MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY

On the solution of systems of matrix integral equations in parametric model order reduction

Manuela Hund, Tim Mitchell, Petar Mlinarić, <u>Jens Saak</u>

ICIAM 2019, Valencia

MS *Recent advances in matrix equations with applications*

July 16$^{\text{th}}$, 2019

Supported by:

DFG pyMOR (SA 3477/1–1)

MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

CSC COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY

On the efficient evaluation of systems of matrix integral equations in parametric model order reduction

Manuela Hund, Tim Mitchell, Petar Mlinarić, Jens Saak

ICIAM 2019, Valencia

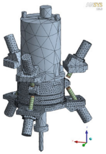
MS *Recent advances in matrix equations with applications*

July 16th, 2019

Supported by:

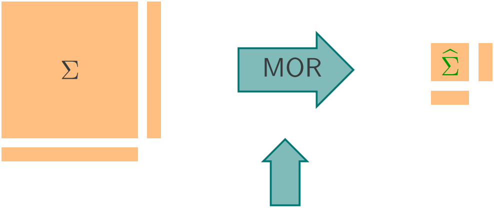
DFG pyMOR
(SA 3477/1–1)


pyMOR School
October 7-11, 2019
MPI Magdeburg
https://school.pymor.org



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

Model Order Reduction (MOR)



$\mathcal{H}_2$-Optimality Conditions

Wilson conditions [Wilson '70]

$$
\begin{aligned}
0 &= \widetilde{Q}^\mathsf{T} A \widetilde{P} + \widehat{Q}^\mathsf{T} \widehat{A} \widehat{P},\\
0 &= \widetilde{Q}^\mathsf{T} E \widetilde{P} + \widehat{Q}^\mathsf{T} \widehat{E} \widehat{P},\\
0 &= \widetilde{Q}^\mathsf{T} B + \widehat{Q}^\mathsf{T} \widehat{B},\\
0 &= C \widetilde{P} - \widehat{C} \widehat{P}.
\end{aligned}
$$

first-order necessary $\mathcal{H}_2$-optimality conditions

Sylvester

$$A \widetilde{P} \widehat{E}^\mathsf{T} + E \widetilde{P} \widehat{A}^\mathsf{T} = -B \widehat{B}^\mathsf{T}$$

$$A^\mathsf{T} \widetilde{Q} \widehat{E} + E^\mathsf{T} \widetilde{Q} \widehat{A} = -C^\mathsf{T} \widehat{C}$$

Lyapunov

$$\widehat{A} \widehat{P} \widehat{E}^\mathsf{T} + \widehat{E} \widehat{P} \widehat{A}^\mathsf{T} = -\widehat{B} \widehat{B}^\mathsf{T}$$

$$\widehat{A}^\mathsf{T} \widehat{Q} \widehat{E} + \widehat{E}^\mathsf{T} \widehat{Q} \widehat{A} = -\widehat{C}^\mathsf{T} \widehat{C}$$

IRKA: Iterative Rational Krylov Algorithm [Gugercin/Antoulas/Beattie '08]

- Given admissible shifts, compute rational subspaces.
- Solve reduced eigenvalue problem and use mirrored eigenvalues as new shifts.
- Repeat until shifts converge.

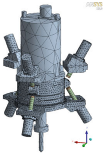
- Good initialization and general convergence proof open problems.
- Property preservation not guaranteed (e.g. asymptotic stability).
- Soft breakdowns (e.g. inadmissible shifts, cycling through local minimums) need to be treated.

TSIA: Two-Sided Iteration Algorithm [Xu/Zeng '11]

- Solve Wilson conditions via fixed-point iteration.

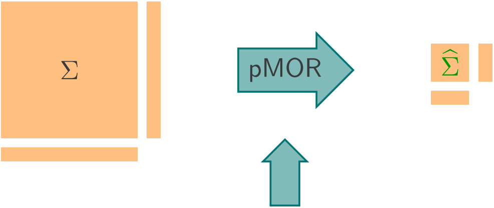
- initial model/subspace?
- possibly slow convergence
- needs careful implementation for numerical stability [Benner/Köhler/S. '11]

Parametric



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

Model Order Reduction (MOR)



$\mathcal{H}_2 \otimes \mathcal{L}_2$-Optimality Conditions

Outline

Notation
Parametric Setting

- Parametric linear time-invariant system:

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

of dimension n with parameter $\mu \in M \subset \mathbb{R}^d$.

pMOR

- Reduced-order model (ROM):

$$\widehat{\Sigma} \begin{cases} \widehat{E}(\mu)\dot{\widehat{x}}(t) = \widehat{A}(\mu)\widehat{x}(t) + \widehat{B}(\mu)u(t), \\ \widehat{y}(t) = \widehat{C}(\mu)\widehat{x}(t), \end{cases}$$

of dimension $r \ll n$ with parameter $\mu \in M \subset \mathbb{R}^d$.

Full-order model (FOM) **"uniformly nice"**

- $\forall \mu \in M$: $E(\mu)$ invertible.
- $\forall \mu \in M$: FOM asymptotically stable.

Full-order model (FOM) — "uniformly nice"

- $\forall \mu \in M$: $E(\mu)$ invertible.
- $\forall \mu \in M$: FOM asymptotically stable.

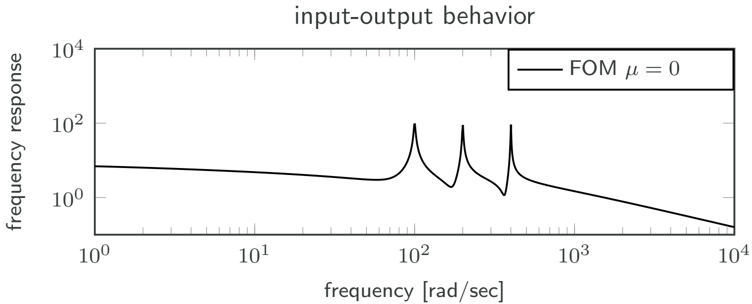
ROM — "uniformly nice and easy to handle"

- $\forall \mu \in M$: ROM asymptotically stable.
- Affine decompositions

$$\widehat{E}(\mu) = \sum_{i=1}^{q_{\widehat{E}}} \widehat{e}_i(\mu) \widehat{E}_i, \qquad \widehat{A}(\mu) = \sum_{j=1}^{q_{\widehat{A}}} \widehat{a}_j(\mu) \widehat{A}_j,$$

$$\widehat{B}(\mu) = \sum_{k=1}^{q_{\widehat{B}}} \widehat{b}_k(\mu) \widehat{B}_k, \qquad \widehat{C}(\mu) = \sum_{\ell=1}^{q_{\widehat{C}}} \widehat{c}_\ell(\mu) \widehat{C}_\ell,$$

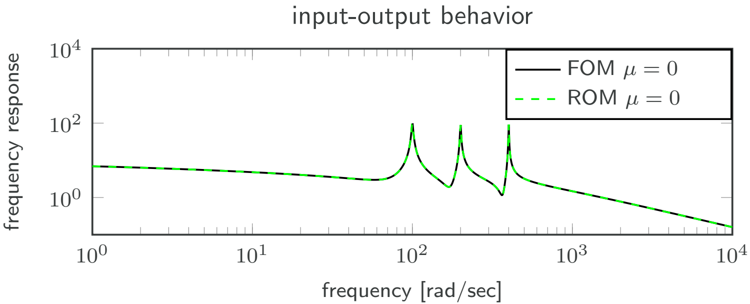
with continuous functions $\widehat{e}_i, \widehat{a}_j, \widehat{b}_k, \widehat{c}_\ell \colon M \to \mathbb{R}$.



- Transfer function for $s \in \mathbb{C}, \mu \in M$:

$$H(s,\mu) = C(\mu)(sE(\mu) - A(\mu))^{-1}B(\mu), \qquad (\Sigma(\mu, E, A, B, C))$$

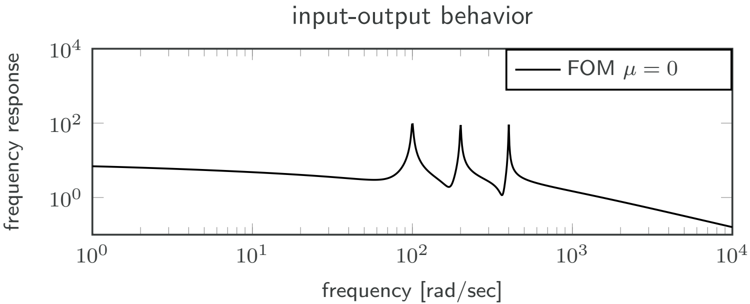
$$\widehat{H}(s,\mu) = \widehat{C}(\mu)(s\widehat{E}(\mu) - \widehat{A}(\mu))^{-1}\widehat{B}(\mu), \qquad (\widehat{\Sigma}(\mu, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C}))$$



- Transfer function for $s \in \mathbb{C}, \mu \in M$:

$$H(s,\mu) = C(\mu)(sE(\mu) - A(\mu))^{-1}B(\mu), \qquad (\Sigma(\mu, E, A, B, C))$$

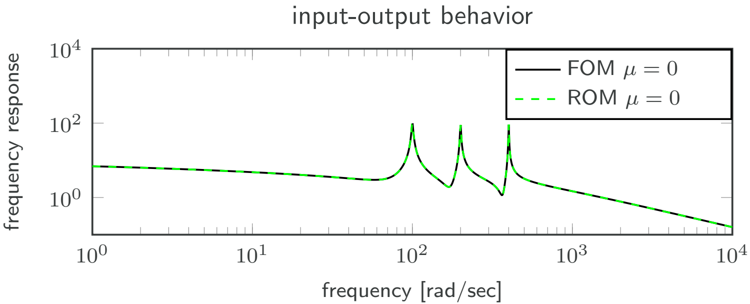
$$\widehat{H}(s,\mu) = \widehat{C}(\mu)(s\widehat{E}(\mu) - \widehat{A}(\mu))^{-1}\widehat{B}(\mu), \qquad (\widehat{\Sigma}(\mu, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C}))$$



- Transfer function for $s \in \mathbb{C}, \mu \in M$:

$$H(s,\mu) = C(\mu)(sE(\mu) - A(\mu))^{-1}B(\mu), \qquad (\Sigma(\mu, E, A, B, C))$$

$$\widehat{H}(s,\mu) = \widehat{C}(\mu)(s\widehat{E}(\mu) - \widehat{A}(\mu))^{-1}\widehat{B}(\mu), \qquad (\widehat{\Sigma}(\mu, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C}))$$



- Transfer function for $s \in \mathbb{C}, \mu \in M$:

$$H(s,\mu) = C(\mu)(sE(\mu) - A(\mu))^{-1}B(\mu), \qquad (\Sigma(\mu, E, A, B, C))$$

$$\widehat{H}(s,\mu) = \widehat{C}(\mu)(s\widehat{E}(\mu) - \widehat{A}(\mu))^{-1}\widehat{B}(\mu). \qquad (\widehat{\Sigma}(\mu, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C}))$$

$$H_e(s,\mu) = H(s,\mu) - \widehat{H}(s,\mu). \qquad (\Sigma_e(\mu, E_e, A_e, B_e, C_e))$$

Notation

The Error System $\Sigma_e(\mu, E_e, A_e, B_e, C_e)$ in Time-Domain

$$\Sigma_e \begin{cases} E_e(\mu)\dot{x}_e(t) = A_e(\mu)x_e(t) + B_e(\mu)u(t), \\ y_e(t) = C_e(\mu)x_e(t). \end{cases}$$

$$A_e(\mu) = \begin{bmatrix} A(\mu) & 0 \\ 0 & \widehat{A}(\mu) \end{bmatrix}, \qquad E_e(\mu) = \begin{bmatrix} E(\mu) & 0 \\ 0 & \widehat{E}(\mu) \end{bmatrix},$$

$$B_e(\mu) = \begin{bmatrix} B(\mu) \\ \widehat{B}(\mu) \end{bmatrix}, \qquad C_e(\mu) = \begin{bmatrix} C(\mu) & -\widehat{C}(\mu) \end{bmatrix}.$$

Definition: $\mathcal{H}_2 \otimes \mathcal{L}_2$-norm [Baur/Beattie/Benner/Gugercin '11]

Given a transfer function $H(s,\mu)$ for a stable system Σ with parameter $\mu \in M$, the **$\mathcal{H}_2 \otimes \mathcal{L}_2$-norm** is defined as

$$\|H\|^2_{\mathcal{H}_2\otimes\mathcal{L}_2} := \int\limits_M \|H(\cdot,\mu)\|^2_{\mathcal{H}_2}\,\mathrm{d}\mu.$$

Definition: $\mathcal{H}_2 \otimes \mathcal{L}_2$-norm [Baur/Beattie/Benner/Gugercin '11]

Given a transfer function $H(s, \mu)$ for a stable system Σ with parameter $\mu \in M$, the **$\mathcal{H}_2 \otimes \mathcal{L}_2$-norm** is defined as

$$\begin{aligned} \|H\|^2_{\mathcal{H}_2 \otimes \mathcal{L}_2} &:= \int\limits_M \|H(\cdot, \mu)\|^2_{\mathcal{H}_2} \,\mathrm{d}\mu, \\ &= \int\limits_M \operatorname{tr}\left(C(\mu)P(\mu)C(\mu)^\mathsf{T}\right) \mathrm{d}\mu, \\ &= \int\limits_M \operatorname{tr}\left(B(\mu)^\mathsf{T}Q(\mu)B(\mu)\right) \mathrm{d}\mu. \end{aligned}$$

- For $\mu \in M$, $P(\mu)$ and $Q(\mu)$ are the solutions of

$$\begin{aligned} 0 &= A(\mu)P(\mu)E(\mu)^\mathsf{T} + E(\mu)P(\mu)A(\mu)^\mathsf{T} + B(\mu)B(\mu)^\mathsf{T}, \\ 0 &= A(\mu)^\mathsf{T}Q(\mu)E(\mu) + E(\mu)^\mathsf{T}Q(\mu)A(\mu) + C(\mu)^\mathsf{T}C(\mu). \end{aligned}$$

Definition: $\mathcal{H}_2 \otimes \mathcal{L}_2$-norm [Baur/Beattie/Benner/Gugercin '11]

Given a transfer function $H(s,\mu)$ for a stable system Σ with parameter $\mu \in M$, the **$\mathcal{H}_2 \otimes \mathcal{L}_2$-norm** is defined as

$$\begin{aligned}\|H\|^2_{\mathcal{H}_2\otimes\mathcal{L}_2} &:= \int_M \|H(\cdot,\mu)\|^2_{\mathcal{H}_2}\,\mathrm{d}\mu.\\ &= \int_M \mathrm{tr}\left(C(\mu)P(\mu)C(\mu)^\mathsf{T}\right)\mathrm{d}\mu,\\ &= \int_M \mathrm{tr}\left(B(\mu)^\mathsf{T}Q(\mu)B(\mu)\right)\mathrm{d}\mu.\end{aligned}$$

- For $\mu \in M$, $P(\mu)$ and $Q(\mu)$ are the solutions of

$$\begin{aligned}P(\mu) &= \mathsf{lyap}(\mu; A, E, B),\\ Q(\mu) &= \mathsf{lyap}(\mu; A^\mathsf{T}, E^\mathsf{T}, C^\mathsf{T}).\end{aligned}$$

Optimization problem

$$\text{Minimize } \|\Sigma_e(\mu, E_e, A_e, B_e, C_e)\|^2_{\mathcal{H}_2 \otimes \mathcal{L}_2},$$

with respect to all reduced-order matrices $\widehat{E}_i, \widehat{A}_j, \widehat{B}_k, \widehat{C}_\ell$.

Optimization problem

$$\text{Minimize } \|\Sigma_e(\mu, E_e, A_e, B_e, C_e)\|^2_{\mathcal{H}_2 \otimes \mathcal{L}_2},$$

$$\Rightarrow \text{Minimize } J(\widehat{\Sigma}) := \int_M \operatorname{tr}\left(C_e(\mu) P_e(\mu) C_e(\mu)^\mathsf{T}\right) \mathrm{d}\mu,$$

with respect to all reduced-order matrices $\widehat{E}_i, \widehat{A}_j, \widehat{B}_k, \widehat{C}_\ell$.

Optimization problem

$$\text{Minimize } \|\Sigma_e(\mu, E_e, A_e, B_e, C_e)\|^2_{\mathcal{H}_2 \otimes \mathcal{L}_2},$$

$$\Rightarrow \text{Minimize } J(\widehat{\Sigma}) := \int_M \operatorname{tr}\left(C_e(\mu) P_e(\mu) C_e(\mu)^\mathsf{T}\right) \mathrm{d}\mu,$$

with respect to all reduced-order matrices $\widehat{E}_i, \widehat{A}_j, \widehat{B}_k, \widehat{C}_\ell$.

- Solutions of Lyapunov equations:

$$P_e(\mu) = \mathsf{lyap}(\mu; A_e, E_e, B_e),$$

$$Q_e(\mu) = \mathsf{lyap}(\mu; A_e^\mathsf{T}, E_e^\mathsf{T}, C_e^\mathsf{T}).$$

Optimization problem

$$\text{Minimize } \left\|\Sigma_e(\mu, E_e, A_e, B_e, C_e)\right\|^2_{\mathcal{H}_2\otimes\mathcal{L}_2},$$

$$\Rightarrow \text{Minimize } J(\widehat{\Sigma}) := \int\limits_M \operatorname{tr}\left(C_e(\mu)P_e(\mu)C_e(\mu)^\mathsf{T}\right) \mathrm{d}\mu,$$

with respect to all reduced-order matrices $\widehat{E}_i, \widehat{A}_j, \widehat{B}_k, \widehat{C}_\ell$.

- Solutions of Lyapunov equations:

$$\begin{bmatrix} P(\mu) & \widetilde{P}(\mu) \\ \widetilde{P}(\mu)^\mathsf{T} & \widehat{P}(\mu) \end{bmatrix} = \mathsf{lyap}(\mu; A_e, E_e, B_e),$$

$$\begin{bmatrix} Q(\mu) & \widetilde{Q}(\mu) \\ \widetilde{Q}(\mu)^\mathsf{T} & \widehat{Q}(\mu) \end{bmatrix} = \mathsf{lyap}(\mu; A_e^\mathsf{T}, E_e^\mathsf{T}, C_e^\mathsf{T}).$$

$\mathcal{H}_2 \otimes \mathcal{L}_2$-gradient

$$\nabla_{\widehat{E}_i} J(\widehat{\Sigma}) = 2 \int_M \widehat{e}_i(\mu) \left(\widetilde{Q}(\mu)^\mathsf{T} A(\mu) \widetilde{P}(\mu) + \widehat{Q}(\mu)^\mathsf{T} \widehat{A}(\mu) \widehat{P}(\mu) \right) \mathrm{d}\mu, \quad i \in [q_{\widehat{E}}],$$

$$\nabla_{\widehat{A}_j} J(\widehat{\Sigma}) = 2 \int_M \widehat{a}_j(\mu) \left(\widetilde{Q}(\mu)^\mathsf{T} E(\mu) \widetilde{P}(\mu) + \widehat{Q}(\mu)^\mathsf{T} \widehat{E}(\mu) \widehat{P}(\mu) \right) \mathrm{d}\mu, \quad j \in [q_{\widehat{A}}],$$

$$\nabla_{\widehat{B}_k} J(\widehat{\Sigma}) = 2 \int_M \widehat{b}_k(\mu) \left(\widetilde{Q}(\mu)^\mathsf{T} B(\mu) + \widehat{Q}(\mu)^\mathsf{T} \widehat{B}(\mu) \right) \mathrm{d}\mu, \quad k \in [q_{\widehat{B}}],$$

$$\nabla_{\widehat{C}_\ell} J(\widehat{\Sigma}) = 2 \int_M \widehat{c}_\ell(\mu) \left(\widehat{C}(\mu) \widehat{P}(\mu) - C(\mu) \widetilde{P}(\mu) \right) \mathrm{d}\mu, \quad \ell \in [q_{\widehat{C}}].$$

$\mathcal{H}_2 \otimes \mathcal{L}_2$-necessary optimality conditions

$$0 = \int_M \widehat{e}_i(\mu) \left(\widetilde{Q}(\mu)^\mathsf{T} A(\mu) \widetilde{P}(\mu) + \widehat{Q}(\mu)^\mathsf{T} \widehat{A}(\mu) \widehat{P}(\mu) \right) \mathrm{d}\mu, \quad i \in [q_{\widehat{E}}],$$

$$0 = \int_M \widehat{a}_j(\mu) \left(\widetilde{Q}(\mu)^\mathsf{T} E(\mu) \widetilde{P}(\mu) + \widehat{Q}(\mu)^\mathsf{T} \widehat{E}(\mu) \widehat{P}(\mu) \right) \mathrm{d}\mu, \quad j \in [q_{\widehat{A}}],$$

$$0 = \int_M \widehat{b}_k(\mu) \left(\widetilde{Q}(\mu)^\mathsf{T} B(\mu) + \widehat{Q}(\mu)^\mathsf{T} \widehat{B}(\mu) \right) \mathrm{d}\mu, \quad k \in [q_{\widehat{B}}],$$

$$0 = \int_M \widehat{c}_\ell(\mu) \left(\widehat{C}(\mu) \widehat{P}(\mu) - C(\mu) \widetilde{P}(\mu) \right) \mathrm{d}\mu, \quad \ell \in [q_{\widehat{C}}].$$

- Optimization problem:

$$\min_{x \in \mathbb{R}^N} f(x),$$

for a differentiable function $f\colon \mathbb{R}^N \to \mathbb{R}$.

- Here: x is the vectorization of all reduced-order matrices, not the state!
- Given an initial value x_0, the new iterate is obtained by

$$x_{k+1} = x_k + \alpha_k p_k,$$

for a step length $\alpha_k > 0$, a search direction $p_k \in \mathbb{R}^N$ and $k = 0, 1, \ldots$.

- Optimization problem:

$$\min_{x \in \mathbb{R}^N} f(x),$$

  for a differentiable function $f\colon \mathbb{R}^N \to \mathbb{R}$.

- Here: x is the vectorization of all reduced-order matrices, not the state!
- Given an initial value x_0, the new iterate is obtained by

$$x_{k+1} = x_k + \alpha_k p_k,$$

  for a step length $\alpha_k > 0$, a search direction $p_k \in \mathbb{R}^N$ and $k = 0, 1, \ldots$.

  - Gradient descent: $p_k = -\nabla f(x_k)$.
  - Newton: $p_k = -H_k^{-1} \nabla f(x_k)$, where $H_k = \nabla^2 f(x_k)$.
  - Quasi-Newton: $p_k = -\widetilde{H}_k^{-1} \nabla f(x_k)$, where $\widetilde{H}_k$ is an approximation of $\nabla^2 f(x_k)$ using $\nabla f(x_0), \ldots, \nabla f(x_k)$.

- Optimization problem:

$$\min_{x \in \mathbb{R}^N} f(x),$$

for a differentiable function $f\colon \mathbb{R}^N \to \mathbb{R}$.

- Here: x is the vectorization of all reduced-order matrices, not the state!
- Given an initial value x_0, the new iterate is obtained by

$$x_{k+1} = x_k + \alpha_k p_k,$$

for a step length $\alpha_k > 0$, a search direction $p_k \in \mathbb{R}^N$ and $k = 0, 1, \ldots$.

  - Gradient descent: $p_k = -\nabla f(x_k)$.
  - Newton: $p_k = -H_k^{-1} \nabla f(x_k)$, where $H_k = \nabla^2 f(x_k)$.
  - Quasi-Newton: $p_k = -\widetilde{H}_k^{-1} \nabla f(x_k)$, where $\widetilde{H}_k$ is an approximation of $\nabla^2 f(x_k)$ using $\nabla f(x_0), \ldots, \nabla f(x_k)$. (here `fminunc`, GRANSO[1])

[1] `http://www.timmitchell.com/software/GRANSO/`

- **Surrogate objective function:**

$$\begin{aligned}\|\Sigma_e\|^2_{\mathcal{H}_2\otimes\mathcal{L}_2} &= \int\limits_M \operatorname{tr}\left(C_e(\mu)P_e(\mu)C_e(\mu)^{\mathsf{T}}\right)\mathrm{d}\mu \\ &= \int\limits_M \operatorname{tr}\left(C(\mu)P(\mu)C(\mu)^{\mathsf{T}}\right)\mathrm{d}\mu + \int\limits_M \operatorname{tr}\left(\widehat{C}(\mu)\widehat{P}(\mu)\widehat{C}(\mu)^{\mathsf{T}}\right)\mathrm{d}\mu \\ &\quad - 2\int\limits_M \operatorname{tr}\left(C(\mu)\widetilde{P}(\mu)\widehat{C}(\mu)^{\mathsf{T}}\right)\mathrm{d}\mu\end{aligned}$$

- **Surrogate objective function:**

$$\|\Sigma_e\|^2_{\mathcal{H}_2\otimes\mathcal{L}_2} = \int\limits_M \operatorname{tr}\left(C_e(\mu)P_e(\mu)C_e(\mu)^\mathsf{T}\right)\mathrm{d}\mu$$

$$= \underbrace{\|\Sigma\|^2_{\mathcal{H}_2\otimes\mathcal{L}_2}}_{\text{constant}} + \left\|\widehat{\Sigma}\right\|^2_{\mathcal{H}_2\otimes\mathcal{L}_2} - 2\int\limits_M \operatorname{tr}\left(C(\mu)\widetilde{P}(\mu)\widehat{C}(\mu)^\mathsf{T}\right)\mathrm{d}\mu$$

- **Surrogate objective function:**

$$f(x) := \left\| \widehat{\Sigma} \right\|^2_{\mathcal{H}_2 \otimes \mathcal{L}_2} - 2 \int\limits_{M} \operatorname{tr}\left(C(\mu) \widetilde{P}(\mu) \widehat{C}(\mu)^{\mathsf{T}} \right) \mathrm{d}\mu$$

Large-sparse structured Lyapunov equation vs. small Lyapunov and skinny Sylvester equations.

- **Surrogate objective function:**

$$f(x) := \left\| \widehat{\Sigma} \right\|_{\mathcal{H}_2 \otimes \mathcal{L}_2}^2 - 2 \int_M \operatorname{tr}\left(C(\mu) \widetilde{P}(\mu) \widehat{C}(\mu)^\mathsf{T} \right) \mathrm{d}\mu$$

  Large-sparse structured Lyapunov equation vs. small Lyapunov and skinny Sylvester equations.

- **Gradient computation**
  - Uses adaptive Gauß-Kronrod quadrature (`quadgk`+`arrayfun`).
  - Requires (many) non-parametric small Lyapunov and skinny Sylvester solutions.
  - Caches μ-evaluations to avoid re-solving matrix equations.

- **Surrogate objective function:**

$$f(x) := \left\|\widehat{\Sigma}\right\|^2_{\mathcal{H}_2\otimes\mathcal{L}_2} - 2\int\limits_M \operatorname{tr}\left(C(\mu)\widetilde{P}(\mu)\widehat{C}(\mu)^\mathsf{T}\right)\mathrm{d}\mu$$

  Large-sparse structured Lyapunov equation vs. small Lyapunov and skinny Sylvester equations.

- **Gradient computation**
  - Uses adaptive Gauß-Kronrod quadrature (`quadgk`+`arrayfun`).
  - Requires (many) non-parametric small Lyapunov and skinny Sylvester solutions.
  - Caches μ-evaluations to avoid re-solving matrix equations.
- **Preservation of asymptotic stability**
  - $\widehat{\Sigma}$ unstable for any $\mu \Rightarrow \left\|\widehat{\Sigma}\right\|_{\mathcal{H}_2\otimes\mathcal{L}_2} = \infty$.
  - Use a `chebfun` for largest real part of $\Lambda(\widehat{A}(\mu), \widehat{E}(\mu))$ for fast evaluation.

- Piecewise $\mathcal{H}_2$-optimal interpolatory parametric MOR.

- Piecewise $\mathcal{H}_2$-optimal interpolatory parametric MOR.
- Assumption: FOM has affine decomposable matrices.

$$\Sigma(\mu, E, A, B, C)$$

- Piecewise $\mathcal{H}_2$-optimal interpolatory parametric MOR.
- Assumption: FOM has affine decomposable matrices.

$$\Sigma(\mu, E, A, B, C)$$

$\mu = \mu_1$ ↙ ↘ $\mu = \mu_{p_s}$

$$\Sigma(\mu_1, E, A, B, C) \quad \cdots \quad \Sigma(\mu_{p_s}, E, A, B, C)$$

p_s parameter samples

- Piecewise $\mathcal{H}_2$-optimal interpolatory parametric MOR.
- Assumption: FOM has affine decomposable matrices.

$$\Sigma(\mu, E, A, B, C)$$

$\mu = \mu_1$ ↙ ↘ $\mu = \mu_{p_s}$ — p_s parameter samples

$$\Sigma(\mu_1, E, A, B, C) \quad \cdots \quad \Sigma(\mu_{p_s}, E, A, B, C)$$

$\downarrow$ $\downarrow$ — $\mathcal{H}_2$-optimal MOR using IRKA

$$\widehat{\Sigma}(\mu_1, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C}) \quad \cdots \quad \widehat{\Sigma}(\mu_{p_s}, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C})$$

- Piecewise $\mathcal{H}_2$-optimal interpolatory parametric MOR.
- Assumption: FOM has affine decomposable matrices.

$$\Sigma(\mu, E, A, B, C)$$

$\mu = \mu_1$ $\qquad$ $\mu = \mu_{p_s}$ $\qquad$ p_s parameter samples

$$\Sigma(\mu_1, E, A, B, C) \quad \cdots \quad \Sigma(\mu_{p_s}, E, A, B, C)$$

$\downarrow \qquad \downarrow$ $\qquad$ $\mathcal{H}_2$-optimal MOR using IRKA

$$\widehat{\Sigma}(\mu_1, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C}) \quad \cdots \quad \widehat{\Sigma}(\mu_{p_s}, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C})$$

$\downarrow \qquad \downarrow$ $\qquad$ projection matrices

$$V_1 \in \mathbb{R}^{n \times r_s} \quad \cdots \quad V_{p_s} \in \mathbb{R}^{n \times r_s}$$

Implementation

Initialization: pIRKA-like

[Baur/Beattie/Benner/Gugercin '11]

- Piecewise $\mathcal{H}_2$-optimal interpolatory parametric MOR.
- Assumption: FOM has affine decomposable matrices.

$$\Sigma(\mu, E, A, B, C)$$

$\mu = \mu_1$ ↙ … ↘ $\mu = \mu_{p_s}$ — p_s parameter samples

$$\Sigma(\mu_1, E, A, B, C) \quad \cdots \quad \Sigma(\mu_{p_s}, E, A, B, C)$$

$\downarrow$ $\mathcal{H}_2$-optimal MOR using IRKA $\downarrow$

$$\widehat{\Sigma}(\mu_1, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C}) \quad \cdots \quad \widehat{\Sigma}(\mu_{p_s}, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C})$$

$\downarrow$ projection matrices $\downarrow$

$$V_1 \in \mathbb{R}^{n \times r_s} \quad \cdots \quad V_{p_s} \in \mathbb{R}^{n \times r_s}$$

↘ concatenate ↙

$$V = \begin{bmatrix} V_1 & \dots & V_{p_s} \end{bmatrix} \in \mathbb{R}^{n \times r}$$

Implementation

Initialization: pIRKA-like

[Baur/Beattie/Benner/Gugercin '11]

- Piecewise $\mathcal{H}_2$-optimal interpolatory parametric MOR.
- Assumption: FOM has affine decomposable matrices.

$$\Sigma(\mu, E, A, B, C)$$

$\mu = \mu_1$ $\swarrow$ $\searrow$ $\mu = \mu_{p_s}$ — p_s parameter samples

$$\Sigma(\mu_1, E, A, B, C) \quad \cdots \quad \Sigma(\mu_{p_s}, E, A, B, C)$$

$\downarrow$ $\mathcal{H}_2$-optimal MOR using IRKA $\downarrow$

$$\widehat{\Sigma}(\mu_1, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C}) \quad \cdots \quad \widehat{\Sigma}(\mu_{p_s}, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C})$$

$\downarrow$ projection matrices $\downarrow$

$$V_1 \in \mathbb{R}^{n \times r_s} \quad \cdots \quad V_{p_s} \in \mathbb{R}^{n \times r_s}$$

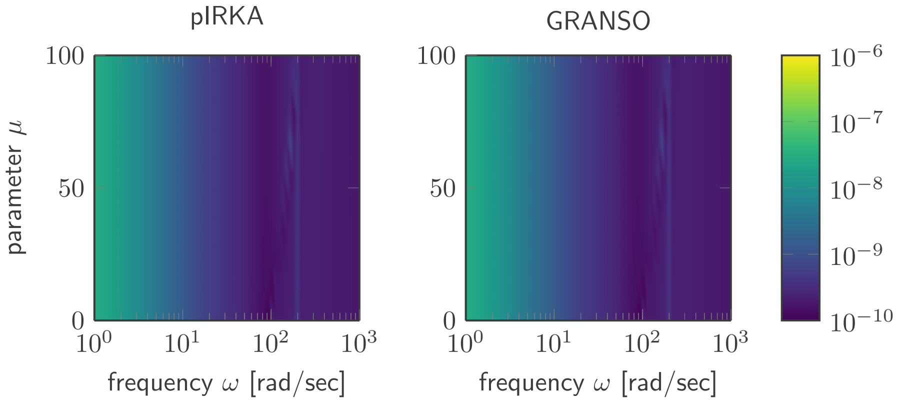
$\searrow$ concatenate $\swarrow$

$$V = \begin{bmatrix} V_1 & \ldots & V_{p_s} \end{bmatrix} \in \mathbb{R}^{n \times r}$$

$\downarrow$ project

$$\widehat{\Sigma}(\mu, \widehat{E}, \widehat{A}, \widehat{B}, \widehat{C})$$

- Pointwise relative frequency response errors:

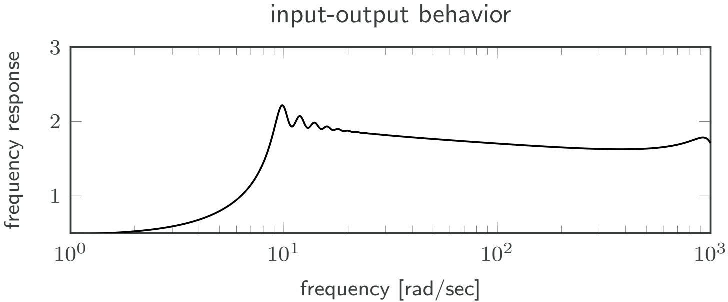


ROM dimension:	parameter samples:	sample dimension:
12	3	6

- FOM:

$$E\dot{x}(t) = (A_1 + \mu A_2)\, x(t) + Bu(t),$$
$$y(t) = Cx(t),$$

of dimension 1000 and parameter $\mu \in [0.1, 1]$.

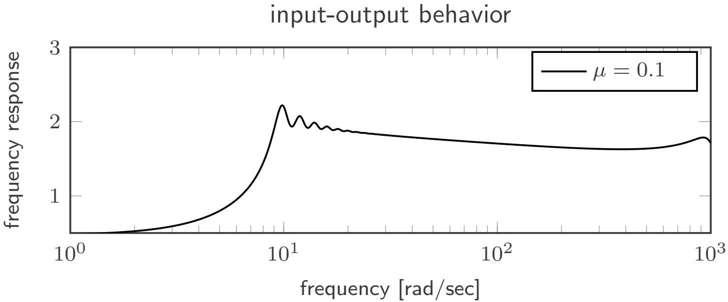


[1] http://modelreduction.org

- FOM:

$$E\dot{x}(t) = (A_1 + \mu A_2)\, x(t) + Bu(t),$$
$$y(t) = Cx(t),$$

of dimension 1000 and parameter $\mu \in [0.1, 1]$.



[1] http://modelreduction.org

- FOM:

$$
\begin{aligned}
E\dot{x}(t) &= (A_1 + \mu A_2)\, x(t) + Bu(t),\\
y(t) &= Cx(t),
\end{aligned}
$$

of dimension 1000 and parameter $\mu \in [0.1, 1]$.

MOR

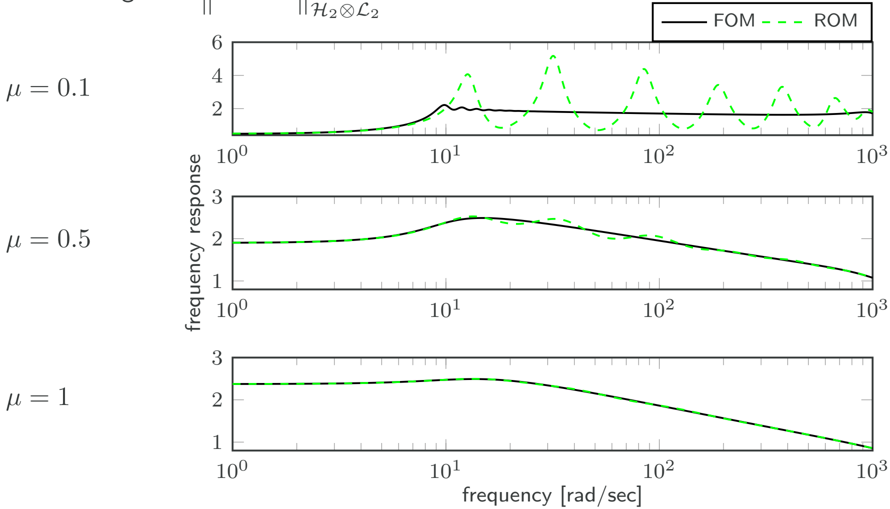
- ROM:

$$
\begin{aligned}
\widehat{E}\dot{\widehat{x}}(t) &= \left(\widehat{A}_1 + \mu \widehat{A}_2\right)\widehat{x}(t) + \widehat{B}u(t),\\
\widehat{y}(t) &= \widehat{C}\widehat{x}(t),
\end{aligned}
$$

of dimension 16 $(p_s = 4, r_s = 4) \Rightarrow N = 800$ optimization variables.

[1] http://modelreduction.org

- Initial guess: $\left\| H - \widehat{H} \right\|^2_{\mathcal{H}_2 \otimes \mathcal{L}_2} = 1.7 \cdot 10^1.$

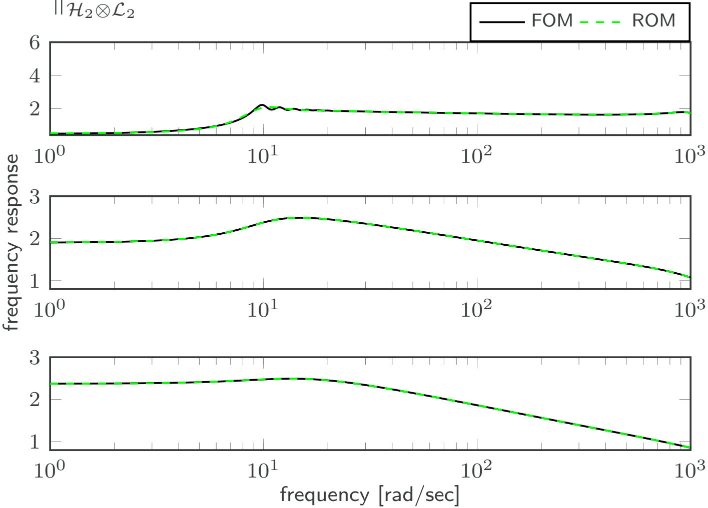


- Result: $\left\| H - \widehat{H} \right\|^2_{\mathcal{H}_2 \otimes \mathcal{L}_2} = 1.7 \cdot 10^{-3}$.

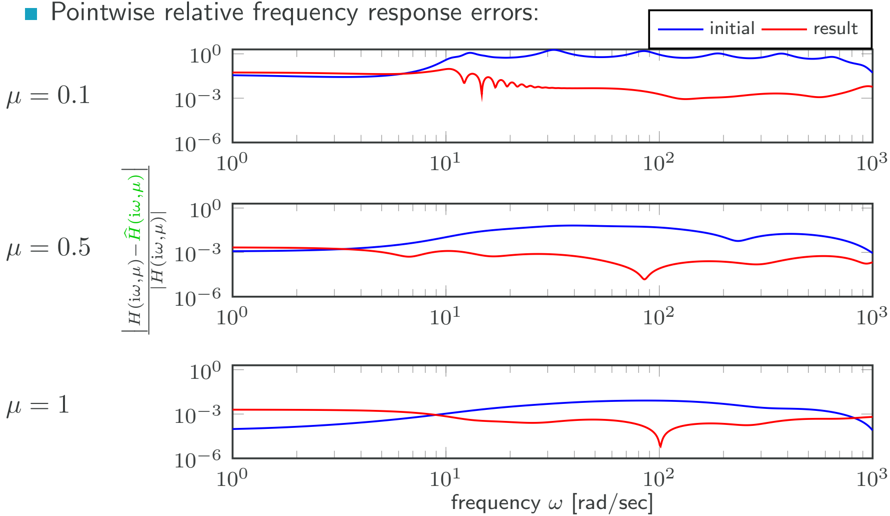
$\mu = 0.1$

$\mu = 0.5$

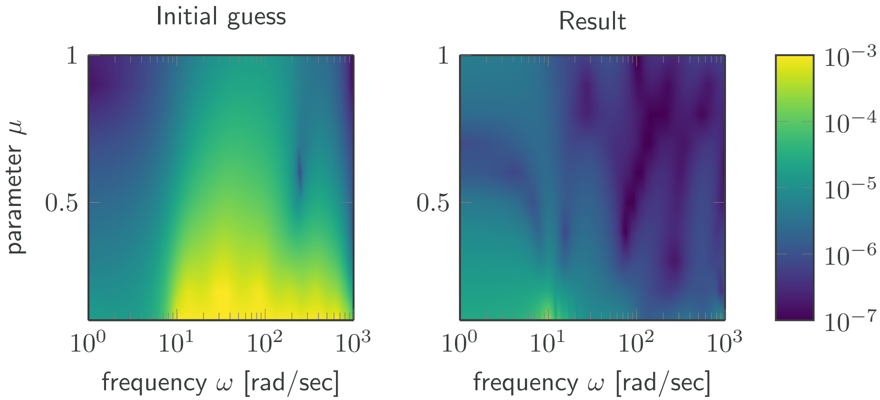
$\mu = 1$



- Pointwise relative frequency response errors:

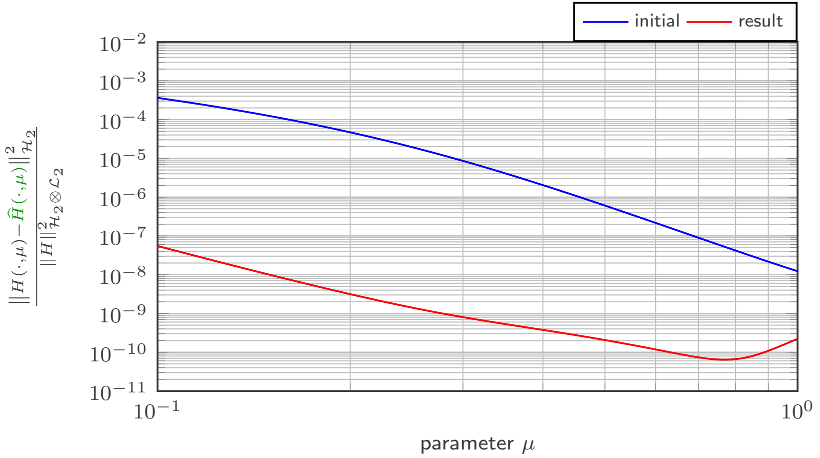


- Pointwise relative frequency response errors:



ROM dimension:	parameter samples:	sample dimension:
16	4	4

- Relative $\mathcal{H}_2$-error as a function of μ.



Conclusions and Outlook

What we have seen

- $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimality conditions for parametric systems.
- Derivation of a quasi-Newton-based algorithm.
- Implementation of the $\mathcal{H}_2 \otimes \mathcal{L}_2$-framework.

Conclusions and Outlook

What we have seen

- $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimality conditions for parametric systems.
- Derivation of a quasi-Newton-based algorithm.
- Implementation of the $\mathcal{H}_2 \otimes \mathcal{L}_2$-framework.

Further steps

- Extend implementation to $\mu \in M = \mathbb{R}^d$ for $d > 1$.
- Investigate different combinations of Hardy and Lebesgue norm ($\mathcal{H}_2 \otimes \mathcal{L}_\infty$, $\mathcal{H}_\infty \otimes \mathcal{L}_2$, $\mathcal{H}_\infty \otimes \mathcal{L}_\infty$).
- Tangential interpolation-based optimality conditions.

References

[1] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, vol. 6 of Adv. Des. Control, SIAM Publications, Philadelphia, PA, 2005, https://doi.org/10.1137/1.9780898718713.

[2] U. Baur, C. A. Beattie, P. Benner, and S. Gugercin, *Interpolatory projection methods for parameterized model reduction*, SIAM J. Sci. Comput., 33 (2011), pp. 2489–2518, https://doi.org/10.1137/090776925.

[3] Chebfun Developers, *Chebfun — numerical computing with functions*, https://www.chebfun.org/.

[4] F. E. Curtis, T. Mitchell, and M. L. Overton, *A BFGS-SQP method for nonsmooth, nonconvex, constrained optimization and its evaluation using relative minimization profiles*, Optim. Methods Softw., 32 (2017), pp. 148–181, https://doi.org/10.1080/10556788.2016.1208749.

[5] S. Gugercin, A. C. Antoulas, and C. Beattie, *$\mathcal{H}_2$ model reduction for large-scale linear dynamical systems*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 609–638, https://doi.org/10.1137/060666123.

[6] M. Hund, P. Mlinarić, and J. Saak, *An $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal model order reduction approach for parametric linear time-invariant systems*, Proc. Appl. Math. Mech., 18 (2018), p. e201800084, https://doi.org/10.1002/pamm.201800084.

[7] The MathWorks, Inc., *MATLAB*, http://www.matlab.com.

[8] T. Mitchell, *GRANSO: GRadient-based Algorithm for Non-Smooth Optimization*. http://timmitchell.com/software/GRANSO.
See also [4].

Thank you for your attention!