August, 2013
NLA and Optimization Workshop
Vancouver

$\mathcal{H}_2$ optimal model order reduction for parametric systems using RBF metamodels

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

MAX PLANCK INSTITUTE FOR DYNAMICS OF COMPLEX TECHNICAL SYSTEMS MAGDEBURG

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

What is MOR?

$$E\,\dot{x}(t) = A\,x(t) + B\,u(t);$$
$$y(t) = C\,x(t)$$

MOR

$$\hat{E}\,\dot{\hat{x}}(t) = \hat{A}\,\hat{x}(t) + \hat{B}\,u(t);$$
$$\hat{y}(t) = \hat{C}\,\hat{x}(t)$$

Projection-Based MOR

LTI System:

$$\dot{x}(t) = \mathbf{A}x(t) + \mathbf{b}u(t),$$
$$y(t) = \mathbf{c}^T x(t), \quad x(0) = 0.$$

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

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

Projection-Based MOR

LTI System:

$$\dot{x}(t) = \mathbf{A}x(t) + \mathbf{b}u(t),$$
$$y(t) = \mathbf{c}^T x(t), \quad x(0) = 0.$$

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

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

Projection-Based MOR

LTI System:

$$\dot{x}(t) = \mathbf{A}x(t) + \mathbf{b}u(t),$$
$$y(t) = \mathbf{c}^T x(t), \quad x(0) = 0.$$

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

$$\dot{\hat{x}}(t) = \hat{\mathbf{A}}\hat{x}(t) + \hat{\mathbf{b}}u(t)$$
$$\hat{y}(t) = \hat{\mathbf{c}}^T\hat{x}(t).$$

Projection-Based MOR

LTI System:

$$\dot{x}(t) = \mathbf{A}x(t) + \mathbf{b}u(t),$$
$$y(t) = \mathbf{c}^T x(t), \quad x(0) = 0.$$

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

$$\dot{\hat{x}}(t) = \hat{\mathbf{A}}\hat{x}(t) + \hat{\mathbf{b}}u(t)$$
$$\hat{y}(t) = \hat{\mathbf{c}}^T\hat{x}(t).$$

$$\begin{Bmatrix} \dot{x} = \mathbf{A}x + \mathbf{b}u \\ y = \mathbf{c}^T x \end{Bmatrix} \xrightarrow{\text{Lapl}} \begin{Bmatrix} sX = \mathbf{A}X + \mathbf{b}U \\ Y = \mathbf{c}^T X \end{Bmatrix} \rightarrow \begin{Bmatrix} X = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}U \\ Y = \mathbf{c}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}U \end{Bmatrix}$$

Projection-Based MOR

LTI System:

$$\dot{x}(t) = \mathbf{A}x(t) + \mathbf{b}u(t),$$
$$y(t) = \mathbf{c}^T x(t), \quad x(0) = 0.$$

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

$$\dot{\hat{x}}(t) = \hat{\mathbf{A}}\hat{x}(t) + \hat{\mathbf{b}}u(t)$$
$$\hat{y}(t) = \hat{\mathbf{c}}^T\hat{x}(t).$$

$$\begin{Bmatrix} \dot{x} = \mathbf{A}x + \mathbf{b}u \\ y = \mathbf{c}^T x \end{Bmatrix} \xrightarrow{\text{Lapl}} \begin{Bmatrix} sX = \mathbf{A}X + \mathbf{b}U \\ Y = \mathbf{c}^T X \end{Bmatrix} \rightarrow \begin{Bmatrix} X = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}U \\ Y = \mathbf{c}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}U \end{Bmatrix}$$

We define the transfer functions

$$\hat{H}(s) = \hat{\mathbf{c}}^T(s\mathbf{I} - \hat{\mathbf{A}})^{-1}\hat{\mathbf{b}} \approx H(s) = \mathbf{c}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}$$

which is a rational function in s of degree r or n.

$\mathcal{H}_2$ Model Order Reduction

Good Reduced Order Model

$$\left\{\begin{array}{l} u \xrightarrow{\Sigma} y \\ u \xrightarrow{\hat{\Sigma}} \hat{y} \end{array}\right\} \quad \|y - \hat{y}\| \text{ small}$$

We know that:

$$\sup_{t \geq 0} |y(t) - \hat{y}(t)| \leq \|H - \hat{H}\|_{\mathcal{H}_2} \|u\|_{L_2}.$$

for $\|H - \hat{H}\|_{\mathcal{H}_2} := \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\imath\omega) - \hat{H}(\imath\omega)|^2 d\omega \right)^{1/2}$.

References

[Absil, Antoulas, Baur, Beattie, Benner, Breiten, Bunse-Gerstner, Gallivan, Gugercin, Kubalinska, Van Dooren, Vossen, Wilczek,...]

$\mathcal{H}_2$ 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. [MEYER, LUENBERGER 1967]

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

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

$$\begin{aligned} (\sigma \mathbf{I} - \mathbf{A})^{-1}\mathbf{b} &\in Ran(\mathbf{V}) \\ (\overline{\sigma}\mathbf{I} - \mathbf{A}^T)^{-1}\mathbf{c} &\in Ran(\mathbf{W}) \\ \Rightarrow H(\sigma) = \hat{H}(\sigma) \quad H'(\sigma) &= \hat{H}'(\sigma) \end{aligned}$$

[GRIMME, YOUSOUFF, SKELETON].

$\mathcal{H}_2$ 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. [Meyer, Luenberger 1967]

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

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

σ_is are not a priori known, but can be found by IRKA

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

[Grimme, Yousouff, Skeleton].

IRKA

[Antoulas,Beattie,Gugercin 2006]

Algorithm 1 Iterative rational Krylov algorithm (IRKA)

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

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

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

Outline

Parametrized Dynamical System

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

$$\dot{x}(t) = \mathbf{A}(p)x(t) + \mathbf{b}(p)u(t),$$
$$y(t) = \mathbf{c}(p)^T x(t), \quad x(0) = 0.$$

Model Reduction:

$$\dot{\hat{x}}(t) = \hat{\mathbf{A}}(p)\hat{x}(t) + \hat{\mathbf{b}}(p)u(t)$$
$$\hat{y}(t) = \hat{\mathbf{c}}(p)^T \hat{x}(t)$$

This means that the approximated transfer function

$$\hat{H}(s,p) = \hat{\mathbf{c}}(p)^T (s\mathbf{I} - \hat{\mathbf{A}}(p))^{-1} \hat{\mathbf{b}}(p) \approx H(s,p) = \mathbf{c}(p)^T (s\mathbf{I} - \mathbf{A}(p))^{-1} \mathbf{b}(p)$$

is a rational function in s, but also a function in p.

Previous Work

Reduced matrices from original matrices

$$\mathbf{A}(p) \to \hat{\mathbf{A}}(p) \quad \mathbf{c}(p) \to \hat{\mathbf{c}}(p) \quad \mathbf{b}(p) \to \hat{\mathbf{b}}(p) \tag{2}$$

Many attempt for parameteric Model Order Reduction exist

- projection matrix independent of parameter
  $\hat{\mathbf{A}} = V^T \mathbf{A}(p) W$
  [Breiten, Damm, Baur, Benner, Beattie, Gugercin]
- matrix interpolation
  $\hat{\mathbf{A}}(p_i)$
  [Panzer et al] or [Amsallam, Farhat]
- transfer function interpolation
  $H(s_i, p_j)$
  [Antoulas, Ionita]

PMOR and $\mathcal{H}_2$

- 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

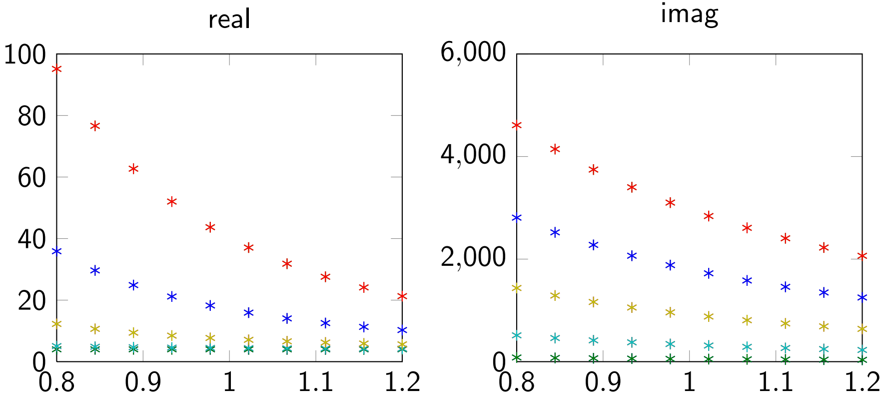
$\Rightarrow$ metamodelling of $\sigma_i(p)$.

Problem

Is this even a function? How smooth?

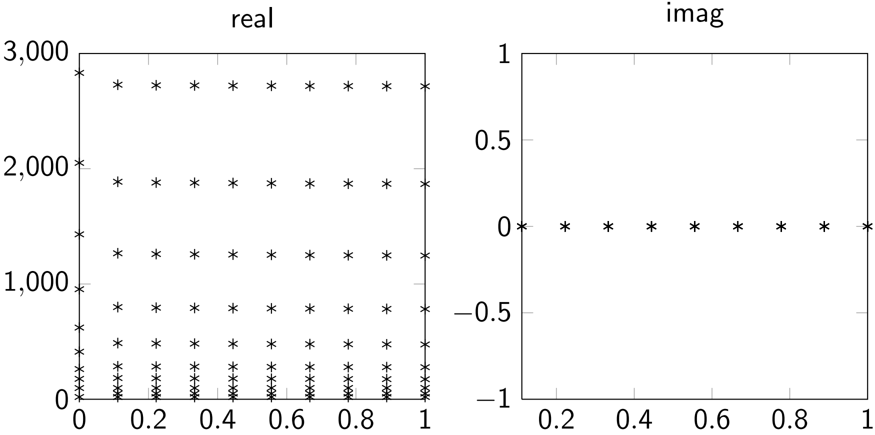
Examples of σ

Beam Model



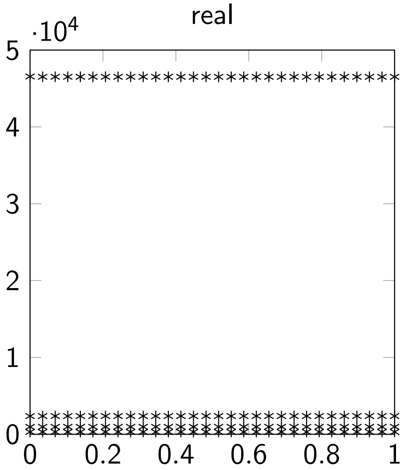
Examples of σ

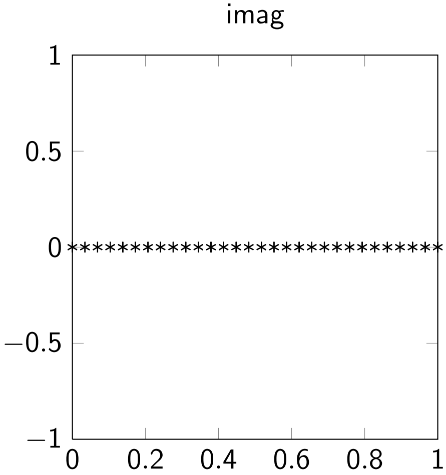
Convection Diffusion Model



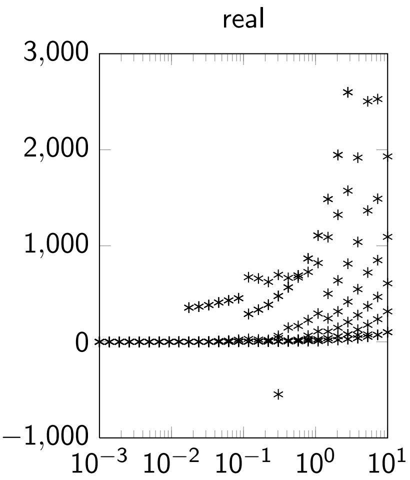
Examples of σ

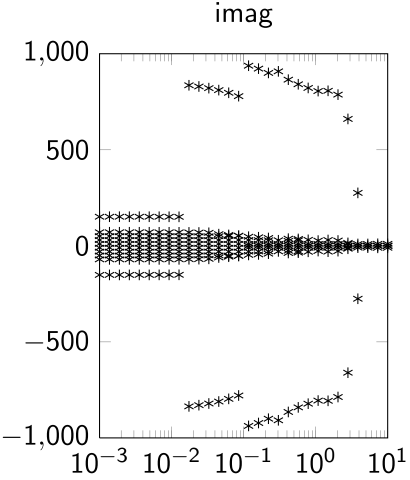
Anemometer



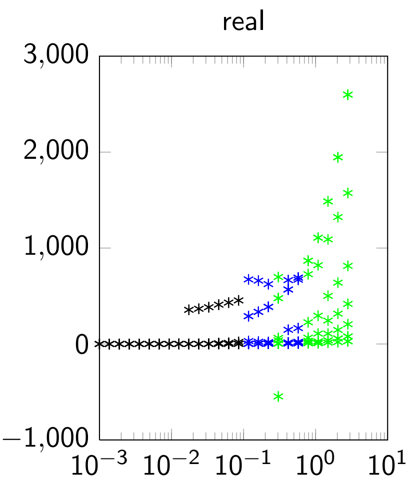


Synthetic Example



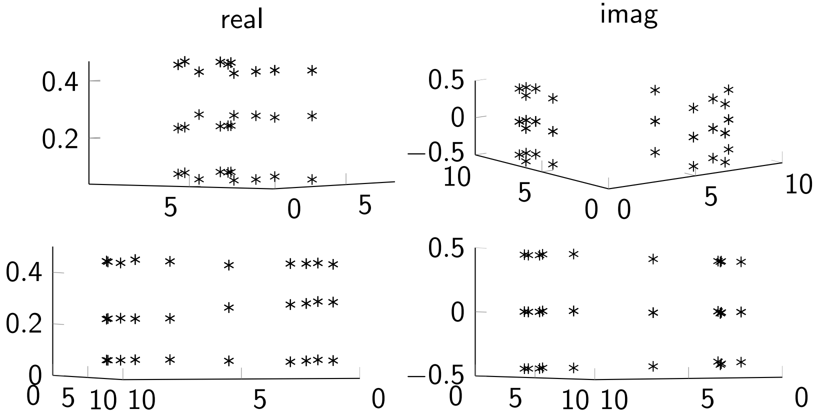


Synthetic Example



2D Example

Scanning Electrochemical Microscopy



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 σ_i
- sort real ones regulary
- sort complex ones by real part first.

Metamodelling using k means

- in most applications the σ_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
- θ found problem dependent
- γ_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

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

This is true since σ 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 σ is a minimizer and the second derivative therefore vanishes.

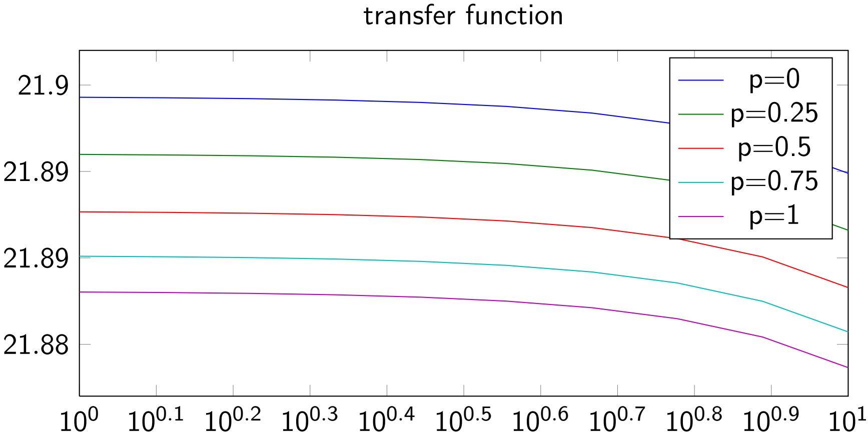
Problems

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

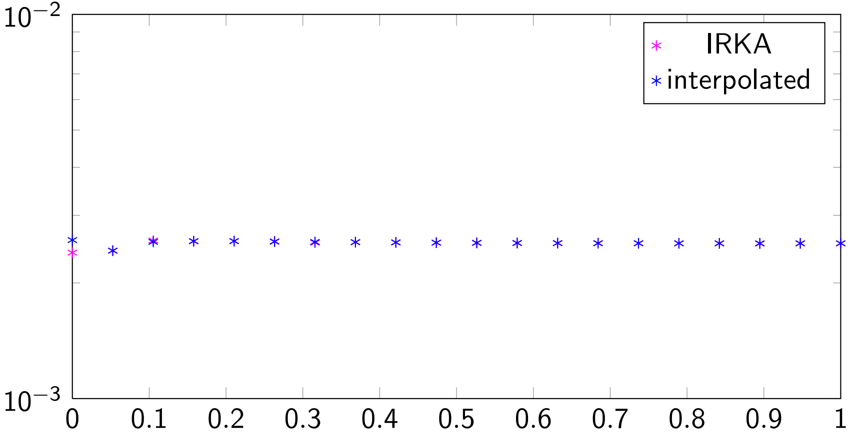
Outline

Anemometer - modelreduction.org

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



Anemometer - modelreduction.org

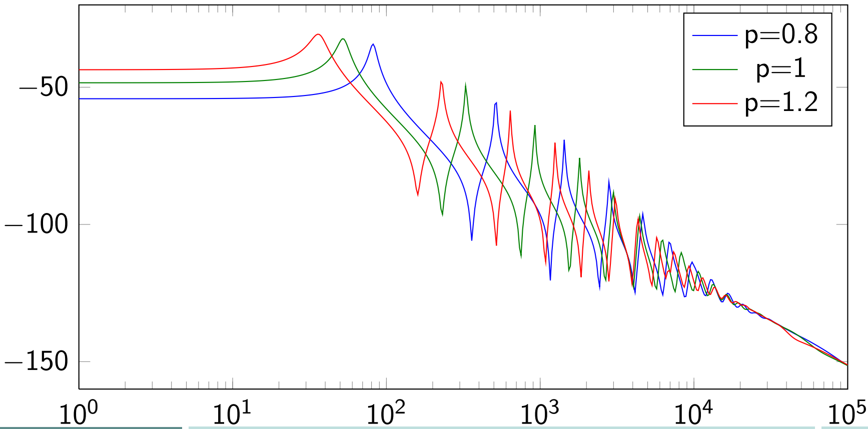


$$\|H\|_{\mathcal{H}_2} \approx 2.7e4$$

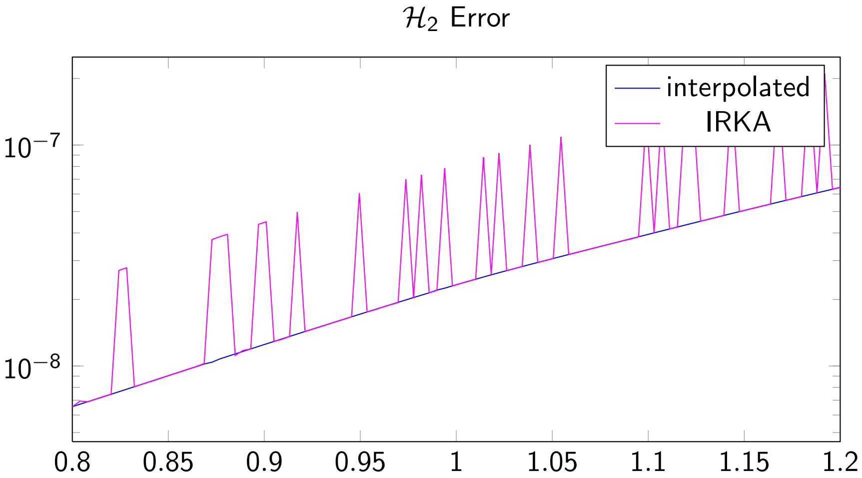
Beam Model

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

transfer function



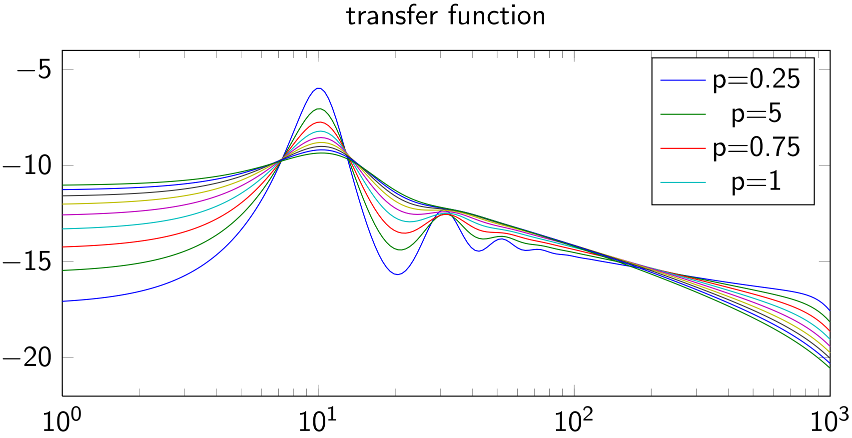
Beam Model



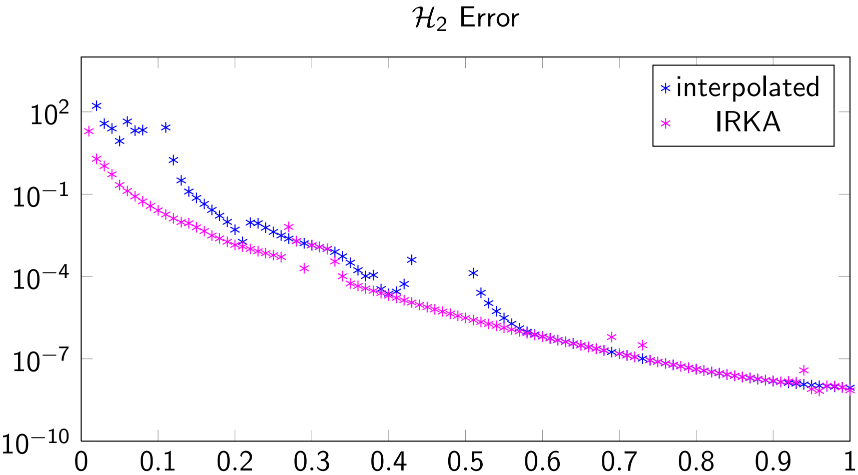
$$\|H\|_{\mathcal{H}_2} \approx 0.0035$$

Synthetic

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



Synthetic



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

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.

Example	r=4	r=6	r=10
create σ model	86s	122 s	382s
one IRKA run	43s	150s	216s
do the projection	3s	4.8s	8s

Outline

Medium Model

General Idea

$$\begin{bmatrix} \mathbf{A}(p) & \mathbf{b}(p) \\ \mathbf{c}^T(p) & \end{bmatrix} \xrightarrow{\text{Medium}} \begin{bmatrix} \mathbf{A}_m(p) & \mathbf{b}_m(p) \\ \mathbf{c}_m^T(p) & \end{bmatrix} \xrightarrow{\tilde{\sigma}\ \text{int}} \begin{bmatrix} \tilde{\mathbf{A}}(p) & \tilde{\mathbf{b}}(p) \\ \tilde{\mathbf{c}}^T(p) & \end{bmatrix}$$

Medium Model

General Idea

$$\begin{bmatrix} \mathbf{A}(p) & \mathbf{b}(p) \\ \mathbf{c}^T(p) & \end{bmatrix} \xrightarrow{\text{Medium}} \begin{bmatrix} \mathbf{A}_m(p) & \mathbf{b}_m(p) \\ \mathbf{c}_m^T(p) & \end{bmatrix} \xrightarrow{\tilde{\sigma}\ \text{int}} \begin{bmatrix} \tilde{\mathbf{A}}(p) & \tilde{\mathbf{b}}(p) \\ \tilde{\mathbf{c}}^T(p) & \end{bmatrix}$$

Remarks

- metamodel of σ 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

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

Remarks

- metamodel of σ 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

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

Algorithm 3 Online Phase Calculation

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

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

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 [Higham 2004, Grundel-Benner 2013]

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_r^m interpolates H^m then

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

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

This is basically related to forward stability of rational interpolation.

$$\mathbb{L}_{ij} = \begin{cases} \frac{H(\sigma_i(p),p) - H(\sigma_j(p),p)}{\sigma_i(p) - \sigma_j(p)} & \text{if } i \neq j \\ \frac{\partial}{\partial \sigma} H(\sigma_i(p), p) & \text{if } i = j \end{cases}$$

Comparison

SECM Example

N=10, r=4, n=16912,

Example	IRKA	large proj	medium proj
$\mathcal{H}_2$ error	5e-7	7.7e-7	7.7e-7
on-line cost	80s	8s	0.1s
off-line cost	0s	1365s	1366s

Comparison

SECM Example

N=10, r=4, n=16912,

Example	IRKA	large proj	medium proj
$\mathcal{H}_2$ error	5e-7	7.7e-7	7.7e-7
on-line cost	80s	8s	0.1s
off-line cost	0s	1365s	1366s

The medium model itelf 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 (severel 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 σ

Thank you

Thank you

Thank you