T, C^T)$.

- For application in "RBM fashion", $r^{du}(\mu), r^{pr}(\mu)$ can be efficiently computed, need to solve sparse linear systems on training set, i.e., one sparse factorization for each sampling point.
- $\beta(s,p) = \sigma_{\min}(G(s,p))$ easily computable on the training set — system solves for evaluation of the transfer function readily available from residual computation!
- Extension to MIMO case possible taking max over all I/O channels.
- Can use Petrov-Galerkin framework using $W = V^{du}$ at no extra cost!

Algorithm 2 Automatic generation of the ROM: adaptively selecting $\mu^{(i)}$

Input: $V = [\,]$; $\epsilon > \epsilon_{tol}$; Initial expansion point: $\hat{\mu}$; $i := -1$;
Ξ_{train}: a set of samples of μ covering the parameter domain.

Output: V.

while $\epsilon > \epsilon_{tol}$ **do**
$\quad i = i + 1$;
$\quad \mu^{(i)} = \hat{\mu}$;
$\quad V_{\mu^{(i)}} =$ orthogonal basis of $\mathcal{K}_{k+1}((\sigma_1^{(i)}\mathcal{M}_1 + \ldots + \sigma_\ell^{(i)}\mathcal{M}_\ell), \mathcal{B})$;
$\quad V = \mathtt{orth}([V, V_{\mu^{(i)}}])$;
$\quad \hat{\mu} = \arg \max_{\mu \in \Xi_{train}} \Delta(\mu)$;
$\quad \epsilon = \Delta(\hat{\mu})$;
end while

PMOR based on Multi-Moment Matching

Numerical Example: Silicon Nitride Membrane

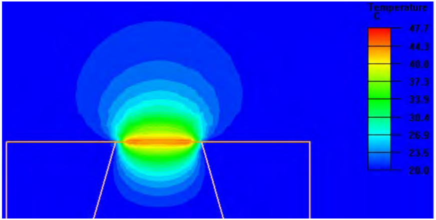
A SiN membrane can be a part of a gas sensor, an infrared sensor, a microthruster, etc. Heat tansfer in the membrane is described by

$$
\begin{aligned}
(E_0 + \rho c_p E_1)\dot{x}(t) &= -(K_0 + \kappa K_1 + h K_2)x(t) + bu(t) \\
y(t) &= Cx(t),
\end{aligned}
$$

with parameters

- density $\rho \in [3000, 3200]$,
- specific heat capacity $c_p \in [400, 750]$,
- thermal conductivity $\kappa \in [2.5, 4]$,
- membrane heat transfer coefficient $h \in [10, 12]$.

and frequency $f \in [0, 25] Hz$.



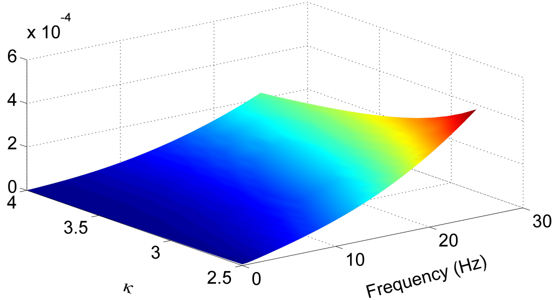
Source: MOR Wiki: http://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Silicon_nitride_membrane

Setting

- Training set: Ξ_{train} = 5 random samples for ρ and c_p, 3 random samples for κ and h, respectively, 10 samples of Laplace variable s.
- Error measures:
  $\varepsilon_{true}^{rel} = \max_{\mu \in \Xi_{train}} |G(\mu) - \hat{G}(\mu)|/|G(\mu)|, \qquad \Delta^{rel}(\mu) = \Delta(\mu)/|\hat{G}(\mu)|$
- $V_{\mu^{(i)}} = \mathrm{span}\{\mathcal{B}, (\sigma_1^{(i)}\mathcal{M}_1 + \ldots + \sigma_\ell^{(i)}\mathcal{M}_\ell)\mathcal{B}\}$, $\epsilon_{tol}^{rel} = 10^{-2}$, $n = 60,020$, $r = 8$.

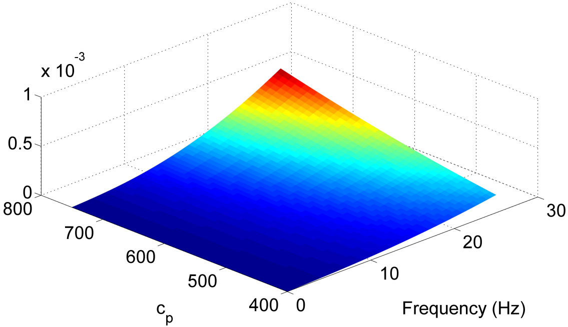
iter.	ε_{true}^{rel}	$\Delta^{rel}(\mu^{(i)})$	s	ρc_p	κ	h
1	1×10^{-3}	3.44	18.94	1.37×10^{6}	2.74	10.97
2	1×10^{-4}	4.59×10^{-2}	0.89	1.31×10^{6}	3.96	11.60
3	2.80×10^{-5}	4.07×10^{-2}	23.98	2.35×10^{6}	3.94	10.28
4	2.58×10^{-6}	2.62×10^{-5}	0.89	2.31×10^{6}	2.74	10.28

Verification of the accuracy of the ROM for κ over set Ξ_{fine} with 16 equidistant samples of κ, 51 equidistant samples of the frequency f, while the other parameters are fixed.



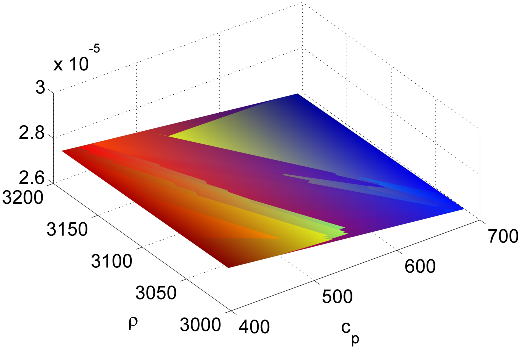
Relative error of the final ROM changing with κ and frequency.

Verification of the accuracy of the ROM for c_p over set Ξ_{fine} with 36 equidistant samples of c_p, 51 equidistant samples of the frequency f, while the other parameters are fixed.



Relative error of the final ROM changing with c_p and frequency.

Verification of the accuracy of the ROM for ρ, c_p over set Ξ_{fine} with 50 random samples of ρ, c_p, respectively, the other parameters are fixed.



Relative error of the final ROM changing with c_p and κ.

Optimal PMOR using Rational Interpolation?

Greedy expansion point selection has a heuristic nature and relies on a training set.

Optimal PMOR using Rational Interpolation?

Greedy expansion point selection has a heuristic nature and relies on a training set.

How to determine the right number of partial derivatives to be matched at the expansion points is an open problem (for potential solutions in the non-parametric case, see [Feng/Korvink/B. 2015, Bonin/Fassbender/Soppa/Zäh 2016, Lee/Chu/Feng 2006,...].

Optimal PMOR using Rational Interpolation?

Greedy expansion point selection has a heuristic nature and relies on a training set.

How to determine the right number of partial derivatives to be matched at the expansion points is an open problem (for potential solutions in the non-parametric case, see [Feng/Korvink/B. 2015, Bonin/Fassbender/Soppa/Zäh 2016, Lee/Chu/Feng 2006,...].

Question

Can we find (necessary) optimality conditions similar to the LTI case, leading to an IRKA-like procedure?

Optimal PMOR using Rational Interpolation?

Greedy expansion point selection has a heuristic nature and relies on a training set.

How to determine the right number of partial derivatives to be matched at the expansion points is an open problem (for potential solutions in the non-parametric case, see [Feng/Korvink/B. 2015, Bonin/Fassbender/Soppa/Zäh 2016, Lee/Chu/Feng 2006,...].

Question

Can we find (necessary) optimality conditions similar to the LTI case, leading to an IRKA-like procedure?

Hence, we investigate the problem: for a given order r of the reduced-order model, can we provide necessary conditions for a rational interpolant to minimize

$$\|G - \hat{G}\|^2_{\mathcal{H}_2 \times L_2(\Omega)} := \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{\Omega} \left\| G(\jmath\omega, p) - \hat{G}(\jmath\omega, p) \right\|_F^2 \mathrm{d}p_1 \ldots \mathrm{d}p_d \, \mathrm{d}\omega \ ?$$

Optimal PMOR using Rational Interpolation?

Greedy expansion point selection has a heuristic nature and relies on a training set.

How to determine the right number of partial derivatives to be matched at the expansion points is an open problem (for potential solutions in the non-parametric case, see [Feng/Korvink/B. 2015, Bonin/Fassbender/Soppa/Zäh 2016, Lee/Chu/Feng 2006,...].

Question

Can we find (necessary) optimality conditions similar to the LTI case, leading to an IRKA-like procedure?

Hence, we investigate the problem: for a given order r of the reduced-order model, can we provide necessary conditions for a rational interpolant to minimize

$$\|G-\hat{G}\|^2_{\mathcal{H}_2\times L_2(\Omega)} := \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{\Omega}\left\|G(\jmath\omega,\,p)-\hat{G}(\jmath\omega,\,p)\right\|_F^2 \mathrm{d}p_1\ldots\mathrm{d}p_d\,\mathrm{d}\omega\ ?$$

Following the non-parametric case, one would need:

- Projection-based framework for tangential rational interpolation. [✓]

Optimal PMOR using Rational Interpolation?

Greedy expansion point selection has a heuristic nature and relies on a training set.

How to determine the right number of partial derivatives to be matched at the expansion points is an open problem (for potential solutions in the non-parametric case, see [Feng/Korvink/B. 2015, Bonin/Fassbender/Soppa/Zäh 2016, Lee/Chu/Feng 2006,...].

Question

Can we find (necessary) optimality conditions similar to the LTI case, leading to an IRKA-like procedure?

Hence, we investigate the problem: for a given order r of the reduced-order model, can we provide necessary conditions for a rational interpolant to minimize

$$\|G - \hat{G}\|^2_{\mathcal{H}_2 \times L_2(\Omega)} := \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{\Omega} \left\| G(\jmath\omega, p) - \hat{G}(\jmath\omega, p) \right\|_F^2 \mathrm{d}p_1 \dots \mathrm{d}p_d \, \mathrm{d}\omega \text{ ?}$$

Following the non-parametric case, one would need:

- Projection-based framework for tangential rational interpolation. [✓]
- Iterative procedure for selecting interpolation points. [χ] ... [✓] for special case.

Theorem [Baur/Beattie/B./Gugercin 2007/11]

Let
$$\begin{aligned}\hat{G}(s,p) &:= \hat{C}(p)(s\hat{E}(p)-\hat{A}(p))^{-1}\hat{B}(p)\\ &= C(p)V(sW^TE(p)V-W^TA(p)V)^{-1}W^TB(p).\end{aligned}$$

Suppose $\hat{p}=[\hat{p}_1,\dots,\hat{p}_d]^T$ and $\hat{s}\in\mathbb{C}$ are chosen such that both $\hat{s}\,E(\hat{p})-A(\hat{p})$ and $\hat{s}\,\hat{E}(\hat{p})-\hat{A}(\hat{p})$ are invertible. If

$$(\hat{s}\,E(\hat{p})-A(\hat{p}))^{-1}B(\hat{p})\in\mathrm{range}(V)$$

or

$$\left(C(\hat{p})\,(\hat{s}\,E(\hat{p})-A(\hat{p}))^{-1}\right)^T\in\mathrm{range}(W),$$

then $G(\hat{s},\hat{p})=\hat{G}(\hat{s},\hat{p})$.

Theorem [Baur/Beattie/B./Gugercin 2007/11]

Let
$$\hat{G}(s,p) := \hat{C}(p)(s\hat{E}(p) - \hat{A}(p))^{-1}\hat{B}(p) = C(p)V(sW^T E(p)V - W^T A(p)V)^{-1} W^T B(p).$$

Suppose $\hat{p} = [\hat{p}_1, \ldots, \hat{p}_d]^T$ and $\hat{s} \in \mathbb{C}$ are chosen such that both $\hat{s}\,E(\hat{p}) - A(\hat{p})$ and $\hat{s}\,\hat{E}(\hat{p}) - \hat{A}(\hat{p})$ are invertible. If

$$(\hat{s}\,E(\hat{p}) - A(\hat{p}))^{-1}\,B(\hat{p}) \in \mathrm{range}(V)$$

or

$$\left(C(\hat{p})\,(\hat{s}\,E(\hat{p}) - A(\hat{p}))^{-1}\right)^T \in \mathrm{range}(W)\,,$$

then $G(\hat{s},\hat{p}) = \hat{G}(\hat{s},\hat{p})$.

Extension to MIMO case using tangential interpolation: let $0 \neq b \in \mathbb{R}^m$, $0 \neq c \in \mathbb{R}^q$.

a) If $(\hat{s}\,E(\hat{p}) - A(\hat{p}))^{-1}\,B(\hat{p})b \in \mathrm{range}(V)$, then $G(\hat{s},\hat{p})b = \hat{G}(\hat{s},\hat{p})b$.

b) If $\left(c^T C(\hat{p})\,(\hat{s}\,E(\hat{p}) - A(\hat{p}))^{-1}\right)^T \in \mathrm{range}(W)$, then $c^T G(\hat{s},\hat{p}) = c^T \hat{G}(\hat{s},\hat{p})$.

Theorem [Baur/Beattie/B./Gugercin 2007/11]

Suppose that $E(p)$, $A(p)$, $B(p)$, $C(p)$ are C^1 in a neighborhood of $\hat{p} = [\hat{p}_1, \ldots, \hat{p}_d]^T$ and that both $\hat{s}\,E(\hat{p}) - A(\hat{p})$ and $\hat{s}\,\hat{E}(\hat{p}) - \hat{A}(\hat{p})$ are invertible.

If

$$(\hat{s}\,E(\hat{p}) - A(\hat{p}))^{-1}\,B(\hat{p}) \in \mathrm{range}(V)$$

and

$$\left(C(\hat{p})\,(\hat{s}\,E(\hat{p}) - A(\hat{p}))^{-1}\right)^T \in \mathrm{range}(W)\,,$$

then

$$\nabla_p G(\hat{s}, \hat{p}) = \nabla_p G_r(\hat{s}, \hat{p}), \qquad \frac{\partial}{\partial s} G(\hat{s}, \hat{p}) = \frac{\partial}{\partial s} \hat{G}(\hat{s}, \hat{p}).$$

Theorem [Baur/Beattie/B./Gugercin 2007/11]

Suppose that $E(p)$, $A(p)$, $B(p)$, $C(p)$ are C^1 in a neighborhood of $\hat{p} = [\hat{p}_1, \ldots, \hat{p}_d]^T$ and that both $\hat{s}\,E(\hat{p}) - A(\hat{p})$ and $\hat{s}\,\hat{E}(\hat{p}) - \hat{A}(\hat{p})$ are invertible.

If

$$(\hat{s}\,E(\hat{p}) - A(\hat{p}))^{-1}\,B(\hat{p}) \in \mathrm{range}(V)$$

and

$$\left(C(\hat{p})\,(\hat{s}\,E(\hat{p}) - A(\hat{p}))^{-1}\right)^T \in \mathrm{range}(W),$$

then

$$\nabla_p G(\hat{s}, \hat{p}) = \nabla_p G_r(\hat{s}, \hat{p}), \qquad \frac{\partial}{\partial s} G(\hat{s}, \hat{p}) = \frac{\partial}{\partial s} \hat{G}(\hat{s}, \hat{p}).$$

Note: result extends to MIMO case using tangential interpolation:

Let $0 \neq b \in \mathbb{R}^m$, $0 \neq c \in \mathbb{R}^q$ be arbitrary. If $(\hat{s}\,E(\hat{p}) - A(\hat{p}))^{-1}\,B(\hat{p})b \in \mathrm{range}(V)$ and $\left(c^T C(\hat{p})\,(\hat{s}\,E(\hat{p}) - A(\hat{p}))^{-1}\right)^T \in \mathrm{range}(W)$, then $\nabla_p c^T G(\hat{s}, \hat{p})b = \nabla_p c^T \hat{G}(\hat{s}, \hat{p})b$.

Theorem [Baur/Beattie/B./Gugercin 2007/11]

Suppose that $E(p)$, $A(p)$, $B(p)$, $C(p)$ are C^1 in a neighborhood of $\hat{p} = [\hat{p}_1, \ldots, \hat{p}_d]^T$ and that both $\hat{s}\, E(\hat{p}) - A(\hat{p})$ and $\hat{s}\, \hat{E}(\hat{p}) - \hat{A}(\hat{p})$ are invertible.

If

$$(\hat{s}\, E(\hat{p}) - A(\hat{p}))^{-1} B(\hat{p}) \in \mathrm{range}(V)$$

and

$$\left(C(\hat{p})\, (\hat{s}\, E(\hat{p}) - A(\hat{p}))^{-1} \right)^T \in \mathrm{range}(W),$$

then

$$\nabla_p G(\hat{s}, \hat{p}) = \nabla_p G_r(\hat{s}, \hat{p}), \qquad \frac{\partial}{\partial s} G(\hat{s}, \hat{p}) = \frac{\partial}{\partial s} \hat{G}(\hat{s}, \hat{p}).$$

1. Reduced-order model satisfies necessary conditions for surrogate models in trust region methods [Alexandrov/Dennis/Lewis/Torczon 1998].
2. Approximation of gradient allows use of reduced-order model for sensitivity analysis.

Optimal PMOR using Rational Interpolation?

Generic implementation of interpolatory PMOR

Define $\mathcal{A}(s,p) := sE(p) - A(p)$.

1. Select "frequencies" $s_1, \ldots, s_k \in \mathbb{C}$ and parameter vectors $p^{(1)}, \ldots, p^{(\ell)} \in \Omega \subset \mathbb{R}^d$.
2. Compute (orthonormal) basis of
$$\mathcal{V} = \operatorname{span}\left\{\mathcal{A}(s_1, p^{(1)})^{-1}B(p^{(1)}), \ldots, \mathcal{A}(s_k, p^{(\ell)})^{-1}B(p^{(\ell)})\right\}.$$
3. Compute (orthonormal) basis of
$$\mathcal{W} = \operatorname{span}\left\{\mathcal{A}(s_1, p^{(1)})^{-T}C(p^{(1)})^T, \ldots, \mathcal{A}(s_k, p^{(\ell)})^{-T}C(p^{(\ell)})^T\right\}.$$
4. Set $V := [v_1, \ldots, v_{k\ell}]$, $\tilde{W} := [w_1, \ldots, w_{k\ell}]$, and $W := \tilde{W}(\tilde{W}^T V)^{-1}$.
(Note: $r = k\ell$).
5. Compute $\begin{cases} \hat{A}(p) := W^T A(p) V, & \hat{B}(p) := W^T B(p) V, \\ \hat{C}(p) := W^T C(p) V, & \hat{E}(p) := W^T E(p) V. \end{cases}$

- If directional derivatives w.r.t. p are included in range(V), range(W), then also the Hessian of $G(\hat{s}, \hat{p})$ is interpolated by the Hessian of $\hat{G}(\hat{s}, \hat{p})$.

Optimal PMOR using Rational Interpolation?

Remarks

- If directional derivatives w.r.t. p are included in range(V), range(W), then also the Hessian of $G(\hat{s}, \hat{p})$ is interpolated by the Hessian of $\hat{G}(\hat{s}, \hat{p})$.
- Choice of optimal interpolation frequencies s_k and parameter vectors $p^{(k)}$ in general is an open problem.

- If directional derivatives w.r.t. p are included in range(V), range(W), then also the Hessian of $G(\hat{s}, \hat{p})$ is interpolated by the Hessian of $\hat{G}(\hat{s}, \hat{p})$.
- Choice of optimal interpolation frequencies s_k and parameter vectors $p^{(k)}$ in general is an open problem.
- For prescribed parameter vectors $p^{(k)}$, we can use corresponding $\mathcal{H}_2$-optimal frequencies $s_{k,\ell}$, $\ell = 1, \ldots, r_k$, computed by IRKA, i.e., reduced-order systems $\hat{G}_*^{(k)}$ so that

$$\|G(., p^{(k)}) - \hat{G}_*^{(k)}(.)\|_{\mathcal{H}_2} = \min_{\substack{\mathrm{order}(\hat{G})=r_k \\ \hat{G}\ \mathrm{stable}}} \|G(., p^{(k)}) - \hat{G}^{(k)}(.)\|_{\mathcal{H}_2},$$

where

$$\|G\|_{\mathcal{H}_2} := \left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} \|G(\jmath\omega)\|_{\mathrm{F}}^2 d\omega\right)^{1/2}.$$

- If directional derivatives w.r.t. p are included in range(V), range(W), then also the Hessian of $G(\hat{s},\hat{p})$ is interpolated by the Hessian of $\hat{G}(\hat{s},\hat{p})$.
- Choice of optimal interpolation frequencies s_k and parameter vectors $p^{(k)}$ in general is an open problem.
- For prescribed parameter vectors $p^{(k)}$, we can use corresponding $\mathcal{H}_2$-optimal frequencies $s_{k,\ell}$, $\ell = 1,\ldots,r_k$, computed by IRKA, i.e., reduced-order systems $\hat{G}_*^{(k)}$ so that

$$\|G(.,p^{(k)}) - \hat{G}_*^{(k)}(.)\|_{\mathcal{H}_2} = \min_{\substack{\mathrm{order}(\hat{G})=r_k\\ \hat{G}\ \mathrm{stable}}} \|G(.,p^{(k)}) - \hat{G}^{(k)}(.)\|_{\mathcal{H}_2},$$

  where

$$\|G\|_{\mathcal{H}_2} := \left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} \|G(\jmath\omega)\|_{\mathrm{F}}^2 d\omega\right)^{1/2}.$$

- Optimal choice of interpolation frequencies s_k and parameter vectors $p^{(k)}$ possible for special cases.

- Important requirement for a compact model of thermal conduction is boundary condition independence.
- The thermal problem is modeled by the heat equation, where heat exchange through device interfaces is modeled by convection boundary conditions containing film coefficients $\{p_i\}_{i=1}^3$ describing the heat exchange at ith interface.
- Spatial semi-discretization leads to

$$E\dot{x}(t) = (A_0 + \sum_{i=1}^{3} p_i A_i)x(t) + bu(t), \quad y(t) = c^T x(t),$$

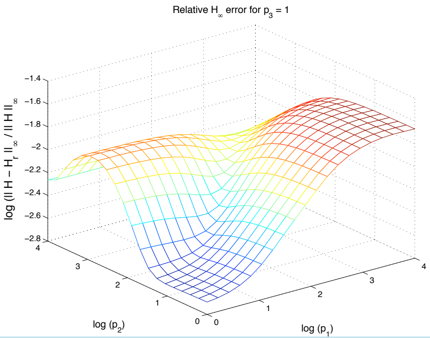
where $n = 4,257$, A_i, $i = 1, 2, 3$, are diagonal.

Source: C.J.M Lasance, *Two benchmarks to facilitate the study of compact thermal modeling phenomena*, IEEE Transactions on Components and Packaging Technologies, 24(4):559–565, 2001.

MOR Wiki: http://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Microthruster_Unit

Choose 2 interpolation points for parameters ("important" configurations), 8/7 H_2-optimal interpolation frequencies selected by IRKA. $\Rightarrow$ $k = 2, \ell = 8, 7$, hence $r = 15$.

$p_3 = 1,\ p_1, p_2 \in [1, 10^4]$.



Theorem [Baur/Beattie/B./Gugercin 2011]

For special parameterized SISO systems,

$$A(p) \equiv A_0,\ E(p) \equiv E_0,\ B(p) = B_0 + p_1 B_1,\ C(p) = C_0 + p_2 C_1,$$

optimal choice possible, necessary conditions:
If $\hat{G}$ minimizes the approximation error w.r.t. $\|G - \hat{G}\|_{\mathcal{H}_2 \times L_2(\Omega)}$, $p \in \Omega \subset \mathbb{R}^d$, and $\Lambda(\hat{A}, \hat{E}) = \{\hat{\lambda}_1, \ldots, \hat{\lambda}_r\}$ (all simple), then the interpolation frequencies satisfy

$$s_i = -\hat{\lambda}_i, \quad i = 1, \ldots, r,$$

and the parameter interpolation points $\{p^{(1)}, \ldots, p^{(r)}\}$ satisfy the interpolation conditions

$$\begin{aligned} G(-\hat{\lambda}_k, p^{(k)}) &= \hat{G}(-\hat{\lambda}, p^{(k)}), \\ \frac{\partial}{\partial s} G(-\hat{\lambda}, p^{(k)}) &= \frac{\partial}{\partial s} \hat{G}(-\hat{\lambda}, p^{(k)}), \quad \nabla_p G(-\hat{\lambda}, p^{(k)}) = \nabla_p \hat{G}(-\hat{\lambda}, p^{(k)}). \end{aligned}$$

Theorem [Baur/Beattie/B./Gugercin 2011]

For special parameterized SISO systems,

$$A(p) \equiv A_0,\ E(p) \equiv E_0,\ B(p) = B_0 + p_1 B_1,\ C(p) = C_0 + p_2 C_1,$$

optimal choice possible, necessary conditions:
If $\hat{G}$ minimizes the approximation error w.r.t. $\|G - \hat{G}\|_{\mathcal{H}_2 \times L_2(\Omega)}$, $p \in \Omega \subset \mathbb{R}^d$, the parameter interpolation points $\{p^{(1)}, \ldots, p^{(r)}\}$ satisfy the interpolation conditions

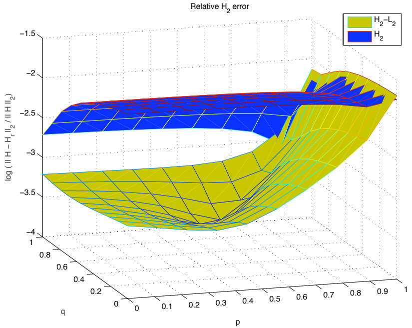
$$\begin{aligned} G(-\hat{\lambda}_k, p^{(k)}) &= \hat{G}(-\hat{\lambda}, p^{(k)}), \\ \frac{\partial}{\partial s} G(-\hat{\lambda}, p^{(k)}) &= \frac{\partial}{\partial s} \hat{G}(-\hat{\lambda}, p^{(k)}), \quad \nabla_p G(-\hat{\lambda}, p^{(k)}) = \nabla_p \hat{G}(-\hat{\lambda}, p^{(k)}). \end{aligned}$$

Proof:

$$\|G\|_{\mathcal{H}_2 \times L_2(\Omega)} = \|L^T \tilde{G} L\|_{\mathcal{H}_2}, \quad \text{where } \tilde{G}(s) = \begin{bmatrix} C_0 \\ C_1 \end{bmatrix} (sE - A)^{-1} [\, B_0, B_1 \,],\ L = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} \end{bmatrix}.$$

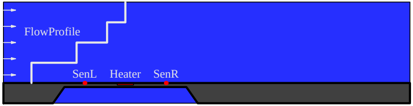
$\Longrightarrow$ **Computation via IRKA applied to $\tilde{G}$.**

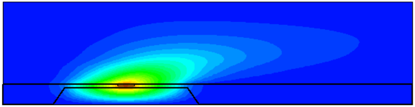
- Model for evolution of temperature distribution on a plate, described by the heat equation.
- FDM SISO model of order $n = 197$.
- Parameter $p_1 \in [0, 1]$ encodes movement of heat source from B_0 to $B_0 + B_1$, analogous for relocation of measurement.



Relative $\mathcal{H}_2 \times L_2(\Omega)$ error: 7.5×10^{-4}.

Consider an anemometer, a flow sensing device located on a membrane used in the context of minimizing heat dissipation.





Source: [Baur/B./Greiner/Korvink/Lienemann/Moosmann 2011]

- FE model:

$$E\dot{x}(t) = (A + pA_1)x(t) + Bu(t), \quad y(t) = Cx(t), \quad x(0) = 0,$$

- $n = 29{,}008$, $m = 1$, $q = 3$, $p_1 \in [0, 1]$ fluid velocity.

Source: MOR Wiki: http://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Anemometer

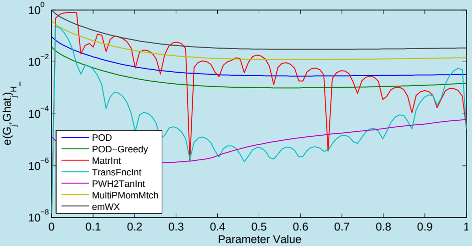
Consider an anemometer, a flow sensing device located on a membrane used in the context of minimizing heat dissipation.

- FE model:

$$E\dot{x}(t) = (A + pA_1)x(t) + Bu(t), \quad y(t) = Cx(t), \quad x(0) = 0,$$

- $n = 29,008$, $m = 1$, $q = 3$, $p_1 \in [0, 1]$ fluid velocity.

$\mathcal{H}_\infty$ error



Source: MOR Wiki: http://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Anemometer

A Comparison of PMOR Methods: Anemometer

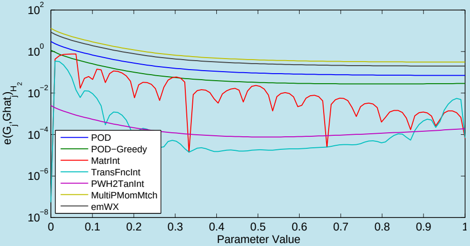
Consider an anemometer, a flow sensing device located on a membrane used in the context of minimizing heat dissipation.

- FE model:

$$E\dot{x}(t) = (A + pA_1)x(t) + Bu(t), \quad y(t) = Cx(t), \quad x(0) = 0,$$

- $n = 29{,}008,\ m = 1,\ q = 3,\ p_1 \in [0, 1]$ fluid velocity.

$\mathcal{H}_2$ error

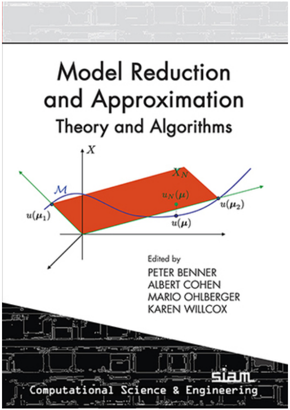


Source: MOR Wiki: http://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Anemometer

For more details of this comparisons, and other tests, see

U. Baur, P. Benner, B. Haasdonk, C. Himpe, I. Maier, and M. Ohlberger.
Comparison of Methods for Parametric Model Order Reduction of Unsteady Problems.
In P. Benner, A. Cohen, M. Ohlberger, and K. Willcox (eds.), *Model Reduction and Approximation: Theory and Algorithms.*
SIAM, Philadelphia, PA, 2017.

Chapter 9 in



Consider bilinear control systems:

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

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

Consider bilinear control systems:

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

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

Key Observation **[B./Breiten 2011]**

Consider parameters as additional inputs, a linear parametric system

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{m_p} a_i(p) A_i x(t) + B_0 u_0(t), \quad y(t) = Cx(t)$$

with $B_0 \in \mathbb{R}^{n\times m_0}$ can be interpreted as bilinear system:

$$u(t) := \begin{bmatrix} a_1(p) & \dots & a_{m_p}(p) & u_0(t) \end{bmatrix}^T,$$

$$B := \begin{bmatrix} 0 & \dots & 0 & B_0 \end{bmatrix} \in \mathbb{R}^{n\times m}, \quad m = m_p + m_0.$$

Linear parametric systems can be interpreted as bilinear systems.

Linear parametric systems can be interpreted as bilinear systems.

Consequence

Model order reduction techniques for bilinear systems can be applied to linear parametric systems!

Here:

- Balanced truncation,
- $\mathcal{H}_2$ optimal model reduction.

Idea (for simplicity, $E = I_n$)

- $\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t), \end{cases}$ with A stable, i.e., $\Lambda(A) \subset \mathbb{C}^-$,

  is balanced, if system Gramians, i.e., solutions P, Q of the Lyapunov equations

  $$AP + PA^T + BB^T = 0, \qquad A^T Q + QA + C^T C = 0,$$

  satisfy: $P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.

Idea (for simplicity, $E = I_n$)

- $\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t), \end{cases}$ with A stable, i.e., $\Lambda(A) \subset \mathbb{C}^-$,
  is balanced, if system Gramians, i.e., solutions P, Q of the Lyapunov equations
  $$AP + PA^T + BB^T = 0, \qquad A^TQ + QA + C^TC = 0,$$
  satisfy: $P = Q = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.
- $\{\sigma_1, \ldots, \sigma_n\}$ are the Hankel singular values (HSVs) of Σ.

Idea (for simplicity, $E = I_n$)

- $\Sigma: \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t), \end{cases}$ with A stable, i.e., $\Lambda(A) \subset \mathbb{C}^-$,
  is balanced, if system Gramians, i.e., solutions P, Q of the Lyapunov equations
  $$AP + PA^T + BB^T = 0, \qquad A^TQ + QA + C^TC = 0,$$
  satisfy: $P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.
- $\{\sigma_1, \ldots, \sigma_n\}$ are the Hankel singular values (HSVs) of Σ.
- Compute balanced realization (needs P, Q!) of the system via state-space transformation

$$\begin{aligned} \mathcal{T} : (A, B, C) &\mapsto (TAT^{-1}, TB, CT^{-1}) \\ &= \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \begin{bmatrix} C_1 & C_2 \end{bmatrix} \right). \end{aligned}$$

Idea (for simplicity, $E = I_n$)

- Σ : $\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t), \end{cases}$ with A stable, i.e., $\Lambda(A) \subset \mathbb{C}^-$,

  is balanced, if system Gramians, i.e., solutions P, Q of the Lyapunov equations

  $$AP + PA^T + BB^T = 0, \qquad A^TQ + QA + C^TC = 0,$$

  satisfy: $P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.
- $\{\sigma_1, \ldots, \sigma_n\}$ are the Hankel singular values (HSVs) of Σ.
- Compute balanced realization (needs P, Q!) of the system via state-space transformation

  $$\begin{aligned} \mathcal{T} : (A, B, C) &\mapsto (TAT^{-1}, TB, CT^{-1}) \\ &= \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \begin{bmatrix} C_1 & C_2 \end{bmatrix} \right). \end{aligned}$$
- Truncation $\rightsquigarrow (\hat{A}, \hat{B}, \hat{C}) = (A_{11}, B_1, C_1)$.

Balanced Truncation for Bilinear Systems

The concept of balanced truncation can be generalized to the case of bilinear systems, where we need the solutions of the generalized Lyapunov equations:

$$AP + PA^T + \sum_{i=1}^{m} A_i P A_i^T + BB^T = 0,$$
$$A^T Q + QA^T + \sum_{i=1}^{m} A_i^T Q A_i + C^T C = 0.$$

- These equations also appear for stochastic control systems, see [B./Damm 2011].
- "Twice-the-trail-of-the-HSVs" error bound does not hold [B./Damm 2014].
- Stability preservation [B./Damm/Redmann/Rodriguez Cruz 2016].

$$AX + XA^T + \sum_{i=1}^{m} A_i X A_i^T + BB^T = 0. \qquad (1)$$

- Need a positive semi-definite symmetric solution X.

$$AX + XA^T + \sum_{i=1}^{m} A_i X A_i^T + BB^T = 0. \tag{1}$$

- Need a positive semi-definite symmetric solution X.
- In standard Lyapunov case, existence and uniqueness guaranteed if A stable $(\Lambda(A) \subset \mathbb{C}^-)$; this is not sufficient here: (1) is equivalent to

$$\left(I_n \otimes A + A \otimes I_n + \sum_{i=1}^{m} A_i \otimes A_i \right) \mathrm{vec}(X) = -\mathrm{vec}(BB^T).$$

A sufficient condition for stable A is smallness of A_i (related to stability radius of $\mathcal{A}$)

$\rightsquigarrow$ **bounded-input bounded-output (BIBO) stability** of bilinear systems.

This will be assumed from here on, hence: existence and uniqueness of positive semi-definite solution $X = X^T$.

$$AX + XA^T + \sum_{i=1}^{m} A_i X A_i^T + BB^T = 0. \tag{1}$$

- Need a positive semi-definite symmetric solution X.
- In standard Lyapunov case, existence and uniqueness guaranteed if A stable ($\Lambda(A) \subset \mathbb{C}^-$); this is not sufficient here: (1) is equivalent to

$$\left(I_n \otimes A + A \otimes I_n + \sum_{i=1}^{m} A_i \otimes A_i \right) \operatorname{vec}(X) = -\operatorname{vec}(BB^T).$$

  A sufficient condition for stable A is smallness of A_i (related to stability radius of $\mathcal{A}$)
  $\rightsquigarrow$ **bounded-input bounded-output (BIBO) stability** of bilinear systems.
  This will be assumed from here on, hence: existence and uniqueness of positive semi-definite solution $X = X^T$.
- Want: solution methods for large scale problems, i.e., only matrix-matrix multiplication with A, A_j, solves with (shifted) A allowed!

$$AX + XA^T + \sum_{i=1}^{m} A_i X A_i^T + BB^T = 0. \tag{1}$$

- Need a positive semi-definite symmetric solution X.
- In standard Lyapunov case, existence and uniqueness guaranteed if A stable ($\Lambda(A) \subset \mathbb{C}^-$); this is not sufficient here: (1) is equivalent to

$$\left(I_n \otimes A + A \otimes I_n + \sum_{i=1}^{m} A_i \otimes A_i \right) \mathrm{vec}(X) = -\mathrm{vec}(BB^T).$$

  A sufficient condition for stable A is smallness of A_i (related to stability radius of $\mathcal{A}$)
  $\rightsquigarrow$ **bounded-input bounded-output (BIBO) stability** of bilinear systems.
  This will be assumed from here on, hence: existence and uniqueness of positive semi-definite solution $X = X^T$.
- Want: solution methods for large scale problems, i.e., only matrix-matrix multiplication with A, A_j, solves with (shifted) A allowed!
- Requires to compute data-sparse approximation to generally dense X; here: $X \approx ZZ^T$ with $Z \in \mathbb{R}^{n \times n_Z}$, $n_Z \ll n$!

Q: Can we expect low-rank approximations $ZZ^T \approx X$ to the solution of

$$AX + XA^T + \sum_{j=1}^{m} A_j X A_j^T + BB^T = 0 \text{ ?}$$

Q: Can we expect low-rank approximations $ZZ^T \approx X$ to the solution of

$$AX + XA^T + \sum_{j=1}^{m} A_j X A_j^T + BB^T = 0 \text{ ?}$$

Theorem **[B./Breiten 2013]**

Assume existence and uniqueness assumption with stable A and $A_j = U_j V_j^T$, with $U_j, V_j \in \mathbb{R}^{n \times r_j}$. Set $r = \sum_{j=1}^{m} r_j$.
Then the solution X of

$$AX + XA^T + \sum_{j=1}^{m} A_j X A_j^T + BB^T = 0$$

can be approximated by X_k of rank $(2k+1)(m+r)$, with an error satisfying

$$\|X - X_k\|_2 \lesssim \exp(-\sqrt{k}).$$

- Generalized Alternating Directions Iteration (ADI) method.
  1. Computing square solution matrix ($\sim n^2$ unknowns) [Damm 2008].
  2. Computing low-rank factors of solutions ($\sim n$ unknowns) [B./Breiten 2013].
- Generalized Extended (or rational) Krylov Subspace Method (EKSM) [B./Breiten 2013].
- Tensorized versions of standard Krylov subspace methods, e.g., PCG, PBiCGStab [Kressner/Tobler 2011, B./Breiten 2013].
- Inexact stationary (fix point) iteration [Shank/Simoncini/Szyld 2016].

$$
\begin{aligned}
E\dot{x}(t) &= (A + p_1(t)A_1 + p_2(t)A_2)x(t) + B,\\
y(t) &= Cx(t), \quad x(0) = x_0 \neq 0,
\end{aligned}
$$

- Rewrite as system with zero initial condition,
- FE model: $n = 16,912$, $m = 3$, $q = 1$,
- $p_j \in [0, 10^9]$ time-varying voltage functions,
- transfer function $G(i\omega, p_1, p_2)$,
- reduced system dimension $r = 67$,
- $\max\limits_{\substack{\omega\in\{\omega_{min},\dots,\omega_{max}\}\\ p_j\in\{p_{min},\dots,p_{max}\}}} \frac{\|G-\hat{G}\|_2}{\|G\|_2} < 6 \cdot 10^{-4}$,
- evaluation times: FOM 4.5h, ROM 38s
  $\rightsquigarrow$ speed-up factor ≈ 426.

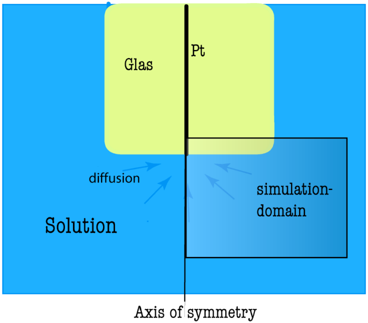
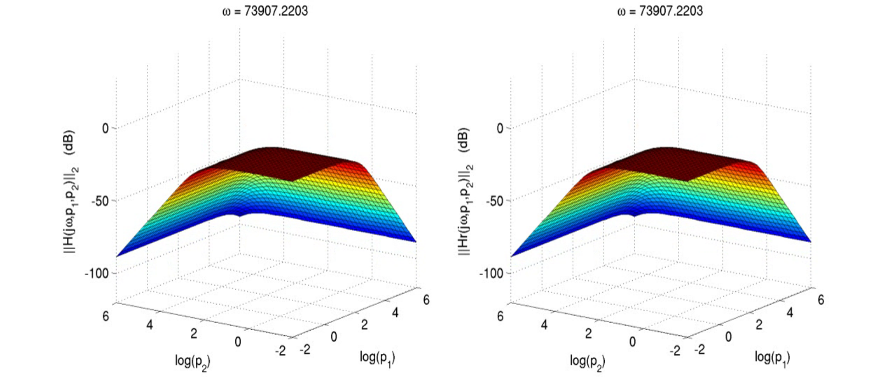
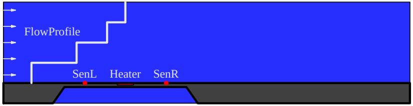


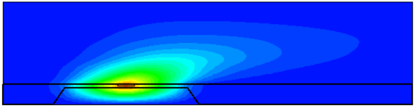
Figure: [FENG ET AL. 2006]

Original. . . and reduced-order model.



Consider again the anemometer example.





$$E\dot{x}(t) = (A + pA_1)x(t) + Bu(t), \quad y(t) = Cx(t), \quad x(0) = 0,$$

- FE model: $n = 29,008$, $m = 1$, $q = 3$,
- $p_1 \in [0,1]$ fluid velocity,
- transfer function $G(i\omega, p_1)$, reduced system dimension $r = 146$,
- $\max\limits_{\substack{\omega \in \{\omega_{min},\dots,\omega_{max}\} \\ p_1 \in \{p_{min},\dots,p_{max}\}}} \frac{\|G(\omega,p) - \hat{G}(\omega,p)\|_2}{\|G(\omega,p)\|_2} < 3 \cdot 10^{-5}$,
- evaluation times: FOM 51min, ROM 21sec.

Consider bilinear system ($m = 1$, i.e. SISO)

$$\Sigma : \left\{ \dot{x}(t) = Ax(t) + A_1 x(t)u(t) + Bu(t), \quad y(t) = Cx(t). \right.$$

Consider bilinear system ($m = 1$, i.e. SISO)

$$\Sigma : \left\{ \dot{x}(t) = Ax(t) + A_1 x(t)u(t) + Bu(t), \quad y(t) = Cx(t). \right.$$

Output Characterization (SISO) via Volterra series:

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

with kernels $K(t_1, \ldots, t_k) = Ce^{At_k} A_1 \cdots e^{At_2} A_1 e^{At_1} B$.

Consider bilinear system ($m = 1$, i.e. SISO)

$$\Sigma : \left\{ \dot{x}(t) = Ax(t) + A_1 x(t)u(t) + Bu(t), \quad y(t) = Cx(t). \right.$$

Output Characterization (SISO) via Volterra series:

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

with kernels $K(t_1, \ldots, t_k) = Ce^{At_k} A_1 \cdots e^{At_2} A_1 e^{At_1} B$.

Multivariate Laplace-transform:

$$G_k(s_1, \ldots, s_k) = C(s_k I - A)^{-1} A_1 \cdots (s_2 I - A)^{-1} A_1 (s_1 I - A)^{-1} B.$$

Consider bilinear system ($m = 1$, i.e. SISO)

$$\Sigma: \ \left\{ \dot{x}(t) = Ax(t) + A_1 x(t)u(t) + Bu(t), \quad y(t) = Cx(t). \right.$$

Output Characterization (SISO) via Volterra series:

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

with kernels $K(t_1, \ldots, t_k) = Ce^{At_k} A_1 \cdots e^{At_2} A_1 e^{At_1} B$.

Multivariate Laplace-transform:

$$G_k(s_1, \ldots, s_k) = C(s_k I - A)^{-1} A_1 \cdots (s_2 I - A)^{-1} A_1 (s_1 I - A)^{-1} B.$$

Bilinear $\mathcal{H}_2$-norm: [Zhang/Lam 2002]

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

Lemma **[B./Breiten 2012]**

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

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

Lemma [B./Breiten 2012]

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

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

Error System

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

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

Theorem (error system squared)

Let Σ and $\hat{\Sigma}$ be the original and reduced bilinear systems, respectively. Then

$$
\begin{aligned}
\mathcal{E} &:= \|\Sigma^{err}\|_{\mathcal{H}_2}^2 := \|\Sigma - \hat{\Sigma}\|_{\mathcal{H}_2}^2 \\
&= (\mathcal{I}_p)^T \left(\begin{bmatrix} C & -\tilde{C} \end{bmatrix} \otimes \begin{bmatrix} C & -\hat{C} \end{bmatrix} \right) \times \\
&\quad \left(- \begin{bmatrix} A & 0 \\ 0 & \Lambda \end{bmatrix} \otimes I_{n+r} - I_{n+r} \otimes \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix} - \sum_{k=1}^{m} \begin{bmatrix} A_k & 0 \\ 0 & \tilde{A}_k^T \end{bmatrix} \otimes \begin{bmatrix} A_k & 0 \\ 0 & \hat{A}_k \end{bmatrix} \right)^{-1} \times \\
&\quad \left(\begin{bmatrix} B \\ \tilde{B}^T \end{bmatrix} \otimes \begin{bmatrix} B \\ \hat{B} \end{bmatrix} \right) \mathcal{I}_m, \\
&= (\mathcal{I}_p)^T \left(\begin{bmatrix} C & -\tilde{C} \end{bmatrix} \otimes \begin{bmatrix} C & -\tilde{C} \end{bmatrix} \right) \times \\
&\quad \left(- \begin{bmatrix} A & 0 \\ 0 & \Lambda \end{bmatrix} \otimes I_{n+r} - I_{n+r} \otimes \begin{bmatrix} A & 0 \\ 0 & \Lambda \end{bmatrix} - \sum_{k=1}^{m} \begin{bmatrix} A_k & 0 \\ 0 & \tilde{A}_k^T \end{bmatrix} \otimes \begin{bmatrix} A_k & 0 \\ 0 & \tilde{A}_k^T \end{bmatrix} \right)^{-1} \times \\
&\quad \left(\begin{bmatrix} B \\ \tilde{B}^T \end{bmatrix} \otimes \begin{bmatrix} B \\ \tilde{B}^T \end{bmatrix} \right) \mathcal{I}_m,
\end{aligned}
$$

where $\mathcal{I}_\ell = \operatorname{vec}(I_\ell)$, $R\Lambda R^{-1} = \hat{A}$ is the spectral decomposition of $\hat{A}$, and $\tilde{B} = \hat{B}^T R^{-T}$, $\tilde{C} = \hat{C}R$, $\tilde{A}_k = R^T \hat{A}_k^T R^{-T}$.

Assume $\hat{\Sigma}$ is given in coordinate system induced by eigendecomposition of $\hat{A}$:

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

Assume $\hat{\Sigma}$ is given in coordinate system induced by eigendecomposition of $\hat{A}$:

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

Using Λ, $\tilde{A}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g., $\frac{\partial \mathcal{E}}{\partial \tilde{C}_{ij}} = 0$ yields:

Assume $\hat{\Sigma}$ is given in coordinate system induced by eigendecomposition of $\hat{A}$:

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

Using Λ, $\tilde{A}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g., $\frac{\partial \mathcal{E}}{\partial \tilde{C}_{ij}} = 0$ yields:

$$\begin{aligned}
&(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_r \otimes A - \sum_{k=1}^{m} \tilde{A}_k \otimes A_k\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m) \\
&= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_n - I_r \otimes \hat{A} - \sum_{k=1}^{m} \tilde{A}_k \otimes \hat{A}_k\right)^{-1} \left(\tilde{B} \otimes \hat{B}\right) \mathrm{vec}(I_m).
\end{aligned}$$

Assume $\hat{\Sigma}$ is given in coordinate system induced by eigendecomposition of $\hat{A}$:

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

Using Λ, $\tilde{A}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g., $\frac{\partial \mathcal{E}}{\partial \tilde{C}_{ij}} = 0$ yields:

$$\begin{aligned}
&(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_r \otimes A - \sum_{k=1}^{m} \tilde{A}_k \otimes A_k\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m) \\
&= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_n - I_r \otimes \hat{A} - \sum_{k=1}^{m} \tilde{A}_k \otimes \hat{A}_k\right)^{-1} \left(\tilde{B} \otimes \hat{B}\right) \mathrm{vec}(I_m).
\end{aligned}$$

Connection to interpolation of transfer functions?

Assume $\hat{\Sigma}$ is given in coordinate system induced by eigendecomposition of $\hat{A}$:

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

Using Λ, $\tilde{A}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g., $\frac{\partial \mathcal{E}}{\partial \tilde{C}_{ij}} = 0$ yields:

$$\begin{aligned}
&(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_r \otimes A - \sum_{k=1}^{m} \tilde{A}_k \otimes A_k\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m) \\
&= (\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes \hat{C}\right) \left(-\Lambda \otimes I_n - I_r \otimes \hat{A} - \sum_{k=1}^{m} \tilde{A}_k \otimes \hat{A}_k\right)^{-1} \left(\tilde{B} \otimes \hat{B}\right) \mathrm{vec}(I_m).
\end{aligned}$$

For $A_k \equiv 0$, this is equivalent to

$$G(-\lambda_\ell)\tilde{B}_\ell^T = \hat{G}(-\lambda_\ell)\tilde{B}_\ell^T, \quad \ell = 1, \ldots, r.$$

$\rightsquigarrow$ Tangential interpolation at mirror images of reduced system poles!

Assume $\hat{\Sigma}$ is given in coordinate system induced by eigendecomposition of $\hat{A}$:

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

Using Λ, $\tilde{A}_i$, $\tilde{B}$, $\tilde{C}$ as optimization parameters, we can derive necessary conditions for $\mathcal{H}_2$-optimality, e.g., $\frac{\partial \mathcal{E}}{\partial \tilde{C}_{ij}} = 0$ yields:

$$(\mathrm{vec}(I_q))^T \left(e_j e_\ell^T \otimes C\right) \left(-\Lambda \otimes I_n - I_r \otimes A - \sum_{k=1}^{m} \tilde{A}_k \otimes A_k\right)^{-1} \left(\tilde{B} \otimes B\right) \mathrm{vec}(I_m)$$

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

For $A_k \equiv 0$, this is equivalent to

$$G(-\lambda_\ell)\tilde{B}_\ell^T = \hat{G}(-\lambda_\ell)\tilde{B}_\ell^T, \quad \ell = 1, \ldots, r.$$

$\rightsquigarrow$ Tangential interpolation at mirror images of reduced system poles!

Note: [Flagg 2011] shows equivalence to interpolating the Volterra series!

A First Iterative Approach — BIRKA

Algorithm 3 Bilinear IRKA

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

Output: $A^{opt},\ A_i^{opt},\ B^{opt},\ C^{opt}$

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

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

$$\operatorname{vec}(V) = \left(-\Lambda \otimes I_n - I_r \otimes A - \sum_{i=1}^{m} \tilde{A}_i \otimes A_i\right)^{-1} \left(\tilde{B} \otimes B\right) \operatorname{vec}(I_m)$$

$$\operatorname{vec}(W) = \left(-\Lambda \otimes I_n - I_r \otimes A^T - \sum_{i=1}^{m} \tilde{A}_i^T \otimes A_i^T\right)^{-1} \left(\tilde{C}^T \otimes C^T\right) \operatorname{vec}(I_q)$$

$\quad V = \operatorname{orth}(V),\ W = \operatorname{orth}(W)$

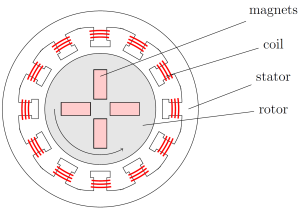
$\quad \hat{A} = \left(W^T V\right)^{-1} W^T A V,\ \hat{A}_i = \left(W^T V\right)^{-1} W^T A_i V,$

$\quad \hat{B} = \left(W^T V\right)^{-1} W^T B,\ \hat{C} = CV$

end while

$A^{opt} = \hat{A},\ A_i^{opt} = \hat{A}_i,\ B^{opt} = \hat{B},\ C^{opt} = \hat{C}$

- Thermal simulations to detect whether temperature changes lead to fatigue or deterioration of employed materials.
- Main heat source: thermal losses resulting from current stator coil/rotor.
- Many different current profiles need to be considered to predict whether temperature on certain parts of the motor remains in feasible region.
- Finite element analysis on rather complicated geometries $\rightsquigarrow$ large-scale linear models with 7/13 parameters.

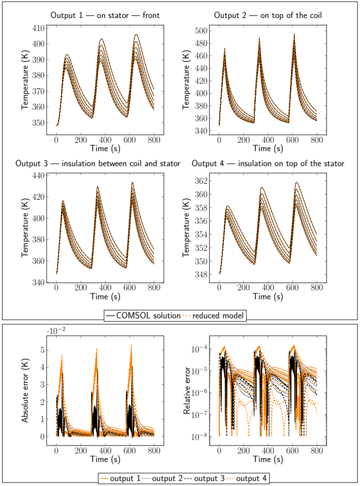


Schematic view of an electrical motor.

Bosch integrated motor generator used in hybrid variants of Porsche Cayenne, VW Touareg.

Pictures: BOSCH

- FEM analysis of thermal model $\rightsquigarrow$ linear parametric systems with $n = 41,199$, $m = 4$ inputs, and $d = 13$ parameters;
- measurements taken at $q = 4$ heat sensors;
- time for 1 transient simulation in COMSOL® $\sim$ 90min;
- ROM order $\hat{n} = 300$, time for 1 transient simulation $\sim$ 15sec.
- Legend: Temperature curves for six different values (5, 25, 45, 65, 85, 100$[W/m^2K]$) of the heat transfer coefficient on the coil.



Conclusions and Outlook

- We have reviewed some of the most popular PMOR methods developed in the last decade, in particular those based on rational interpolation.

  Open problem in general: optimal interpolation points.

Conclusions and Outlook

- We have reviewed some of the most popular PMOR methods developed in the last decade, in particular those based on rational interpolation.

  Open problem in general: optimal interpolation points.
- We have established a connection between special linear parametric and bilinear systems that automatically yields structure-preserving model reduction techniques for linear parametric systems.

Conclusions and Outlook

- We have reviewed some of the most popular PMOR methods developed in the last decade, in particular those based on rational interpolation.

  Open problem in general: optimal interpolation points.
- We have established a connection between special linear parametric and bilinear systems that automatically yields structure-preserving model reduction techniques for linear parametric systems.
- PMOR via $\mathcal{H}_2$ optimal model reduction for bilinear systems:
  - Yields competitive approach, proven in industrial context.
  - Still high offline cost (= time for generating reduced-order model).
  - May need to switch to one-sided projection ($W = V$) to preserve stability.

Conclusions and Outlook

- We have reviewed some of the most popular PMOR methods developed in the last decade, in particular those based on rational interpolation.

  Open problem in general: optimal interpolation points.
- We have established a connection between special linear parametric and bilinear systems that automatically yields structure-preserving model reduction techniques for linear parametric systems.
- PMOR via $\mathcal{H}_2$ optimal model reduction for bilinear systems:
  - Yields competitive approach, proven in industrial context.
  - Still high offline cost (= time for generating reduced-order model).
  - May need to switch to one-sided projection ($W = V$) to preserve stability.
- PMOR via balanced truncation for bilinear systems:
  - Requires solution of certain generalized Lyapunov equations.
  - Novel numerical algorithms developed, but efficiency needs improvement.

Conclusions and Outlook

- We have reviewed some of the most popular PMOR methods developed in the last decade, in particular those based on rational interpolation.

  Open problem in general: optimal interpolation points.
- We have established a connection between special linear parametric and bilinear systems that automatically yields structure-preserving model reduction techniques for linear parametric systems.
- PMOR via $\mathcal{H}_2$ optimal model reduction for bilinear systems:
  - Yields competitive approach, proven in industrial context.
  - Still high offline cost (= time for generating reduced-order model).
  - May need to switch to one-sided projection ($W = V$) to preserve stability.
- PMOR via balanced truncation for bilinear systems:
  - Requires solution of certain generalized Lyapunov equations.
  - Novel numerical algorithms developed, but efficiency needs improvement.
- Several extensions to nonlinear systems, but just starting.

Conclusions and Outlook

- We have reviewed some of the most popular PMOR methods developed in the last decade, in particular those based on rational interpolation.

  Open problem in general: optimal interpolation points.
- We have established a connection between special linear parametric and bilinear systems that automatically yields structure-preserving model reduction techniques for linear parametric systems.
- PMOR via $\mathcal{H}_2$ optimal model reduction for bilinear systems:
  - Yields competitive approach, proven in industrial context.
  - Still high offline cost (= time for generating reduced-order model).
  - May need to switch to one-sided projection ($W = V$) to preserve stability.
- PMOR via balanced truncation for bilinear systems:
  - Requires solution of certain generalized Lyapunov equations.
  - Novel numerical algorithms developed, but efficiency needs improvement.
- Several extensions to nonlinear systems, but just starting.
- New direction: data-enhanced approaches, merging ideas from Loewner framework with model-based methods.

Conclusions and Outlook

- We have reviewed some of the most popular PMOR methods developed in the last decade, in particular those based on rational interpolation.

  Open problem in general: optimal interpolation points.
- We have established a connection between special linear parametric and bilinear systems that automatically yields structure-preserving model reduction techniques for linear parametric systems.
- PMOR via $\mathcal{H}_2$ optimal model reduction for bilinear systems:
  - Yields competitive approach, proven in industrial context.
  - Still high offline cost (= time for generating reduced-order model).
  - May need to switch to one-sided projection ($W = V$) to preserve stability.
- PMOR via balanced truncation for bilinear systems:
  - Requires solution of certain generalized Lyapunov equations.
  - Novel numerical algorithms developed, but efficiency needs improvement.
- Several extensions to nonlinear systems, but just starting.
- New direction: data-enhanced approaches, merging ideas from Loewner framework with model-based methods.
- Most of the methods can be used to significantly accelerate UQ by Monte Carlo or Stochastic Collocation methods!

N. Banagaaya, P. Benner, L. Feng, P. Meuris, and W. Schoenmaker.
An index-aware parametric model order reduction method for parametrized quadratic differential-algebraic equations. *Applied Mathematics and Computation*, 319:409–424, 2018.

U. Baur, C. Beattie, P. Benner, and S. Gugercin.
Interpolatory projection methods for parameterized model reduction.
SIAM Journal on Scientific Computing 33(5):2489–2518, 2011.

U. Baur and P. Benner.
Model reduction for parametric systems using balanced truncation and interpolation.
at-Automatisierungstechnik 57(8):411–419, 2009.

U. Baur, P. Benner, A. Greiner, J.G. Korvink, J. Lienemann, and C. Moosmann.
Parameter preserving model order reduction for mems applications.
Mathematical and Computer Modelling of Dynamical Systems 17(4):297–317, 2011.

P. Benner and T. Breiten.
On $\mathcal{H}_2$ model reduction of linear parameter-varying systems.
Proceedings in Applied Mathematics and Mechanics 11:805–806, 2011.

P. Benner and T. Breiten.
Interpolation-based $\mathcal{H}_2$-model reduction of bilinear control systems.
SIAM Journal on Matrix Analysis and Applications 33(3):859–885, 2012.

P. Benner and T. Breiten.
Low rank methods for a class of generalized Lyapunov equations and related issues.
Numerische Mathematik, 124(3):441–470, 2013.

P. Benner and A. Bruns.
Parametric model order reduction of thermal models using the bilinear interpolatory rational Krylov algorithm. *Mathematical and Computer Modelling of Dynamical Systems*, 21(2):103–129, 2015.

References

P. Benner, A. Cohen, M. Ohlberger, and K. Willcox (eds.).
Model Reduction and Approximation: Theory and Algorithms.
SIAM, Philadelphia, PA, 2017.

P. Benner and T. Damm.
Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems.
SIAM Journal on Control and Optimization 49(2):686–711, 2011.

P. Benner and L. Feng.
A robust algorithm for parametric model order reduction based on implicit moment matching.
In A. Quarteroni and G. Rozza (eds.), Reduced Order Methods for Modeling and Computational Reduction, MS&A – Modeling, Simulation and Applications, Vol. 9, pp. 159–185, Springer International Publishing, 2014.

P. Benner, S. Gugercin, and K. Willcox.
A survey of model reduction methods for parametric systems.
SIAM Review 57(4):483–531, 2015.

L. Feng, A.C. Antoulas, and P. Benner.
Some a posteriori error bounds for reduced order modelling of (non-)parametrized linear systems.
ESAIM: Mathematical Modelling and Numerical Analysis, 51(6):2127–2158, 2017.

L. Feng and P. Benner.
A new error estimator for reduced-order modeling of linear parametric systems.
IEEE Transactions on Microwave Theory and Techniques, 67(12):4848–4859, 2019.

L. Feng, Y. Yue, N. Banagaaya, P. Meuris, W. Schoenmaker, and P. Benner.
Parametric modeling and model order reduction for (electro-)thermal analysis of nanoelectronic structures. *Journal of Mathematics in Industry* 6.10 (16pp.), 2016.