$H_2$ optimal model order reduction for parametric systems using RBF metamodels

Peter Benner, Sara Grundel, Nils Hornung
MPI Magdeburg and Fraunhofer SCAI
Abstract

Model Order Reduction Methods for linear systems are well studied and many successful methods exist. We will review some and explain more recent advances in Parametric Model Order Reduction. The focus will be on methods where we interpolate certain significant measures, that are computed for specific values of the parameter by Radial Basis Function Interpolation. These measures have a disadvantage as they behave like eigenvalues of matrices depending on parameters and we will explain how that can be dealt with in practice. We will furthermore need to introduce a technique to create a medium size model.
Abstract

Model Order Reduction Methods for linear systems are well studied and many successful methods exist. We will review some and explain more recent advances in Parametric Model Order Reduction. The focus will be on methods where we interpolate certain significant measures, that are computed for specific values of the parameter by Radial Basis Function Interpolation. These measures have a disadvantage as they behave like eigenvalues of matrices depending on parameters and we will explain how that can be dealt with in practice. We will furthermore need to introduce a technique to create a medium size model.
Abstract

Model Order Reduction Methods for linear systems are well studied and many successful methods exist. We will review some and explain more recent advances in Parametric Model Order Reduction. The focus will be on methods where we interpolate certain significant measures, that are computed for specific values of the parameter by Radial Basis Function Interpolation. These measures have a disadvantage as they behave like eigenvalues of matrices depending on parameters and we will explain how that can be dealt with in practice. We will furthermore need to introduce a technique to create a medium size model.
Outline

1. $\mathcal{H}_2$ MOR
2. Parametric MOR
3. Numerik
4. Medium Model
What is MOR?

\[ E \dot{x}(t) = A x(t) + B u(t); \]
\[ y(t) = C x(t) \]
\[ \dot{x}(t) = \hat{A} \hat{x}(t) + \hat{B} u(t); \]
\[ \hat{y}(t) = \hat{C} \hat{x}(t) \]
Projection-Based MOR

LTI System:

\[
\dot{x}(t) = Ax(t) + bu(t), \\
y(t) = c^T x(t), \quad x(0) = 0.
\]

Model Reduction Idea: Find \( W, V \in \mathbb{C}^{n \times r} \) with \( W^T V = I \) and \( x(t) \approx V \hat{x}(t) \), here \( r \ll n \)

\[
W^T V \dot{\hat{x}}(t) = W^T A V \hat{x}(t) + W^T b u(t) \\
\hat{y}(t) = c^T V \hat{x}(t).
\]

We define the transfer functions

\[
\hat{H}(s) = c^T (sI - \hat{A})^{-1} b \approx H(s) = c^T (sI - A)^{-1} b
\]

which is a rational function in \( s \) of degree \( r \) or \( n \).
Projection-Based MOR

LTI System:

\[
\dot{x}(t) = Ax(t) + bu(t), \\
y(t) = c^T x(t), \quad x(0) = 0.
\]

Model Reduction Idea: Find \( W, V \in \mathbb{C}^{n \times r} \) with \( W^T V = I \) and

\[
x(t) \approx V \hat{x}(t), \text{ here } r \ll n
\]

\[
\dot{\hat{x}}(t) = W^T A V \hat{x}(t) + W^T b u(t)
\]

\[
\hat{y}(t) = c^T V \hat{x}(t).
\]
Projection-Based MOR

LTI System:

\[
\dot{x}(t) = Ax(t) + bu(t), \\
y(t) = c^T x(t), \quad x(0) = 0.
\]

Model Reduction Idea: Find \(W, V \in \mathbb{C}^{n \times r}\) with \(W^T V = I\) and
\(x(t) \approx V \hat{x}(t)\), here \(r \ll n\)

\[
\dot{\hat{x}}(t) = \hat{A} \hat{x}(t) + \hat{b} u(t) \\
\hat{y}(t) = \hat{c}^T \hat{x}(t).
\]
Projection-Based MOR

LTI System:

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + bu(t), \\
y(t) &= c^T x(t), \\
x(0) &= 0.
\end{align*}
\]

Model Reduction Idea: Find \( W, V \in \mathbb{C}^{n \times r} \) with \( W^T V = I \) and \( x(t) \approx V \hat{x}(t) \), here \( r \ll n \)

\[
\begin{align*}
\dot{\hat{x}}(t) &= \hat{A} \hat{x}(t) + \hat{b} u(t) \\
\hat{y}(t) &= \hat{c}^T \hat{x}(t).
\end{align*}
\]

\[
\begin{align*}
\begin{cases}
\dot{x} = Ax + bu \\
y = c^T x
\end{cases} \quad \xrightarrow{\text{Lapl}} \quad \begin{cases}
sX = AX + bU \\
Y = c^T X
\end{cases} \quad \rightarrow \quad \begin{cases}
X = (sI - A)^{-1} bU \\
Y = c^T (sI - A)^{-1} bU
\end{cases}
\end{align*}
\]
Projection-Based MOR

LTI System:

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + bu(t), \\
y(t) &= c^T x(t), \quad x(0) = 0.
\end{align*}
\]

Model Reduction Idea: Find \( W, V \in \mathbb{C}^{n \times r} \) with \( W^T V = I \) and \( x(t) \approx V \hat{x}(t) \), here \( r \ll n \)

\[
\begin{align*}
\dot{\hat{x}}(t) &= \hat{A} \hat{x}(t) + \hat{b} u(t) \\
\hat{y}(t) &= \hat{c}^T \hat{x}(t).
\end{align*}
\]

\[
\begin{align*}
\{ \dot{x} = Ax + bu \} \quad \xrightarrow{\text{Lapl}} \quad \{ sX = AX + bU \} \quad \rightarrow \quad \{ X = (sl - A)^{-1}bU \} \\
\{ y = c^T x \} \quad \rightarrow \quad \{ Y = c^T X \} \quad \rightarrow \quad \{ Y = c^T (sl - A)^{-1}bU \}
\end{align*}
\]

We define the transfer functions

\[
\hat{H}(s) = \hat{c}^T (sl - \hat{A})^{-1}\hat{b} \approx H(s) = c^T (sl - A)^{-1}b
\]

which is a rational function in \( s \) of degree \( r \) or \( n \).
**$H_2$ Model Order Reduction**

Good Reduced Order Model

\[
\begin{align*}
\left\{ \begin{array}{ccc}
u & \Sigma & y \\
u & \hat{\Sigma} & \hat{y}
\end{array} \right\} & & \|y - \hat{y}\| \text{ small}
\end{align*}
\]

We know that:

\[
\sup_{t \geq 0} |y(t) - \hat{y}(t)| \leq \|H - \hat{H}\|_{H_2} \|u\|_{L_2}.
\]

for \(\|H - \hat{H}\|_{H_2} := \left( \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega) - \hat{H}(\omega)|^2 d\omega \right)^{1/2}\).

References

[Absil, Antoulas, Baur, Beattie, Benner, Breiten, Bunse-Gerstner, Gallivan, Gugercin, Kubalinska, Van Dooren, Vossen, Wilczek, ...]
**\( H_2 \) Model Order Reduction**

**How does it work**

We know that the optimal order \( r \) reduced transfer function \( \hat{H} \) hermite interpolates the true transfer function at the mirror poles \( \sigma_1, \ldots, \sigma_r \) of the reduced system.  

\[
H(\sigma_i) = \hat{H}(\sigma_i), \quad H'(\sigma_i) = \hat{H}'(\sigma_i)
\]

Given \( \sigma \) a rational function of degree \((r - 1, r)\) is uniquely defined.

\[
(\sigma \mathbf{I} - \mathbf{A})^{-1} \mathbf{b} \in Ran(\mathbf{V}) \\
(\overline{\sigma} \mathbf{I} - \mathbf{A}^T)^{-1} \mathbf{c} \in Ran(\mathbf{W}) \\
\Rightarrow H(\sigma) = \hat{H}(\sigma) \quad H'(\sigma) = \hat{H}'(\sigma)
\]

[GRIMME, YOUSOUFF, SKELETON].
**H₂ Model Order Reduction**

**How does it work**

We know that the optimal order $r$ reduced transfer function $\hat{H}$ hermite interpolates the true transfer function at the mirror poles $\sigma_1, \ldots, \sigma_r$ of the reduced system. [Meyer, Luenberger 1967]

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

Given $\sigma$ a rational function of degree $(r-1,r)$ is uniquely defined. [Grimme, Yousouff, Skeleton].

$$\begin{align*}
(\sigma I - A)^{-1}b &\in \text{Ran}(V) \\
(\overline{\sigma} I - A^T)^{-1}c &\in \text{Ran}(W) \\
\Rightarrow H(\sigma) = \hat{H}(\sigma) &\quad H'(\sigma) = \hat{H}'(\sigma)
\end{align*}$$
**Algorithm 1** Iterative rational Krylov algorithm (IRKA)

**Input:** Initial selection of interpolation points $\sigma_i$, closed under conjugation and a convergence tolerance $tol$.

**Output:** $\hat{A}$, $\hat{b}$, $\hat{c}$

1. Choose $V$ and $W$ s.t. $\text{range}(V) = \{(\sigma_1 I - A)^{-1}b, \ldots, (\sigma_r I - A)^{-1}b\}$ and $\text{range}(W) = \{(\sigma_1 I - A^T)^{-1}c, \ldots, (\sigma_r I - A^T)^{-1}c\}$ and $W^T V = I$.

2. **while** relative change in $\{\sigma_i\} > tol$ **do**
3. \[ \hat{A} = W^T AV, \]
4. assign $\sigma_i \leftarrow -\lambda_i(\hat{A})$ for $i = 1, \ldots, r$,
5. update $V$ and $W$ s.t. $\text{range}(V) = \{(\sigma_1 I - A)^{-1}b, \ldots, (\sigma_r I - A)^{-1}b\}$ and $\text{range}(W) = \{(\sigma_1 I - A^T)^{-1}c, \ldots, (\sigma_r I - A^T)^{-1}c\}$ and $W^T V = I$.
6. **end while**
7. $\hat{A} = W^T AV$, $\hat{b} = W^T b$, $\hat{c}^T = c^T V$
Outline

1. $\mathcal{H}_2$ MOR
2. Parametric MOR
3. Numerik
4. Medium Model
Parametrized Dynamical System

LTI System: \( p \in \mathcal{P} \subset \mathbb{R}^p \)

\[
\begin{align*}
\dot{x}(t) &= A(p)x(t) + b(p)u(t), \\
y(t) &= c(p)^T x(t), & x(0) = 0.
\end{align*}
\]

Model Reduction:

\[
\begin{align*}
\dot{\hat{x}}(t) &= \hat{A}(p)\hat{x}(t) + \hat{b}(p)u(t) \\
\hat{y}(t) &= \hat{c}(p)^T \hat{x}(t)
\end{align*}
\]

This means that the approximated transfer function

\[
\hat{H}(s, p) = \hat{c}(p)^T (sI - \hat{A}(p))^{-1} \hat{b}(p) \approx H(s, p) = c(p)^T (sI - A(p))^{-1} b(p)
\]

is a rational function in \( s \), but also a function in \( p \).
Previous Work

Reduced matrices from original matrices

\[ A(p) \rightarrow \hat{A}(p) \quad c(p) \rightarrow \hat{c}(p) \quad b(p) \rightarrow \hat{b}(p) \] (2)

Many attempts for parametric Model Order Reduction exist:

- projection matrix independent of parameter
  \[ \hat{A} = V^T A(p) W \]
  [Breiten, Damm, Baur, Benner, Beattie, Gugercin]

- matrix interpolation
  \( \hat{A}(p_i) \)
  [Panzer et al] or [Amsallam, Farhat]

- transfer function interpolation
  \( H(s_i, p_j) \)
  [Antoulas, Ionita]
Knowing $\sigma_1(p), \ldots, \sigma_r(p)$ seems to be crucial

With that we can create the reduced order model via projection

We would then get the reduced order system that minimizes

$$\| H(p) - \hat{H}(p) \|_{\mathcal{H}_2}$$

for each $p$.

**Idea**

$\Rightarrow$ metamodelling of $\sigma_i(p)$.

**Problem**

Is this even a function? How smooth?
Examples of $\sigma$

Beam Model

real

imag

Max Planck Institute Magdeburg

Grundel, MOR and RBF
Examples of $\sigma$
Examples of $\sigma$

Anemometer

\[ \cdot 10^4 \]

\[
\begin{array}{cccccc}
0 & 0.2 & 0.4 & 0.6 & 0.8 & 1 \\
0 & 1 & 2 & 3 & 4 & 5 \\
0 & 0.5 & 1 & 0 & 0.5 & 1
\end{array}
\]

real

imag
Synthetic Example

The graphs depict the real and imaginary parts of some complex data, with the x-axis ranging from $10^{-3}$ to $10^1$ and the y-axis ranging from $-1,000$ to $3,000$ for the real part, and from $0$ to $1,000$ for the imaginary part.
Synthetic Example

H₂ MOR
Parametric MOR
Numerik
Medium Model

Max Planck Institute Magdeburg
Grundel, MOR and RBF 14/34
2D Example

Scanning Electrochemical Microscopy

real

imag
Ordering

Complex Valued function

We have to order the interpolation points in order to make a complex valued function out of the set-valued function

- separate purely real and complex conjugate \( \sigma_i \);
- sort real ones regularly
- sort complex ones by real part first.
Metamodelling using k means

- in most applications the $\sigma_i$ behave quit nicely
- create metamodels for different clusters (clustering)
- given $p_1, \ldots, p_N$ consider tuples

$$(C_1 p_i, \sigma(p_i), C_2 n_i) \in \mathbb{R}^p \times \mathbb{C}^r \times \mathbb{N}$$

where $n_i$ measures the number of real values and $1 < C_1 < C_2$.

### k means

1. Set initial means for all $K$ clusters
2. assign each tuple to the cluster with the nearest mean
3. Calculate new mean
4. repeat until convergence
Radial Basis Interpolation

Ansatz

Given \( p_1, \ldots, p_N \) and function values \( \sigma(p_1), \ldots, \sigma(p_N) \) the interpolant is created by

\[
\tilde{\sigma}(p) = \sum \gamma_i R(\|p - p_i\|)
\]

where \( R(x) = \exp(-\theta x^2) \)

- simple interpolation technique
- \( \theta \) found problem dependent
- \( \gamma_i \) found by solving a linear system (interpolation condition)
- different model for each cluster
Smoothness "Theorem"

"Theorem"

If the matrices $A(p), B(p), C(p) \in C^\infty(D)$ then the function $\sigma(p) \in C^\infty(D)$ at least locally

Proof Ideas:

- Implicit Function Theorem on Wilson Condition $\Rightarrow \hat{A}(p)$ is smooth
- eigenvalues of parametrized function behave smooth typically
Error Analysis

$\mathcal{H}_2$ Error

If we assume that the metamodel is such that $\|\tilde{\sigma}(p) - \sigma(p)\| \leq \epsilon$ then we know that

$$\|H - \tilde{H}\|_{\mathcal{H}_2} \leq \|H - \hat{H}\|_{\mathcal{H}_2} + O(\epsilon^2)$$

This is true since $\sigma$ is a minimizer and the second derivative therefore vanishes.
Error Analysis

$\mathcal{H}_2$ Error

If we assume that the metamodel is such that $\|\tilde{\sigma}(p) - \sigma(p)\| \leq \epsilon$
then we know that

$$\|H - \tilde{H}\|_{\mathcal{H}_2} \leq \|H - \hat{H}\|_{\mathcal{H}_2} + \mathcal{O}(\epsilon^2)$$

This is true since $\sigma$ is a minimizer and the second derivative
therefore vanishes.

Problems

- just local not global minimizer
- clustering is heuristic
- RBF has no error bound
Outline

1. $\mathcal{H}_2$ MOR
2. Parametric MOR
3. Numerik
4. Medium Model
Anemometer - modelreduction.org

- $n=29008$, $r=6$
- $K=1$
- $p=[0,1]$, $N=5$

**transfer function**

- $p=0$
- $p=0.25$
- $p=0.5$
- $p=0.75$
- $p=1$

Max Planck Institute Magdeburg

Grundel, MOR and RBF 22/34
\[ \| H \|_{\mathcal{H}_2} \approx 2.7 \times 10^4 \]
Beam Model

- $n=240$, $r=10$
- $K=1$ (number of clusters)
- $p=[0.8,1.2]$, $N=3$

transfer function
Beam Model

$H_2$ Error

$\|H\|_{H_2} \approx 0.0035$

interpolated
IRKA
Synthetic

- $n=100$, $r=10$
- $K=4$
- $p=[0,1], N=50$

Transfer function:

- $p=0.25$
- $p=5$
- $p=0.75$
- $p=1$
Synthetic

$\mathcal{H}_2$ Error

$\|H\|_{\mathcal{H}_2} \approx 10$

* interpolated
* IRKA
On-line versus Off-line

**Off-line**
- precomputation
- time is not so important
- possible bigger computing resources

**On-line**
- simulate the reduced order model for different parameter or input functions
- computing time crucial
- phase 1: compute the reduced state space system
- phase 2: simulate it (system size $r$ is crucial)
**Anemometer Timings**

- \( N=5 \) (number of interpolation points in parameter domain)
- the error of the interpolated and projected function is very close to the error of a reduced order model computed by IRKA directly
- Depending on the application this may however be problematic timewise.

<table>
<thead>
<tr>
<th>Example</th>
<th>( r=4 )</th>
<th>( r=6 )</th>
<th>( r=10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>create ( \sigma ) model</td>
<td>86s</td>
<td>122 s</td>
<td>382 s</td>
</tr>
<tr>
<td>one IRKA run</td>
<td>43s</td>
<td>150s</td>
<td>216 s</td>
</tr>
<tr>
<td>do the projection</td>
<td>3s</td>
<td>4.8s</td>
<td>8s</td>
</tr>
</tbody>
</table>
Outline

1. $\mathcal{H}_2$ MOR
2. Parametric MOR
3. Numerik
4. Medium Model
Medium Model

General Idea

\[
\begin{bmatrix}
A(p) & b(p) \\
c^T(p)
\end{bmatrix}
\xrightarrow{\text{Medium}}
\begin{bmatrix}
A_m(p) & b_m(p) \\
c_m^T(p)
\end{bmatrix}
\xrightarrow{\tilde{\sigma} \text{ int}}
\begin{bmatrix}
\tilde{A}(p) & \tilde{b}(p) \\
\tilde{c}^T(p)
\end{bmatrix}
\]
Medium Model

General Idea

\[
\begin{bmatrix}
A(p) & b(p) \\
c^T(p)
\end{bmatrix}
\xrightarrow{\text{Medium}}
\begin{bmatrix}
A_m(p) & b_m(p) \\
c_m^T(p)
\end{bmatrix}
\xrightarrow{\tilde{\sigma} \text{ int}}
\begin{bmatrix}
\tilde{A}(p) & \tilde{b}(p) \\
\tilde{c}^T(p)
\end{bmatrix}
\]

Remarks

- Metamodel of $\sigma$ is created from original model
- Interpolation condition leads to system solve of moderate size (medium model)
- Generally one could use any medium size model that approximates the original one well
Medium Model

General Idea

\[
\begin{bmatrix}
A(p) & b(p) \\
c^T(p)
\end{bmatrix}
\xrightarrow{\nu \ \text{proj}}
\begin{bmatrix}
V^T A(p) V & V^T b(p)
\end{bmatrix}
\xrightarrow{\tilde{\sigma} \ \text{int}}
\begin{bmatrix}
\tilde{A}(p) & \tilde{b}(p)
\end{bmatrix}
\]

Remarks

- Metamodel of \(\sigma\) is created from original model
- Interpolation condition leads to system solve of moderate size (medium model)
- Generally one could use any medium size model that approximates the original one well
- \(V\) is created such that the medium size model interpolates at many points in frequency and parameter
**Algorithm 2** Offline Phase Calculation

1. Pick parameter points $p_1, \ldots, p_N$
2. **for** $i = 1$ to $N$ **do**
3. Compute via IRKA $\sigma(p_i)$ and $V_i$, $W_i$ projection matrices
4. **end for**
5. Create metamodel
6. Compute $V$ from all $V_i$ and $W_i$
7. Precompute medium size matrices with $V$
Algorithm 3 Online Phase Calculation

**Input:** $p \in \mathcal{P}$

**Output:** Reduced state space system $\tilde{A}, \tilde{b}, \tilde{c}$

1. Compute $\tilde{\sigma}(p)$

2. Solve $2r$ linear systems of medium size to create $V, W$

3. Project medium size model onto small model via $V, W$
Error Bounds

Lemma

Assuming that $\|H - H^m\|_\infty \leq \epsilon \|H\|_\infty$ and $\sigma_1, \ldots, \sigma_r$ given interpolation points. If $H_r$ interpolates $H$ and $H^m_r$ interpolates $H^m$ then

$$\|H_r - H^m_r\| \leq (\epsilon + \delta + \epsilon\delta)\|H_r\| + \delta\|H^m_r\|$$

where $\delta = \sum_{k=1}^{\infty} (\|L\|\|L^{-1}\|\epsilon)^k$

This is basically related to forward stability of rational interpolation.

$$L_{ij} = \begin{cases} H(\sigma_i(p), p) - H(\sigma_j(p), p) & \text{if } i \neq j \\ \frac{\sigma_i(p) - \sigma_j(p)}{\sigma_i(p) - \sigma_j(p)} & \text{if } i = j \end{cases}$$
### Comparison

#### SECM Example

\[ N=10, \ r=4, \ n=16912, \]

<table>
<thead>
<tr>
<th>Example</th>
<th>IRKA</th>
<th>large proj</th>
<th>medium proj</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathcal{H}_2 ) error</td>
<td>5e-7</td>
<td>7.7e-7</td>
<td>7.7e-7</td>
</tr>
<tr>
<td>on-line cost</td>
<td>80s</td>
<td>8s</td>
<td>0.1s</td>
</tr>
<tr>
<td>off-line cost</td>
<td>0s</td>
<td>1365s</td>
<td>1366s</td>
</tr>
</tbody>
</table>
Comparison

SECM Example

\[ N=10, \; r=4, \; n=16912, \]

<table>
<thead>
<tr>
<th>Example</th>
<th>IRKA</th>
<th>large proj</th>
<th>medium proj</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_2 ) error</td>
<td>5e-7</td>
<td>7.7e-7</td>
<td>7.7e-7</td>
</tr>
<tr>
<td>on-line cost</td>
<td>80s</td>
<td>8s</td>
<td>0.1s</td>
</tr>
<tr>
<td>off-line cost</td>
<td>0s</td>
<td>1365s</td>
<td>1366s</td>
</tr>
</tbody>
</table>

The medium model itself is not a good approximation but its projection on almost optimal points is close to the true best.

- online cost is just to cost to create the reduced order model, not to simulate anything.
- off-line cost is cost to create the metamodel and medium size model (several IRKA runs mainly)
Summary

1. introduction to $\mathcal{H}_2$ Model Order Reduction
2. new approach to Parametric Model Order Reduction using RBFs
3. the direct method needs some extra online computation time
4. medium model can reduce that to a small amount
5. some open problems in clustering, related to the smoothness of the function $\sigma$
Thank you
Thank you

Max Planck Institute Magdeburg

Grundel, MOR and RBF
Thank you