Parametric Model Order Reduction Using Sparse Grids

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Sina Brunsch, geb. Meister

aus

Burgstädt

Bonn 2016

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Michael Griebel

2. Gutachter: Prof. Dr. Peter Benner

Tag der Promotion: 28.07.2017

Erscheinungsjahr: 2017

Abstract

In applications such as very large scale integration chip design models are typically huge. The same is true in mechanical engineering, especially when the models are complex finite element discretizations. To speed up simulations the large full order model is replaced by a smaller reduced order model. This is called model order reduction. The challenge is to find a reduced order model that closely resembles the full order model in order not to lose too much accuracy.

Many models depend on parameters. The goal of parametric model order reduction is to preserve this dependence in the reduced order model to avoid repeatedly performing time-consuming model order reduction for every new parameter value. This is particularly interesting for parameter studies.

In this thesis we develop, analyze and test a parametric model order reduction method for large symmetric linear time-invariant dynamical systems which preserves stability and is efficient in higher-dimensional parameter spaces. This method is based on sparse grid interpolation. In a pre-computation step, local reduced order models are computed at several discrete points in parameter space. Whenever we need to evaluate the model at an arbitrary point in parameter space, a reduced order model at that point is obtained by interpolating the system matrices of the local reduced order models on matrix manifolds.

As a theoretical foundation we introduce appropriate norms for parametric models and state conditions for the parameter dependence such that these norms are finite. In our analysis we then derive an upper bound for the interpolation error expressed in these norms, which shows a good qualitative behavior in computational experiments.

We demonstrate that interpolation on sparse grids is more efficient than interpolation on full grids and that interpolation with global polynomial basis functions is more efficient than interpolation with piece-wise linear basis functions when the parameter dependence is smooth. Furthermore, we consider a benchmark studied in a recent parametric model order reduction method comparison survey and show that parametric model order reduction methods based on matrix interpolation can be competitive to other methods when exploiting the symmetry of that system.

Zusammenfassung

In vielen Anwendungen wie dem Entwurf von VLSI-Schaltkreisen treten sehr große Simulationsmodelle auf. Ein weiteres Beispiel sind physikalische Modelle aus dem Maschinen- oder Fahrzeugbau, die mit finiten Elementen sehr fein diskretisiert wurden. Man kann die Simulation solcher Modelle mithilfe von Modellreduktion beschleunigen. Dabei ersetzt man das große Modell durch ein kleineres, reduziertes Modell. Die Herausforderung besteht darin, ein passendes reduziertes Modell zu finden. Um möglichst wenig Genauigkeit einzubüßen, sollten sich das ursprüngliche Modell und das reduzierte Modell möglichst ähnlich verhalten.

Oft sind die Modelle außerdem parameterabhängig. Bei parametrischer Modellreduktion versucht man, diese Abhängigkeit im reduzierten Modell zu erhalten, um dadurch das wiederholte Ausführen der zeitaufwändigen Modellreduktion für jeden neuen Parameterwert zu vermeiden. Dies ist besonders für Parameterstudien von wesentlicher Bedeutung.

In dieser Dissertation entwickeln, analysieren und testen wir ein stabilitätserhaltendes parametrisches Modellreduktionsverfahren für große symmetrische lineare zeit-invariante dynamische Systeme, welches aufgrund seiner Effizienz auch für den Einsatz in höher-dimensionalen Parameterräumen geeignet ist. Dieses Verfahren basiert auf Dünngitterinterpolation. In einem Vorverarbeitungsschritt werden lokale reduzierte Modelle an wenigen festen Punkten im Parameterraum erzeugt. Durch Interpolation der Systemmatrizen der lokalen reduzierten Modelle auf Matrixmannigfaltigkeiten erhält man das reduzierte Modell an einem beliebigen Punkt im Parameterraum.

Als theoretische Grundlage führen wir passende Normen für parametrische Modelle ein und formulieren Bedingungen an die Parameterabhängigkeit des Modells, sodass diese Normen endlich sind. In unserer Analyse leiten wir dann eine obere Schranke für den Interpolationsfehler in diesen Normen her. Diese zeigt in numerischen Experimenten ein qualitativ gutes Verhalten.

Wir demonstrieren, dass Interpolation auf dünnen Gittern effizienter ist als Interpolation auf vollen Gittern und dass Interpolation mit globalen Polynomen effizienter ist als Interpolation mit stückweise linearen Basisfunktionen, vorausgesetzt die Parameterabhängigkeit ist glatt. Außerdem zeigen wir anhand eines Benchmarks, welches in einem aktuellen Vergleichsartikel über parametrische Modellreduktionsverfahren verwendet wird, dass Verfahren, die auf der Interpolation von Systemmatrizen beruhen, mit anderen parametrischen Modellreduktionsverfahren konkurrieren können, wenn man die Symmetrie des Modells ausnutzt.

Acknowledgments

I thank especially

- Michael Griebel for his kind guidance, support and for providing excellent working conditions,
- Christian Rieger for several illuminating discussions and for his precious ideas and comments,
- my other helpful colleagues for amusing lunch time conversations,
- Peter Benner for valuable hints and for inviting me to interesting group meetings,
- Ulrike Baur for many informative discussions and for useful comments on parts of this thesis,
- Serkan Gugercin for providing his implementation of the Iterative Rational KRYLOV Algorithm,
- my husband for continually asking critical questions, for motivating me and for proof-reading this thesis
- and my wonderful daughter for being a constant source of joy.

Contents

Abstract iii

Zusammenfassung v

Acknowledgments vii

Abbreviations xiii

Notation xv

1 Introduction 1
1.1 Related Work . . . 2
1.1.1 Model Order Reduction . . . 2
1.1.2 Parametric Model Order Reduction . . . 3
1.1.3 Sparse Grids . . . 3
1.2 Contributions of This Thesis . . . 4
1.3 Sources of Error . . . 5
1.4 Outline . . . 6

2 Fundamentals 7
2.1 Basic Definitions . . . 7
2.1.1 Norms of Vectors and Matrices . . . 7
2.1.2 Multi-indices . . . 8
2.1.3 Lebesgue–Bochner Spaces and Sobolev Spaces . . . 9
2.1.4 Spaces of Continuous and Continuously Differentiable Functions . 10
2.1.5 Hardy Spaces . . . 11
2.1.6 Functions of Matrices . . . 11
2.2 Model Order Reduction . . . 13
2.2.1 Model . . . 13
2.2.2 Model Order Reduction by Projection . . . 15
2.2.3 Error Measures . . . 16
2.2.4 Methods . . . 17
2.3 Interpolation on Sparse Grids . . . 20
2.3.1 Piece-wise Linear Basis Functions . . . 21
2.3.2 Global Polynomial Basis Functions . . . 30
2.3.3 Basis Representations . . . 30
2.3.4 Entry-wise Interpolation of Matrix-valued Functions . . . 34

3 Parametric Model Order Reduction **35**
3.1 Parametric Model . . . 35
3.1.1 Requirements . . . 35
3.2 Parametric Model Order Reduction . . . 36
3.3 Error Measures . . . 36
3.4 Methods . . . 39
3.4.1 Interpolation of Local Reduced System Matrices on Manifolds . . . 40

4 A Stability-preserving Parametric Model Order Reduction Method **43**
4.1 Algorithm . . . 43
4.1.1 Performance Improvements . . . 46
4.2 Design Decisions . . . 46
4.2.1 System Structure . . . 47
4.2.2 Reference Basis . . . 47
4.2.3 Model Order Reduction Method . . . 47
4.2.4 Switch Projection and Reduced Order Basis Transformation . . . 49
4.2.5 Additional Reduced Order Basis Transformation to Simplify System 50
4.2.6 Reference Point for Tangent Space . . . 51
4.3 Properties . . . 51
4.3.1 Stability Preservation . . . 51
4.3.2 Preservation of Model Requirements for Well-defined Error Measures 52
4.3.3 Transfer Function Interpolation Property . . . 58

5 Interpolation Error Bounds **61**
5.1 Bound for $\mathcal{H}_\infty$-norm at Fixed Point . . . 63
5.1.1 Final Theorem . . . 63
5.1.2 The $\Delta\mathbf{M}$-term . . . 65
5.1.3 The $\Delta\mathbf{B}$- and the $\Delta\mathbf{C}$-term . . . 69
5.2 Bound for $\mathcal{H}_2$-norm at Fixed Point . . . 69
5.2.1 Final Theorem . . . 69
5.2.2 The $\Delta\mathbf{M}$-term . . . 71
5.2.3 The $\Delta\mathbf{B}$- and the $\Delta\mathbf{C}$-term . . . 72
5.3 Bound for $\mathcal{L}_\infty(\mathcal{P}, \mathcal{H}_\infty)$-norm . . . 73
5.4 Bound for $\mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$-norm . . . 74
5.5 Technical Lemmas . . . 75

6 Error Convergence Rates for Interpolation on Sparse Grids **83**
6.1 Interpolation with Piece-wise Linear Basis Functions . . . 83
6.1.1 Interpolation of Scalar-valued Functions . . . 84
6.1.2 Entry-wise Interpolation of Matrix-valued Functions . . . 84
6.1.3 Transfer Function . . . 87
6.2 Interpolation with Global Polynomial Basis Functions . . . 89
6.2.1 Interpolation of Scalar-valued Functions . . . 89
6.2.2 Entry-wise Interpolation of Matrix-valued Functions . . . 89

6.2.3 Transfer Function 90

7 Computational Experiments 93

7.1 Settings 93

7.1.1 Benchmarks 93

7.1.2 Model Order Reduction 94

7.1.3 Discrete Error Measures 95

7.1.4 Sparse Grid Interpolation Toolbox 95

7.1.5 Types of Sparse Grids 96

7.2 Tests 97

7.2.1 Full Grids and Different Types of Sparse Grids 97

7.2.2 Influence of Additional Reduced Order Basis Transformation . . . 98

7.2.3 Reduction Order 99

7.2.4 Multi-input, Multi-output and Single-input, Single-output Systems 100

7.2.5 Dimension of Parameter Space 101

7.2.6 Performance of Matrix Interpolation PMOR Methods 101

7.2.7 Silicon Nitride Membrane Benchmark 105

7.3 Comparison with Theoretical Results 106

7.3.1 Interpolation Error Bound 106

7.3.2 Interpolation Error Convergence Rates 108

8 Conclusions and Outlook 111

8.1 Conclusions 111

8.2 Outlook on Future Work 112

8.2.1 Model 112

8.2.2 Model Order Reduction 112

8.2.3 Interpolation 113

A Definitions and Theorems from Literature 115

A.1 Matrix Basics 115

A.1.1 Properties 115

A.1.2 Transformations 116

A.1.3 Decompositions 116

A.1.4 Projections 117

A.1.5 Bounds for Matrix Exponential and Matrix Logarithm 117

A.2 Perturbation Theory 118

A.2.1 Matrix Inverse 118

A.2.2 Matrix Exponential 118

A.3 Miscellaneous 118

A.3.1 HÖLDER's Inequalities 118

A.3.2 LAPLACE Transform 118

A.3.3 Integration Formula 119

B Proofs **121**
B.1 Proofs for chapter 2 121
B.1.1 Proofs for subsection 2.2.3 121
B.1.2 Proofs for subsection 2.3.1 122
B.2 Proofs for chapter 3 125
B.3 Proofs for chapter 4 127
B.3.1 Proofs for subsection 4.3.1 127
B.3.2 Proofs for subsection 4.3.2 127
B.4 Proofs for chapter 5 128
B.4.1 Technical Lemmas 129
B.4.2 Bound for $\mathcal{H}_\infty$-norm at Fixed Point 132
B.4.3 Bound for $\mathcal{H}_2$-norm at Fixed Point 132

List of Algorithms **137**

List of Figures **139**

List of Tables **141**

Bibliography **143**

Index **149**

Abbreviations

LTI	linear time-invariant (system)
SISO	single-input, single-output (system)
MIMO	multi-input, multi-output (system)
MOR	model order reduction
PMOR	parametric model order reduction
FOM	full order model
ROM	reduced order model
ROB	reduced order basis
IRKA	Iterative Rational KRYLOV Algorithm
SVD	singular value decomposition

Notation

$\imath$	imaginary unit ($\imath = \sqrt{-1}$)
$\mathrm{Re}(z)$	real part of a complex number z
$\mathrm{Im}(z)$	imaginary part of a complex number z
$\overline{z}$	complex conjugate of a complex number z
$\dot{f}(t)$	derivative with respect to time of a function $f(t)$
δ_{ij}	Kronecker delta ($\delta_{ij} = 0$ for $i \neq j$, $\delta_{ii} = 1$)
$i = 1\colon n$	$i \in \{1, \dots, n\}$

Sets

$\mathbb{N}$	natural numbers (excluding 0)
$\mathbb{N}_0$	natural numbers including 0 ($\mathbb{N}_0 = \mathbb{N} \cup \{0\}$)
$\mathbb{C}$	complex numbers
$\mathbb{C}^-$	open left half of the complex plane ($\mathbb{C}^- = \{c \in \mathbb{C} : \mathrm{Re}(c) < 0\}$)
$\mathbb{C}^+$	open right half of the complex plane ($\mathbb{C}^+ = \{c \in \mathbb{C} : \mathrm{Re}(c) > 0\}$)
$\overline{\mathbb{C}^-}$	closed left half of the complex plane ($\mathbb{C}^- = \{c \in \mathbb{C} : \mathrm{Re}(c) \leq 0\}$)
$\overline{\mathbb{C}^+}$	closed right half of the complex plane ($\mathbb{C}^+ = \{c \in \mathbb{C} : \mathrm{Re}(c) \geq 0\}$)
$\mathbb{R}$	real numbers
$\mathbb{R}^-$	negative real numbers ($\mathbb{R}^- = \{a \in \mathbb{R} : a < 0\}$)
$\mathbb{R}_0^-$	negative real numbers including 0 ($\mathbb{R}^- = \{a \in \mathbb{R} : a \leq 0\}$)
$\mathbb{R}^+$	positive real numbers ($\mathbb{R}^+ = \{a \in \mathbb{R} : a > 0\}$)
$\mathbb{R}_0^+$	positive real numbers including 0 ($\mathbb{R}^+ = \{a \in \mathbb{R} : a \geq 0\}$)
$\mathrm{span}(S)$	set of linear combinations of elements of a set S

Matrices, Vectors and Their Norms

$\mathbf{M}^T$	transpose of a matrix $\mathbf{M}$
$\mathbf{M}^H$	conjugate transpose of a matrix $\mathbf{M}$
$\mathbf{M}^{-1}$	inverse of a matrix $\mathbf{M}$
$\Lambda(\mathbf{M})$	spectrum of a matrix $\mathbf{M}$ (set of eigenvalues)
$\lambda_i(\mathbf{M})$	eigenvalue of a matrix $\mathbf{M}$ ($\lambda_i(\mathbf{M}) \in \Lambda(\mathbf{M})$)
$\lambda_{\max}(\mathbf{M})$	largest eigenvalue of a matrix $\mathbf{M}$ with real spectrum
$\lambda_{\min}(\mathbf{M})$	smallest eigenvalue of a matrix $\mathbf{M}$ with real spectrum
$\sigma_i(\mathbf{M})$	singular value of a matrix $\mathbf{M}$ (see Theorem A.10)
$\sigma_{\max}(\mathbf{M})$	largest singular value of a matrix $\mathbf{M}$
$\mathrm{rank}(\mathbf{M})$	rank of a matrix $\mathbf{M}$
$\mathrm{range}(\mathbf{M})$	range of a matrix $\mathbf{M}$ (span of column vectors)

$\texttt{orth}(\mathbf{M})$	orthogonalization of the columns of a matrix $\mathbf{M}$
$\exp(\mathbf{M})$, $e^{\mathbf{M}}$	matrix exponential of a matrix $\mathbf{M}$ (see Definition 2.5)
$\log(\mathbf{M})$	matrix logarithm of a matrix $\mathbf{M}$ (see Theorem 2.6)
$\mathbf{0}$; $\mathbf{1}$; $\mathbf{2}$	vector whose entries are all equal to zero/ one/ two
$\mathbf{e}_i$	ith unit vector ($\mathbf{e}_i = [e_{i1}, \ldots, e_{in}]$ with $e_{ij} = \delta_{ij}$)
$\mathbf{O}$, $\mathbf{O}_{n,m}$	zero matrix (of dimension $n \times m$)
$\mathbf{I}$, $\mathbf{I}_n$	identity matrix (of dimension $n \times n$)
$\mathrm{diag}(d_1, \ldots, d_n)$	diagonal matrix with entries $d_1, \ldots, d_n$ (of dimension $n \times n$)
$\Vert\mathbf{v}\Vert_p$	p-norm of a vector $\mathbf{v}$ (see Definition 2.1)
$\Vert\mathbf{v}\Vert_1$; $\vert\boldsymbol{\alpha}\vert_1$	sum norm of a vector $\mathbf{v}$/ multi-index $\boldsymbol{\alpha}$
$\Vert\mathbf{v}\Vert_2$	Euclidean norm of a vector $\mathbf{v}$
$\Vert\mathbf{v}\Vert_\infty$; $\vert\boldsymbol{\alpha}\vert_\infty$	maximum norm of a vector $\mathbf{v}$/ multi-index $\boldsymbol{\alpha}$
$\Vert\mathbf{M}\Vert_{S,p}$	Schatten p-norm of a matrix $\mathbf{M}$ (see Definition 2.2)
$\Vert\mathbf{M}\Vert_F$	Frobenius norm of a matrix $\mathbf{M}$ ($\Vert\mathbf{M}\Vert_F = \Vert\mathbf{M}\Vert_{S,2}$; see (2.1))
$\Vert\mathbf{M}\Vert_2$	spectral norm of a matrix $\mathbf{M}$ ($\Vert\mathbf{M}\Vert_2 = \Vert\mathbf{M}\Vert_{S,\infty}$; see (2.2))
$\Vert\mathbf{M}\Vert_p$	p-norm of a matrix $\mathbf{M}$ (see (2.3))
$\kappa_p(\mathbf{M})$	condition number of a matrix $\mathbf{M}$ ($\kappa_p(\mathbf{M}) = \Vert\mathbf{M}\Vert_p \cdot \Vert\mathbf{M}^{-1}\Vert_p$)

Function Spaces and Their Norms

Ω	domain
$\partial\Omega$	boundary of domain Ω
$\mathcal{L}_2$	Lebesgue–Bochner space with norm $\Vert f\Vert_{\mathcal{L}_2}^2 := \int_\Omega \Vert f(\mathbf{p})\Vert_{\mathcal{B}}^2 \,\mathrm{d}\mathbf{p}$ (see (2.6))
$\mathcal{L}_\infty$	Lebesgue–Bochner space with norm $\Vert f\Vert_{\mathcal{L}_\infty} := \operatorname{ess\,sup}_{\mathbf{p}\in\Omega} \Vert f(\mathbf{p})\Vert_{\mathcal{B}}$ (see (2.7))
$\mathcal{S}_2^R$	Sobolev space (see (2.8)) with norm $\Vert f\Vert_{\mathcal{S}_2^R}^2 := \sum_{\vert\boldsymbol{\alpha}\vert_1 \le R} \Vert D^{\boldsymbol{\alpha}} f\Vert_{\mathcal{L}_2}^2$ (see (2.9))
$\mathcal{S}_\infty^R$	Sobolev space (see (2.8)) with norm $\Vert f\Vert_{\mathcal{S}_\infty^R} := \max_{\vert\boldsymbol{\alpha}\vert_1 \le R} \Vert D^{\boldsymbol{\alpha}} f\Vert_{\mathcal{L}_\infty}$ (see (2.10))
$\overset{0}{\mathcal{S}}{}_p^R$	Sobolev space with zero boundary (see (2.11))
$\mathcal{W}_2^R$	Sobolev space with bounded mixed derivatives (see (2.12)) with norm $\Vert f\Vert_{\mathcal{W}_2^R}^2 := \sum_{\vert\boldsymbol{\alpha}\vert_\infty \le R} \Vert D^{\boldsymbol{\alpha}} f\Vert_{\mathcal{L}_2}^2$ (see (2.13)) and semi-norm $\vert f\vert_{\boldsymbol{\alpha},2} := \Vert D^{\boldsymbol{\alpha}} f\Vert_{\mathcal{L}_2}$ (see (2.15))
$\mathcal{W}_\infty^R$	Sobolev space with bounded mixed derivatives (see (2.12)) with norm $\Vert f\Vert_{\mathcal{W}_\infty^R} := \max_{\vert\boldsymbol{\alpha}\vert_\infty \le R} \Vert D^{\boldsymbol{\alpha}} f\Vert_{\mathcal{L}_\infty}$ (see (2.14)) and semi-norm $\vert f\vert_{\boldsymbol{\alpha},\infty} := \Vert D^{\boldsymbol{\alpha}} f\Vert_{\mathcal{L}_\infty}$ (see (2.16))
$\overset{0}{\mathcal{W}}{}_p^R$	Sobolev space with bounded mixed derivatives and zero boundary (see (2.17))
$\mathcal{C}_\infty^0$	space of continuous functions with norm $\Vert f\Vert_{\mathcal{C}_\infty^0} := \sup_{\mathbf{p}\in\Omega} \Vert f(\mathbf{p})\Vert_{\mathcal{B}}$ (see (2.18))
$\mathcal{C}_\infty^R$	space of R-times continuously differentiable functions with norm $\Vert f\Vert_{\mathcal{C}_\infty^R} := \max_{\vert\boldsymbol{\alpha}\vert_\infty \le R} \Vert D^{\boldsymbol{\alpha}} f\Vert_{\mathcal{C}_\infty^0}$ (see (2.19))

$\mathcal{H}_2$ Hardy space (see (2.21))
with norm $\|\mathbf{F}\|_{\mathcal{H}_2}^2 := \sup_{x>0} \int_{\mathbb{R}} \|\mathbf{F}(x+\imath y)\|_{S,2}^2 \, \mathrm{d}y$ (see (2.20))

$\mathcal{H}_\infty$ Hardy space (see (2.21))
with norm $\|\mathbf{F}\|_{\mathcal{H}_\infty} := \sup_{z\in\mathbb{C}^+} \|\mathbf{F}(z)\|_{S,\infty}$ (see (2.20))

Models and Their Norms

$(\mathbf{E},)$ $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$	system matrices of full order model
$(\mathbf{E}_r,)$ $\mathbf{A}_r$, $\mathbf{B}_r$, $\mathbf{C}_r$	system matrices of reduced order model
$\mathbf{H}(s)$	$= \mathbf{C}(s\mathbf{E}-\mathbf{A})^{-1}\mathbf{B} \in \mathbb{C}^{q\times m}$ (see (2.27)) transfer function of full order model
$\mathbf{H}_r(s)$	$= \mathbf{C}_r(s\mathbf{E}_r-\mathbf{A}_r)^{-1}\mathbf{B}_r \in \mathbb{C}^{q\times m}$ (see (2.29)) transfer function of reduced order model
$\|\mathbf{H}\|_{\mathcal{H}_2}$	$:= (\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\mathbf{H}(s)\|_F^2 \, \mathrm{d}s)^{1/2}$ (see (2.33))
$\|\mathbf{H}\|_{\mathcal{H}_\infty}$	$:= \sup_{s\in\imath\mathbb{R}} \|\mathbf{H}(s)\|_2$ (see (2.34))
$\mathcal{P}$	$= [0,1]^D$ (if not stated otherwise) parameter space
$\mathcal{P}_{\text{test}}$	discretization of parameter space for computation of error
$\mathbf{E}(\mathbf{p})$, $\mathbf{A}(\mathbf{p})$, $\mathbf{B}(\mathbf{p})$, $\mathbf{C}(\mathbf{p})$	system matrices of parametric full order model
$\check{\mathbf{E}}_r(\mathbf{p})$, $\check{\mathbf{A}}_r(\mathbf{p})$, $\check{\mathbf{B}}_r(\mathbf{p})$, $\check{\mathbf{C}}_r(\mathbf{p})$	system matrices of parametric reduced order model
$\mathbf{A}_r(\mathbf{p})$, $\mathbf{B}_r(\mathbf{p})$, $\mathbf{C}_r(\mathbf{p})$	system matrices of transformed reduced order model
$\widetilde{\mathbf{A}}_r(\mathbf{p})$, $\widetilde{\mathbf{B}}_r(\mathbf{p})$, $\widetilde{\mathbf{C}}_r(\mathbf{p})$	system matrices of interpolated reduced order model
$\mathbf{H}(s;\mathbf{p})$	$= \mathbf{C}(\mathbf{p})\cdot(s\mathbf{E}(\mathbf{p})-\mathbf{A}(\mathbf{p}))^{-1}\cdot\mathbf{B}(\mathbf{p})$ (see (3.2)) transfer function of parametric full order model
$\mathbf{H}_r(s;\mathbf{p})$	$= \mathbf{C}_r(\mathbf{p})\cdot(s\mathbf{E}_r(\mathbf{p})-\mathbf{A}_r(\mathbf{p}))^{-1}\cdot\mathbf{B}_r(\mathbf{p})$ (see (3.4)) transfer function of parametric reduced order model
$\widetilde{\mathbf{H}}_r(s;\mathbf{p})$	$= \widetilde{\mathbf{C}}_r(\mathbf{p})\cdot(s\widetilde{\mathbf{E}}_r(\mathbf{p})-\widetilde{\mathbf{A}}_r(\mathbf{p}))^{-1}\cdot\widetilde{\mathbf{B}}_r(\mathbf{p})$ transfer function of interpolated reduced order model
$\|\mathbf{H}\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)}$	$:= (\frac{1}{2\pi}\int_{\mathcal{P}}\int_{\imath\mathbb{R}} \|\mathbf{H}(s;\mathbf{p})\|_F^2 \, \mathrm{d}s \, \mathrm{d}\mathbf{p})^{1/2}$ (see (3.7))
$\|\mathbf{H}\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)}$	$:= \operatorname{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \sup_{s\in\imath\mathbb{R}} \|\mathbf{H}(s;\mathbf{p})\|_2$ (see (3.8))

Grids

$\mathcal{V}_n^{(\infty)}$ $:= \bigoplus_{1\le|\boldsymbol{\ell}|_\infty\le n} \mathcal{W}_{\boldsymbol{\ell}}$
full grid space without boundary (see (2.46))

$\mathcal{V}_n^{(1)}$ $:= \bigoplus_{1\le|\boldsymbol{\ell}|_1-D+1\le n} \mathcal{W}_{\boldsymbol{\ell}}$
sparse grid space without boundary (see (2.48))

$\underline{\mathcal{V}}_n^{(\infty)}$ $:= \bigoplus_{0\le|\boldsymbol{\ell}|_\infty\le n} \mathcal{W}_{\boldsymbol{\ell}}$
full grid space with boundary (see (2.51))

$\underline{\mathcal{V}}_n^{(1)}$ $:= \bigoplus_{0\le|\boldsymbol{\ell}|_1-(D-z(\boldsymbol{\ell}))+1\le n} \mathcal{W}_{\boldsymbol{\ell}}$, $z(\boldsymbol{\ell}) := |\{d\in\{1,\dots,D\} : \ell_d = 0\}|$
sparse grid space with boundary (see (2.53))

1 Introduction

In this thesis we develop and analyze an efficient sparse grid based parametric model order reduction method with stability guarantee for large symmetric linear time-invariant dynamical systems with a higher-dimensional parameter space.

Model order reduction (MOR) is important to speed up simulations by reducing the size of a given model without losing too much accuracy in the quantities of interest. Typical applications are optimal control, very large scale integration chip design, and mechanical engineering. Models are for example electrical circuits or mechanical systems. Often, especially in optimization, they are simulated many times, for instance with varying input or external forces.

Many models depend on parameters such as material or geometry parameters. There is a growing interest in preserving this parameter dependence during model order reduction. This is called parametric model order reduction (PMOR). A parametric model of reduced size is useful for example in parameter studies, which might involve simulation, optimization, and uncertainty quantification.

Most PMOR methods base on a non-parametric MOR method, which is applied at several discrete points in parameter space. Since the number of points is crucial for the performance of those PMOR methods, *efficient* methods should use as few points as possible for a given accuracy. It is not known so far how to choose these points in parameter space optimally, despite for very special cases [BBBG11].

The problem of efficiently choosing good sample points in parameter space becomes even more substantial when the dimension of the parameter space, that is, the number of parameters, is higher than 2 or 3. In such settings the so-called "curse of dimensionality" makes full grids unfeasible. The usage of sparse grids can significantly lessen the exponential dependence on the dimension.

The only PMOR method that is suitable for dealing with higher-dimensional parameter spaces and can preserve stability, an important physical system property, is described in [AF11]. But for a fully automatic PMOR method a concrete MOR method for generating the reduced order models at the fixed points in parameter space is needed. Most PMOR papers do not specify how the MOR step should be done. A common technique in the examples is projection onto a so-called reduced order basis. However, no strategies how to obtain a suitable and accurate basis are suggested.

An often used technique to combine the local non-parametric reduced order models is interpolation over the parameter space. For efficient interpolation it is essential that a smooth parameter dependence of the full order model is preserved by the model order reduction method. This is considered a complicated problem [Maz14] and therefore often one reduced order basis is used for the whole parameter space.

The main disadvantage of this approach is that it is unlikely that one can find a reduced order basis that is an appropriate choice for all parameter points simultaneously. Hence, a large reduction order is required to achieve a good accuracy. Furthermore, in many methods that follow this approach the number of points in parameter space influences the size of the reduced order basis and hence the size of the reduced order model.

We present a PMOR method with a suitable and efficient MOR method inside and an intrinsic stability preservation framework for symmetric systems. It overcomes the aforementioned problems. Our method is called efficient since we minimize the number of interpolation points in parameter space while keeping almost the same accuracy. We achieve this by applying sparse grid interpolation with hierarchical bases. This only works since we compute the local reduced order models of good accuracy with a MOR strategy that preserves smooth parameter dependence.

The theoretical foundations of PMOR have not been given much attention yet. Therefore we introduce appropriate norms for parametric models and formulate general conditions for the parameter dependence of the model such that those norms can be applied.

For our PMOR method we present also theoretical results. We prove that it preserves stability and fulfills interpolation conditions in the joint space of frequency and parameters. Moreover, we derive an upper bound for the interpolation error in parameter space in terms of the transfer function. Furthermore, we prove convergence statements under certain smoothness conditions.

In computational experiments we demonstrate the compelling performance of our PMOR method and the good qualitative behavior of the theoretical findings. We also show that for symmetric systems our PMOR method is competitive with other PMOR methods [BBH$^+$15] even in a one-dimensional parameter space. So PMOR methods based on matrix interpolation can yield good results if they exploit the symmetry.

1.1 Related Work

In this section we give a short overview over previous work about (parametric) model order reduction and sparse grids. We mainly mention methods that are used for or similar to the parametric model order reduction method in this thesis.

1.1.1 Model Order Reduction

In many applications, e.g., micro-electro-mechanical systems, the models that need to be reduced are extremely large, especially when they stem from partial differential equations that are finely discretized in space. Hence, we focus on large-scale dynamical systems. The size r of the reduced order model should be much smaller than the size n of the full order model, for example $r \sim \log(n)$.

There exist many established methods (see [Ant05]), especially for linear time-invariant dynamical systems, where the coefficients are not time-dependent. Well-known examples that are not simulation-based but act directly on the systems are the $\mathcal{H}_2$-optimal Iterative Rational KRYLOV Algorithm [GAB08] and low-rank Balanced Truncation [BBM+06], for which an $\mathcal{H}_\infty$-error bound exists.

The Iterative Rational KRYLOV Algorithm, an interpolatory MOR method, is an obvious choice for our PMOR method, in which we interpolate also in parameter space. For example one can prove that joint interpolation conditions are fulfilled.

1.1.2 Parametric Model Order Reduction

Parametric model order reduction is a quite new and active field of research (see e.g., [BGW15] for a survey of PMOR methods or [BBH+15, BBH+16] for a comparison of PMOR methods including computational experiments). Several PMOR methods exist. However, not all of them are compatible with interpolation on sparse grids or can otherwise be made suitable for a higher-dimensional parameter space. For example, often the size of the parametric reduced order model depends on the number of sample points in parameter space or the method requires sampling on a full grid. Moreover, only a few PMOR methods preserve stability.

PMOR methods which base on interpolation of system matrices are good candidates for PMOR methods which are compatible with interpolation on sparse grids. Recently two PMOR methods where system matrices are interpolated were proposed. However, while the first method preserves stability, only the second method can be combined with sparse grids. Our goal is to have the advantages of both.

In [GPWL14] a stability-preserving PMOR method for arbitrary systems is presented. System matrices at discrete points in parameter space are reduced, transformed and then interpolated. However, there is a restriction on the interpolation method which does not allow for interpolation on sparse grids.

In [GBP+14] another PMOR method for arbitrary systems is proposed. System matrices at discrete points in parameter space are reduced, transformed and then interpolated on matrix manifolds [AF11]. There is no restriction on hte interpolation method, hence interpolation is done on sparse grids. However, stability of the resulting reduced order model is not guaranteed as this is hard to achieve for arbitrary systems. For symmetric systems stability can be preserved by the PMOR method from [AF11] and it is still compatible with interpolation on sparse grids. This is the core of the PMOR method presented in this thesis.

1.1.3 Sparse Grids

Sparse grids have been introduced in [Zen91] for the solution of partial differential equations but reach back to [Smo63] about numerical quadrature. They have also been successfully applied in data analysis, finance and physics [GG13]. See [BG04] for a survey and [Gar13] for a short overview over sparse grids and interpolation on sparse grids with piece-wise linear basis functions.

In *higher-dimensional* spaces, that is dimension up to about 10, interpolation should be done on sparse grids instead of full grids, since sparse grids have much less grid points than full grids, while having comparable interpolation accuracies when the function belongs to a Sobolev space with bounded mixed derivatives [BG04].

In [BNR00] interpolation on sparse grids with global polynomial basis functions is considered. When the function is sufficiently smooth, then the number of interpolation points required to reach some prescribed accuracy is even smaller than for interpolation with piece-wise linear basis functions.

1.2 Contributions of This Thesis

In this thesis we contribute to parametric model order reduction both practically and theoretically.

We develop a fully automatic parametric model order reduction method suitable for and efficient in higher-dimensional parameter spaces, which includes collective model order reduction for a discrete set of parameter points that does not destroy the smooth parameter dependence of the full order model. The PMOR method is designed for symmetric systems such that stability is preserved throughout the method.

In order to achieve a given accuracy we aim at a smallest possible number of points in parameter space. This influences memory consumption and computation time positively.

The PMOR method we develop in this thesis consists of two distinct phases. The first one is the offline or reduction phase, where we perform MOR at discrete points in the parameter space, which results in a set of local reduced order models. The second one is the online or interpolation phase, where we interpolate the system matrices of those reduced order models in the parameter space. The offline phase is for precomputation and is executed only once. It can hence be relatively expensive. The online phase on the other hand is executed many times—each time the parametric reduced order model is evaluated for a given parameter value—and should hence be quite cheap.

Our method is based on the PMOR method described in [AF11] and the $\mathcal{H}_2$-optimal model order reduction method Iterative Rational Krylov Algorithm (IRKA) [GAB08], which is an interpolatory MOR method. We interpolate over the parameter space on sparse grids with piece-wise linear or global polynomial basis functions [BG04, BNR00].

We prove that the resulting model fulfills interpolation conditions in a joint frequency and parameter space since we connect two interpolation methods.

Our method is restricted to symmetric stable systems although we present ideas how to extend it to more general systems in the outlook. We provably preserve the stability of those systems by using one-sided projection in the reduction phase and by interpolating on matrix manifolds.

We formulate model requirements for parametric models and prove that the suggested norms for measuring errors are well-defined.

Preserving a smooth parameter dependence during the reduction phase is essential for an efficient interpolation phase. We propose a strategy based on the IRKA at the center of the parameter space, which preserves the smoothness on the one hand and yields reduced order models of good accuracy on the other hand.

Moreover, we include an additional step that almost halves the interpolation effort for generalized systems. In fact we eliminate one of the two square system matrices.

The method's good performance is presented in several computational experiments. Furthermore, we demonstrate that for symmetric systems PMOR methods that base on interpolation of system matrices can achieve better results than presented in the PMOR method comparison [BBH$^+$15].

The overall error of the proposed method consists of the reduction error and the interpolation error. In this thesis we consider the reduction error as fixed once the size of the reduced order model is chosen. The interpolation error becomes smaller when the number of interpolation points in parameter space is increased. It is important to balance the interpolation error and the reduction error. For this we derive an upper bound for the interpolation error, measured as the transfer function error over the parameter space. It relates the transfer function error to the error in matrices that are interpolated. Although the bound is worse than the actual error by some orders of magnitude, we demonstrate in computational experiments that the bound's qualitative behavior is similar to the error's behavior. Furthermore, we use the bound to prove convergence rates of the interpolation error for piece-wise linear or global polynomial basis functions.

1.3 Sources of Error

Simulation involves various sources of error. In this thesis, we focus on the error resulting from model order reduction and the error resulting from interpolation in parameter space. We aim at balancing both errors. Neither do we want the interpolation error to marginalize the reduction error nor do we want to put enormous efforts into the interpolation phase to achieve an interpolation error that is negligible compared to the reduction error.

The reduction error is considered fixed, given a reduction order r. We focus on the interpolation error, which is balanced with the reduction error to achieve the maximal accuracy with a minimal number of interpolation points.

For the sake of completeness, let us mention further sources of errors that can also be of interest, but that are not considered within this thesis. Errors can stem from

- modeling physical processes,
- linearizing non-linear models,
- discretizing continuous models (in space),
- discretizing (in time) and actually solving the model, and
- computing in finite-precision arithmetic.

1.4 Outline

In chapter 2 we provide fundamental definitions, facts and methods. Furthermore, we introduce model order reduction and interpolation on sparse grids. The subject of chapter 3 is parametric model order reduction, including error measures and methods. We present our PMOR method in chapter 4, explaining all decisions. The chapter 5 is devoted to the derivation of an upper bound for the interpolation error. In chapter 6 we prove convergence rates of the interpolation error for piece-wise linear and global polynomial basis functions. We present the results of our computational experiments and compare our method with other PMOR methods in chapter 7. Furthermore, we examine the quality of our theoretical results. In chapter 8 we give a summary of this thesis, pose open questions and suggest directions for further research.

2 Fundamentals

In this chapter we state essential definitions and introduce model order reduction as well as interpolation on sparse grids.

2.1 Basic Definitions

In this section we define several norms and function spaces. Furthermore, we introduce the concept of functions of matrices.

2.1.1 Norms of Vectors and Matrices

In this section we recapitulate some norms for vectors and matrices together with their properties. These can be found in many textbooks, see for example [GVL89], [HJ12].

Definition 2.1. The *p-norm of a vector* $\mathbf{v} = [v_1, \dots, v_n] \in \mathbb{C}^n$ is defined by

$$\|\mathbf{v}\|_p := \begin{cases} \Big(\sum_{i=1}^{n} |v_i|^p \Big)^{1/p} & \text{for } 1 \leq p < \infty, \\ \max_{i=1:\, n} |v_i| & \text{for } p = \infty. \end{cases}$$

The vector 2-norm is unitarily invariant.

Various *matrix norms* exist. We consider the following class of matrix norms, which are unitarily invariant.

Definition 2.2. The SCHATTEN *p-norm* of a matrix $\mathbf{M} \in \mathbb{C}^{n \times m}$ is defined by

$$\|\mathbf{M}\|_{S,p} := \begin{cases} \Big(\sum_{i=1}^{\min\{n,m\}} \sigma_i^p(\mathbf{M}) \Big)^{1/p} & \text{for } 1 \leq p < \infty, \\ \sigma_{\max}(\mathbf{M}) & \text{for } p = \infty, \end{cases}$$

where $\sigma_i(\mathbf{M})$, $i = 1, \dots, \min\{n, m\}$, are the singular values of the matrix $\mathbf{M}$ (see Theorem A.10).

That is, the SCHATTEN p-norm of a matrix is defined as the p-norm applied to the vector of its singular values.

We focus on SCHATTEN p-norms with $p = 2$ and $p = \infty$, which are better known by their alternative definitions: $\|\mathbf{M}\|_{S,2} = \|\mathbf{M}\|_F$ and $\|\mathbf{M}\|_{S,\infty} = \|\mathbf{M}\|_2$.

The FROBENIUS *norm* of matrix $\mathbf{M} = [M_{ij}]_{i,j=1}^{n,m}$ is defined as

$$\|\mathbf{M}\|_F := \Big(\sum_{i=1}^{n}\sum_{j=1}^{m}|M_{ij}|^2\Big)^{1/2}. \tag{2.1}$$

The *spectral norm* is the matrix 2-norm

$$\|\mathbf{M}\|_2 = \lambda_{\max}(\mathbf{M}\mathbf{M}^H)^{1/2} = \lambda_{\max}(\mathbf{M}^H\mathbf{M})^{1/2} \tag{2.2}$$

and belongs to the *matrix p-norms*, which are induced by the vector p-norms,

$$\|\mathbf{M}\|_p := \sup_{\mathbf{v}\neq\mathbf{0}} \frac{\|\mathbf{M}\mathbf{v}\|_p}{\|\mathbf{v}\|_p}. \tag{2.3}$$

If $\mathbf{M}$ is HERMITIAN, then $\|\mathbf{M}\|_2 = |\lambda|_{\max}(\mathbf{M})$.

All norms above satisfy the *submultiplicativity* property

$$\|\mathbf{M}\mathbf{N}\| \leq \|\mathbf{M}\| \cdot \|\mathbf{N}\|. \tag{2.4}$$

We use this inequality extensively in our analysis without referring to it. For the FROBENIUS norm and a matrix $\mathbf{M} \in \mathbb{C}^{r\times r}$ which is unitarily diagonalizable we proved an improvement of this inequality (see Lemma 5.18).

The following inequalities are useful to relate the matrix spectral norm and the FROBENIUS norm, which are equivalent

$$\|\mathbf{M}\|_2 \leq \|\mathbf{M}\|_F \leq \sqrt{\min\{n,m\}} \cdot \|\mathbf{M}\|_2. \tag{2.5}$$

2.1.2 Multi-indices

In this section we introduce multi-indices based on [BG04, Section 3.1] and [RR93, Section 2.1.1]. A multi-index is essentially an integer vector with non-negative entries and some special notational conventions.

For a D-dimensional *multi-index* $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_D] \in \mathbb{N}_0^D$ we define the norms (cf. vector p-norms for $p = 1$ and $p = \infty$)

$$|\boldsymbol{\alpha}|_1 := \sum_{j=1}^{D} \alpha_j,$$
$$|\boldsymbol{\alpha}|_\infty := \max_{j=1:\,D} \alpha_j.$$

Multi-indices simplify the notation of a partial derivative of a function f defined on a subset of $\mathbb{R}^D$, which can be written as

$$D^{\boldsymbol{\alpha}} f := \frac{\partial^{|\boldsymbol{\alpha}|_1} f}{\partial x_1^{\alpha_1} \ldots \partial x_D^{\alpha_D}}.$$

For multi-indices we define the component-wise arithmetic operations

$$\boldsymbol{\alpha} \cdot \boldsymbol{\beta} := (\alpha_1\beta_1, \ldots, \alpha_D\beta_D)$$

and

$$2^{\boldsymbol{\alpha}} := (2^{\alpha_1}, \ldots, 2^{\alpha_D})$$

and the relational operator

$$\boldsymbol{\alpha} \leq \boldsymbol{\beta} \quad :\Longleftrightarrow \quad \alpha_i \leq \beta_i \text{ for } i = 1, \ldots, D.$$

2.1.3 Lebesgue–Bochner Spaces and Sobolev Spaces

Lebesgue spaces for scalar-valued functions can be found in many analysis textbooks (see, e.g., [Rud74]). We deal with a more general concept, the so-called Lebesgue–Bochner spaces (see, e.g., [Zaa67]), which are Lebesgue spaces of functions whose values lie in a Banach space.

Sobolev spaces for scalar-valued functions can be found in many textbooks about partial differential equations (see, e.g., [RR93]). We deal with a more general variant that bases on Lebesgue–Bochner spaces.

In this thesis, the following Banach spaces $\mathcal{B}$ are used:

- scalars: $\mathbb{R}$ (with $|\cdot|$),
- vectors: $\mathbb{R}^n$ with $\|\cdot\|_p$,
- matrices: $\mathbb{R}^{n\times m}$ with $\|\cdot\|_{S,p}$ for $p \in \{2, \infty\}$ (i.e., $\|\cdot\|_F$ and $\|\cdot\|_2$),
- transfer functions ($\overline{\mathbb{C}^+} \to \mathbb{C}^{q\times m}$): $\mathcal{H}_p$ with $\|\cdot\|_{\mathcal{H}_p}$ for $p \in \{2, \infty\}$.

In this thesis, the domain Ω is one of the following:

- time: $\mathbb{R}$, $\mathbb{R}_+$ or $\mathbb{R}_-$ (only for $\mathcal{L}_p$),
- parameter: $\mathcal{P} \subset \mathbb{R}^D$ ($\mathcal{P} = [0,1]^D$).

We define the Lebesgue–Bochner *spaces*

$$\mathcal{L}_2 = \mathcal{L}_2(\Omega, \mathcal{B}) := \left\{ f : \Omega \to \mathcal{B} \mid \int_\Omega \|f(\mathbf{p})\|_{\mathcal{B}}^2 \,\mathrm{d}\mathbf{p} < \infty \right\} \tag{2.6a}$$

$$\text{with } \|f\|_{\mathcal{L}_2(\Omega,\mathcal{B})}^2 := \int_\Omega \|f(\mathbf{p})\|_{\mathcal{B}}^2 \,\mathrm{d}\mathbf{p} \tag{2.6b}$$

and

$$\mathcal{L}_\infty = \mathcal{L}_\infty(\Omega, \mathcal{B}) := \{ f : \Omega \to \mathcal{B} \mid f \text{ essentially bounded} \} \tag{2.7a}$$

$$\text{with } \|f\|_{\mathcal{L}_\infty(\Omega,\mathcal{B})} := \operatorname*{ess\,sup}_{\mathbf{p}\in\Omega} \|f(\mathbf{p})\|_{\mathcal{B}}. \tag{2.7b}$$

For $\Omega = \mathcal{P}$ we include weak derivatives.

The weak derivative of a vector- or matrix-valued function is the vector or matrix of weak derivatives of its scalar-valued entry functions, respectively.

For $p \in \{2, \infty\}$ and $R \in \mathbb{N}_0$ we define the SOBOLEV *spaces*

$$\mathcal{S}_p^R = \mathcal{S}_p^R(\Omega, \mathcal{B}) := \{f \in \mathcal{L}_p \mid D^{\boldsymbol{\alpha}} f \in \mathcal{L}_p \text{ for } |\boldsymbol{\alpha}|_1 \leq R\} \tag{2.8}$$

with the norms

$$\|f\|_{\mathcal{S}_2^R}^2 := \sum_{|\boldsymbol{\alpha}|_1 \leq R} \|D^{\boldsymbol{\alpha}} f\|_{\mathcal{L}_2}^2, \tag{2.9}$$

$$\|f\|_{\mathcal{S}_\infty^R} := \max_{|\boldsymbol{\alpha}|_1 \leq R} \|D^{\boldsymbol{\alpha}} f\|_{\mathcal{L}_\infty}. \tag{2.10}$$

When the functions are zero on the boundary they belong to

$$\overset{0}{\mathcal{S}}{}_p^R = \overset{0}{\mathcal{S}}{}_p^R(\Omega, \mathcal{B}) := \left\{f \in \mathcal{S}_p^R \mid f|_{\partial\Omega} = 0\right\}. \tag{2.11}$$

In this thesis we mostly consider SOBOLEV spaces with bounded mixed derivatives. These are SOBOLEV spaces where the usual $|\boldsymbol{\alpha}|_1 \leq R$ is replaced by $|\boldsymbol{\alpha}|_\infty \leq R$ (as in [BG04]). For $p \in \{2, \infty\}$ and $R \in \mathbb{N}_0$ we define the SOBOLEV *spaces with bounded mixed derivatives*

$$\mathcal{W}_p^R = \mathcal{W}_p^R(\Omega, \mathcal{B}) := \{f \in \mathcal{L}_p \mid D^{\boldsymbol{\alpha}} f \in \mathcal{L}_p \text{ for } |\boldsymbol{\alpha}|_\infty \leq R\} \tag{2.12}$$

with the norms

$$\|f\|_{\mathcal{W}_2^R}^2 := \sum_{|\boldsymbol{\alpha}|_\infty \leq R} \|D^{\boldsymbol{\alpha}} f\|_{\mathcal{L}_2}^2, \tag{2.13}$$

$$\|f\|_{\mathcal{W}_\infty^R} := \max_{|\boldsymbol{\alpha}|_\infty \leq R} \|D^{\boldsymbol{\alpha}} f\|_{\mathcal{L}_\infty}, \tag{2.14}$$

and the semi-norms

$$|f|_{\boldsymbol{\alpha},2} := \|D^{\boldsymbol{\alpha}} f\|_{\mathcal{L}_2}, \tag{2.15}$$

$$|f|_{\boldsymbol{\alpha},\infty} := \|D^{\boldsymbol{\alpha}} f\|_{\mathcal{L}_\infty}, \tag{2.16}$$

respectively. In this thesis we mostly have $R = 2$.

When the functions are zero on the boundary they belong to

$$\overset{0}{\mathcal{W}}{}_p^R = \overset{0}{\mathcal{W}}{}_p^R(\Omega, \mathcal{B}) := \left\{f \in \mathcal{W}_p^R \mid f|_{\partial\Omega} = 0\right\}. \tag{2.17}$$

2.1.4 Spaces of Continuous and Continuously Differentiable Functions

In this section we introduce spaces of continuous and continuously differentiable functions for matrix-valued functions. For this, we generalize the definitions of [BNR00] for scalar-valued functions to BANACH-valued functions. A vector- or matrix-valued function is continuous if all its scalar-valued entry functions are continuous. The derivative of a vector- or matrix-valued function is the vector or matrix of derivatives of its scalar-valued entry functions, respectively.

We define the *space of continuous functions*

$$\mathcal{C}_\infty^0 = \mathcal{C}_\infty^0(\Omega, \mathcal{B}) := \{f : \Omega \to \mathcal{B} \mid f \text{ continuous}\} \tag{2.18a}$$

$$\text{with } \|f\|_{\mathcal{C}_\infty^0} := \sup_{\mathbf{p}\in\Omega} \|f(\mathbf{p})\|_{\mathcal{B}} \tag{2.18b}$$

and the *space of R-times continuously differentiable functions*

$$\mathcal{C}_\infty^R = \mathcal{C}_\infty^R(\Omega, \mathcal{B}) := \{f : \Omega \to \mathcal{B} \mid D^{\boldsymbol{\alpha}} f \text{ continuous for } |\boldsymbol{\alpha}|_\infty \le R\} \tag{2.19a}$$

$$\text{with } \|f\|_{\mathcal{C}_\infty^R} := \max_{|\boldsymbol{\alpha}|_\infty \le R} \|D^{\boldsymbol{\alpha}} f\|_{\mathcal{C}_\infty^0}. \tag{2.19b}$$

The ∞-symbol in the notation $\mathcal{C}_\infty^0$ and $\mathcal{C}_\infty^R$ is not necessary, because we do not consider spaces $\mathcal{C}_p^0$ and $\mathcal{C}_p^R$ for $p \ne \infty$. But we added it to stress the connection with the Lebesgue–Bochner and Sobolev spaces of $p = \infty$.

2.1.5 Hardy Spaces

In this section we introduce Hardy spaces of matrix-valued functions following [Ant05, Section 5.1.3].

For a matrix-valued function $\mathbf{F} : \mathbb{C} \to \mathbb{C}^{q\times m}$, which is analytic on $\mathbb{C}^+$, we define the norm

$$\|\mathbf{F}\|_{\mathcal{H}_p} := \begin{cases} \left(\sup\limits_{x>0} \int_{\mathbb{R}} \|\mathbf{F}(x + \imath y)\|_p^p \, \mathrm{d}y\right)^{1/p} & \text{for } 1 \le p < \infty, \\ \sup\limits_{s\in\mathbb{C}^+} \|\mathbf{F}(s)\|_p & \text{for } p = \infty, \end{cases} \tag{2.20}$$

where $\|\mathbf{F}(s_0)\|_p$ is chosen to be the Schatten p-norm (see Definition 2.2), and get the Hardy *space*

$$\mathcal{H}_p = \mathcal{H}_p(\mathbb{C}^+; \mathbb{C}^{q\times m}) := \left\{\mathbf{F} : \mathbb{C} \to \mathbb{C}^{q\times m} \mid \|\mathbf{F}\|_{\mathcal{H}_p} < \infty\right\}. \tag{2.21}$$

When, in addition, the function $\mathbf{F}$ is continuous on $\overline{\mathbb{C}^+}$, the maximum is attained on the boundary and the norm definitions (2.20) for $p = 2$ and $p = \infty$ become

$$\|\mathbf{F}\|_{\mathcal{H}_2} = \left(\int_{\imath\mathbb{R}} \|\mathbf{F}(s)\|_F^2 \, \mathrm{d}s\right)^{1/2}, \tag{2.22}$$

$$\|\mathbf{F}\|_{\mathcal{H}_\infty} = \sup_{s\in\imath\mathbb{R}} \|\mathbf{F}(s)\|_2. \tag{2.23}$$

2.1.6 Functions of Matrices

In this section we deal with functions of matrices [Hig08], meaning generalizations of scalar functions which map a matrix to a matrix of the same dimensions (but not entry-wise).

These functions should not be confused with matrix-valued functions, which map an argument from a space of any dimension to a matrix.

For the ease of definition and handling we restrict ourselves to diagonalizable matrices and primary matrix functions.

Definition 2.3 (*function of a matrix*)**.** The function f of a diagonalizable matrix $\mathbf{M} \in \mathbb{C}^{n\times n}$ with $\mathbf{X}^{-1}\mathbf{M}\mathbf{X} = \mathbf{D} = \operatorname{diag}(\lambda_1, \dots, \lambda_r)$ is defined by $f(\mathbf{M}) := \mathbf{X} \cdot f(\mathbf{D}) \cdot \mathbf{X}^{-1}$, where $f(\mathbf{D}) = \operatorname{diag}(f(\lambda_1), \dots, f(\lambda_n))$.

We list some properties that are needed later in this thesis.

Theorem 2.4 ([Hig08, Theorem 1.13])**.** *Let $\mathbf{M} \in \mathbb{C}^{n\times n}$ and let f be defined on the spectrum of the matrix $\mathbf{M}$. Then*

(a) $f(\mathbf{M})$ commutes with $\mathbf{M}$;

(b) the eigenvalues of $f(\mathbf{M})$ are $f(\lambda_i)$, where the λ_i are the eigenvalues of $\mathbf{M}$.

For example, when we have a scalar rational function $f(z)$ to achieve the corresponding matrix function $f(\mathbf{M})$, the scalar variable z is substituted by $\mathbf{M}$, division is replaced by matrix inversion and 1 replaced by the identity matrix $\mathbf{I}$, i.e., $f(z) = 1/(s - z)$ becomes $f(\mathbf{M}) = (s\mathbf{I} - \mathbf{M})^{-1}$.

Three matrix functions are of particular interest in this thesis: the matrix exponential, the matrix logarithm and the matrix square root.

Definition 2.5 ([Hig08, (10.1)])**.** The *matrix exponential* of a matrix $\mathbf{M} \in \mathbb{C}^{n\times n}$ is defined by

$$e^{\mathbf{M}} := \sum_{k=0}^{\infty} \frac{\mathbf{M}^k}{k!}.$$

Theorem 2.6 ([Hig08, Theorem 1.31])**.** *Let $\mathbf{M} \in \mathbb{C}^{n\times n}$ have no eigenvalues on $\mathbb{R}_0^-$. Then there is a unique logarithm $\mathbf{X}$ of $\mathbf{M}$ all of whose eigenvalues lie in the strip $\{z : -\pi < \operatorname{Im}(z) < \pi\}$. We refer to $\mathbf{X}$ as the* (principal) matrix logarithm *of $\mathbf{M}$ and write $\mathbf{X} = \log(\mathbf{M})$. If $\mathbf{M}$ is real then $\log(\mathbf{M})$ is real.*

Theorem 2.7 ([Hig08, Theorem 1.29])**.** *Let $\mathbf{M} \in \mathbb{C}^{n\times n}$ have no eigenvalues on $\mathbb{R}_0^-$. There is a unique square root $\mathbf{X}$ of $\mathbf{M}$ all of whose eigenvalues lie in the open right half-plane, and it is a primary matrix function of $\mathbf{M}$. We refer to $\mathbf{X}$ as the* (principal) matrix square root *of $\mathbf{M}$ and write $\mathbf{X} = \mathbf{M}^{1/2}$. If $\mathbf{M}$ is real then $\mathbf{M}^{1/2}$ is real.*

Corollary 2.8 ([Hig08, Corollary 1.30])**.** *A* HERMITIAN *positive definite matrix has a unique* HERMITIAN *positive definite matrix square root.*

Corollary 2.9. *A symmetric positive definite matrix has a unique symmetric positive definite matrix square root.*

2.2 Model Order Reduction

In this section we present an introduction to (non-parametric) models and model order reduction. The book of Antoulas [Ant05] provides a thorough treatment of large linear time-invariant dynamical systems and model order reduction. The survey [BGW15] deals with parametric model order reduction. However, it also gives a good overview over model order reduction.

2.2.1 Model

Input–state–output systems model the interaction over time $t \geq 0$ between the input and output via an internal state (see Figure 2.1).

$$\begin{array}{ccc} \text{input} & \text{state} & \text{output} \\ \longrightarrow & \boxed{\text{system}} & \longrightarrow \\ \mathbf{u}(t) \in \mathbb{R}^m & \mathbf{x}(t) \in \mathbb{R}^n & \mathbf{y}(t) \in \mathbb{R}^q \end{array}$$

Figure 2.1: Input–state–output model.

The vector-valued input function $u : \mathbb{R}_0^+ \to \mathbb{R}^m$ can be a control or an excitation. The vector-valued state function $x : \mathbb{R}_0^+ \to \mathbb{R}^n$ is only an internal variable. The vector-valued output function $y : \mathbb{R}_0^+ \to \mathbb{R}^q$ can be an observation or a measurement.

The interaction of input, state and output is described by the two equations

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \\ \mathbf{y}(t) &= \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t)) \end{aligned}$$

together with an initial condition $\mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{R}^n$.

We assume that the functions $\mathbf{f}$ and $\mathbf{g}$ are linear with time-invariant coefficients, so we can write the system as

$$\dot{\mathbf{x}}(t) = \mathbf{A}\ \mathbf{x}(t) + \mathbf{B}\ \mathbf{u}(t), \tag{2.24a}$$

$$\mathbf{y}(t) = \mathbf{C}\ \mathbf{x}(t), \tag{2.24b}$$

with constant matrices $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$ and $\mathbf{C} \in \mathbb{R}^{q\times n}$. Including an invertible mass matrix $\mathbf{E} \in \mathbb{R}^{n\times n}$ we obtain our model, a so-called *linear time-invariant (LTI) (continuous-time) system* in generalized state space form

$$\mathbf{E}\ \dot{\mathbf{x}}(t) = \mathbf{A}\ \mathbf{x}(t) + \mathbf{B}\ \mathbf{u}(t), \tag{2.25a}$$

$$\mathbf{y}(t) = \mathbf{C}\ \mathbf{x}(t), \tag{2.25b}$$

with constant matrices $\mathbf{A}$, $\mathbf{E} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$ and $\mathbf{C} \in \mathbb{R}^{q\times n}$.

Since we require the matrix $\mathbf{E}$ to be invertible, the system (2.25) could actually be transformed into a system of the standard form (2.24). However, this is unadvisable for several reasons. First, replacing the matrix $\mathbf{A}$ by $\mathbf{E}^{-1}\mathbf{A}$ might destroy the sparsity of the matrix $\mathbf{A}$. Second, it might cause unnecessary problems in model order reduction, when the matrix $\mathbf{E}$ is not reduced separately.

The first equation (2.25a) is called the *state equation*. It is a system of first-order linear differential equations. The second equation (2.25b) is called the *output equation*. It is a system of linear algebraic equations.

In many physical systems the original model is a partial differential equation, which is first discretized in space to result in a system of ordinary differential equations. There are also models which are discrete from the beginning, e.g., multi-body dynamics in mechanical systems simulation and electrical circuit simulation for chip design [MS05].

The dimension n of the state space is called the *order* of the system. For very complex systems, e.g., when the system is a finite element discretization, the order n is large and the square matrices $\mathbf{A}$ and $\mathbf{E}$ are sparse.

When $m = q = 1$, we call (2.25) a *single-input, single-output (SISO) system* and a *multi-input, multi-output (MIMO) system* otherwise.

We consider systems where the dimensions of the input and output space, m and q, respectively, are small compared to the large dimension n of the state space. Hence, the matrix $\mathbf{B}$ is tall rectangular and the matrix $\mathbf{C}$ is wide rectangular. In the SISO case, the matrices $\mathbf{B}$ and $\mathbf{C}$ become a column and a row vector, respectively.

Definition 2.10. A system (2.25) is called *(asymptotically) stable* if all eigenvalues of $\mathbf{E}^{-1}\mathbf{A}$ lie in the open left half of the complex plane, i.e., $\mathrm{Re}(\lambda) < 0$ for all $\lambda \in \Lambda(\mathbf{E}^{-1}\mathbf{A})$.

Equivalently, a system is asymptotically stable if $\lim_{t\to\infty} \mathbf{x}(t) = \mathbf{0}$ for all solutions $\mathbf{x}(t)$ of $\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$. In other words, the solution $\mathbf{x}(t)$ of the state equation (2.25a) tends to zero when there is no input $\mathbf{u}(t)$.

Definition 2.11. The eigenvalues of $\mathbf{E}^{-1}\mathbf{A}$ are called the *poles* of the system (2.25).

Applying the LAPLACE *transform* (see Definition A.27) to system (2.25) with initial condition $\mathbf{x}(0) = \mathbf{x}_0 = \mathbf{0}$ we get

$$\begin{aligned} \mathbf{E}\; s\widehat{\mathbf{x}}(s) &= \mathbf{A}\;\widehat{\mathbf{x}}(s) + \mathbf{B}\;\widehat{\mathbf{u}}(s), \\ \widehat{\mathbf{y}}(s) &= \mathbf{C}\;\widehat{\mathbf{x}}(s), \end{aligned}$$

where the time t is replaced by the frequency variable s. We eliminate the (LAPLACE transformed) state vector $\widehat{\mathbf{x}}(s)$ and get a linear relation between the (LAPLACE transformed) input $\widehat{\mathbf{u}}(s)$ and the (LAPLACE transformed) output $\widehat{\mathbf{y}}(s)$. We have $\widehat{\mathbf{y}}(s) = \mathbf{H}(s)\widehat{\mathbf{u}}(s)$, where

$$\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} \in \mathbb{C}^{q\times m} \quad \text{for } s \in \overline{\mathbb{C}^+} \tag{2.27}$$

is the *transfer function* of system (2.25).

The transfer function of a stable system is analytic on $\mathbb{C}^+$, bounded and continuous on $\overline{\mathbb{C}^+}$ and hence belongs to the HARDY spaces (see subsection 2.1.5) $\mathcal{H}_\infty(\mathbb{C};\mathbb{C}^{q\times m})$ [HP05, page 148] and $\mathcal{H}_2(\mathbb{C};\mathbb{C}^{q\times m})$ [ABG10, page 16].

The values of the transfer function $\mathbf{H} : \overline{\mathbb{C}^+} \to \mathbb{C}^{q\times m}$ for a given $s \in \overline{\mathbb{C}^+}$ are matrices, whose sizes depend on the dimensions of the input and output space, m and q, respectively. In the SISO case, the transfer function is scalar-valued.

When we transform the state vector $\mathbf{x}(t) \in \mathbb{R}^n$ with an invertible matrix $\mathbf{T} \in \mathbb{R}^{n\times n}$, i.e., change coordinates $\widetilde{\mathbf{x}}(t) = \mathbf{T}\mathbf{x}(t)$, the system (2.25) becomes

$$\begin{aligned} \mathbf{TET}^{-1}\,\dot{\widetilde{\mathbf{x}}}(t) &= \mathbf{TAT}^{-1}\,\widetilde{\mathbf{x}}(t) + \mathbf{TB}\,\mathbf{u}(t),\\ \mathbf{y}(t) &= \mathbf{CT}^{-1}\,\widetilde{\mathbf{x}}(t). \end{aligned}$$

We denote the transformed system matrices $\widetilde{\mathbf{E}} = \mathbf{TET}^{-1}$, $\widetilde{\mathbf{A}} = \mathbf{TAT}^{-1}$, $\widetilde{\mathbf{B}} = \mathbf{TB}$ and $\widetilde{\mathbf{C}} = \mathbf{CT}^{-1}$. The transformed system $(\widetilde{\mathbf{E}}, \widetilde{\mathbf{A}}, \widetilde{\mathbf{B}}, \widetilde{\mathbf{C}})$ is *equivalent* to the original system $(\mathbf{E}, \mathbf{A}, \mathbf{B}, \mathbf{C})$, since the transfer functions are the same

$$\begin{aligned} \mathbf{H}(s) &= \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\\ &= \mathbf{CT}^{-1}\mathbf{T}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{T}^{-1}\mathbf{TB}\\ &= \mathbf{CT}^{-1}(s\mathbf{TET}^{-1} - \mathbf{TAT}^{-1})^{-1}\mathbf{TB}\\ &= \widetilde{\mathbf{C}}(s\widetilde{\mathbf{E}} - \widetilde{\mathbf{A}})^{-1}\widetilde{\mathbf{B}}\\ &= \widetilde{\mathbf{H}}(s). \end{aligned}$$

2.2.2 Model Order Reduction by Projection

The goal of model order reduction is to replace the *full order model (FOM)* (2.25) with the *full order* n by a model of smaller order $r \ll n$. The so-called *reduced order model (ROM)* with the *reduction order* r is denoted

$$\mathbf{E}_r\,\dot{\mathbf{x}}_r(t) = \mathbf{A}_r\,\mathbf{x}_r(t) + \mathbf{B}_r\,\mathbf{u}(t), \tag{2.28a}$$

$$\mathbf{y}_r(t) = \mathbf{C}_r\,\mathbf{x}_r(t), \tag{2.28b}$$

with constant matrices $\mathbf{A}_r$, $\mathbf{E}_r \in \mathbb{R}^{r\times r}$, $\mathbf{B}_r \in \mathbb{R}^{r\times m}$ and $\mathbf{C}_r \in \mathbb{R}^{q\times r}$. The output error $\|\mathbf{y} - \mathbf{y}_r\|$ should be small. We do not care about the state error $\|\mathbf{x} - \mathbf{x}_r\|$, since we are interested in the input–output behavior of the systems.

The *transfer function of the reduced order model* (2.28) is

$$\mathbf{H}_r(s) = \mathbf{C}_r\,(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\,\mathbf{B}_r \in \mathbb{C}^{q\times m} \quad \text{for } s \in \overline{\mathbb{C}^+}. \tag{2.29}$$

We obtain the system matrices $\mathbf{A}_r$, $\mathbf{E}_r$, $\mathbf{B}_r$, $\mathbf{C}_r$ of the reduced order model from the system matrices $\mathbf{A}$, $\mathbf{E}$, $\mathbf{B}$, $\mathbf{C}$ of the full order model (2.25) with the *projection framework*.

The idea is to map the state onto a lower-dimensional subspace $\mathcal{V} \subset \mathbb{R}^n$ of dimension r, represented by the matrix $\mathbf{V} \in \mathbb{R}^{n\times r}$ with $\operatorname{rank}(\mathbf{V}) = r$, whose orthonormal column vectors constitute the *reduced order basis (ROB)* that spans $\mathcal{V}$. The approximation $\mathbf{x}(t) \approx \mathbf{V}\,\mathbf{x}_r(t)$ leads to the residual

$$\mathbf{E}\mathbf{V}\,\dot{\mathbf{x}}_r(t) - \mathbf{A}\mathbf{V}\,\mathbf{x}_r(t) - \mathbf{B}\,\mathbf{u}(t),$$

which we force to be orthogonal to another subspace $\mathcal{W} \subset \mathbb{R}^n$ of dimension r, represented by the matrix $\mathbf{W} \in \mathbb{R}^{n\times r}$ with $\operatorname{rank}(\mathbf{W}) = r$,

$$\mathbf{W}^T \left(\mathbf{E}\mathbf{V}\,\dot{\mathbf{x}}_r(t) - \mathbf{A}\mathbf{V}\,\mathbf{x}_r(t) - \mathbf{B}\,\mathbf{u}(t)\right) \overset{!}{=} 0.$$

So finally we get a reduced order model with the system matrices

$$\begin{aligned} \mathbf{E}_r &= \mathbf{W}^T\mathbf{E}\mathbf{V}, \\ \mathbf{A}_r &= \mathbf{W}^T\mathbf{A}\mathbf{V}, \\ \mathbf{B}_r &= \mathbf{W}^T\mathbf{B} \quad \text{and} \\ \mathbf{C}_r &= \mathbf{C}\mathbf{V}. \end{aligned} \tag{2.30}$$

If $\mathbf{V} = \mathbf{W}$ (and hence $\mathcal{V} = \mathcal{W}$), then we call it *one-sided projection* and *two-sided projection* otherwise.

When the system matrices $\mathbf{A}$ and $\mathbf{E}$ are symmetric, one-sided projection ensures that the reduced system matrices $\mathbf{A}_r$ and $\mathbf{E}_r$ are symmetric, too.

2.2.3 Error Measures

We want to examine the input–output behavior of the system and not the state approximation, so we consider the output error and not the state error.

The output error in the time domain is measured by

$$\|\mathbf{y} - \mathbf{y}_r\|_{\mathcal{L}_2} = \left(\int_0^\infty \|\mathbf{y}(t) - \mathbf{y}_r(t)\|_2^2 \;\mathrm{d}t\right)^{1/2} \tag{2.31}$$

or

$$\|\mathbf{y} - \mathbf{y}_r\|_{\mathcal{L}_\infty} = \sup_{t\geq 0} \|\mathbf{y}(t) - \mathbf{y}_r(t)\|_\infty. \tag{2.32}$$

We are interested in an error measure that allows us to check the output error not only for a particular input, but for a range of inputs at once. For this we use the transfer functions (2.27) and (2.29) of the models (2.25) and (2.28). They belong to the Hardy spaces $\mathcal{H}_2$ and $\mathcal{H}_\infty$ (see subsection 2.2.1) with the $\mathcal{H}_2$-*norm*

$$\|\mathbf{H}\|_{\mathcal{H}_2} := \left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\mathbf{H}(s)\|_F^2 \;\mathrm{d}s\right)^{1/2}, \tag{2.33}$$

where $\|\cdot\|_F$ denotes the Frobenius *norm* (see (2.1)), and the $\mathcal{H}_\infty$-*norm*

$$\|\mathbf{H}\|_{\mathcal{H}_\infty} := \sup_{s\in\imath\mathbb{R}} \|\mathbf{H}(s)\|_2. \tag{2.34}$$

The transfer function error (in the frequency domain) is measured by the $\mathcal{H}_2$-norm

$$\|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_2} = \left(\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\mathbf{H}(s) - \mathbf{H}_r(s)\|_F^2 \ \mathrm{d}s \right)^{1/2} \tag{2.35}$$

or the $\mathcal{H}_\infty$-norm

$$\|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_\infty} = \sup_{s \in \imath\mathbb{R}} \|\mathbf{H}(s) - \mathbf{H}_r(s)\|_2. \tag{2.36}$$

In the SISO case, we have

$$\|\mathbf{H}(s) - \mathbf{H}_r(s)\|_F = \|\mathbf{H}(s) - \mathbf{H}_r(s)\|_2 = |\mathbf{H}(s) - \mathbf{H}_r(s)| \text{ for } s \in \imath\mathbb{R}.$$

For a given input $\mathbf{u} \in \mathcal{L}_2(\mathbb{R}_0^+; \mathbb{R}^m)$, it holds (see Theorem B.2)

$$\|\mathbf{y} - \mathbf{y}_r\|_{\mathcal{L}_\infty} \leq \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_2} \cdot \|\mathbf{u}\|_{\mathcal{L}_2} \tag{2.37}$$

and

$$\|\mathbf{y} - \mathbf{y}_r\|_{\mathcal{L}_2} \leq \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_\infty} \cdot \|\mathbf{u}\|_{\mathcal{L}_2}. \tag{2.38}$$

Hence, the transfer function error is an upper bound of the output error for the whole class of input functions with $\mathcal{L}_2$-norm less than or equal one.

2.2.4 Methods

In this section we introduce an important class of model order reduction methods.

The reduced order model (2.28) should produce a small output error $\|\mathbf{y} - \mathbf{y}_r\|$. Since the output error is related to the transfer function error (see (2.37) and (2.38)) we aim at minimizing the transfer function error $\|\mathbf{H} - \mathbf{H}_r\|$.

In the $\mathcal{H}_\infty$-norm, we can produce a (possibly suboptimal) stable solution with the model order reduction method *Balanced Truncation*. For its solution there is a computable upper bound of the error. So the reduced order r can be chosen a priori in dependence on a prescribed accuracy. The original method [Moo81] involves the solution of two LYAPUNOV equations. This is not tractable for large system. Fortunately, there are some low-rank versions of Balanced Truncation (e.g., [BBM⁺06]), which can be applied to large (and sparse) systems.

In the $\mathcal{H}_2$-norm, we can produce a locally optimal solution with the model order reduction method *Iterative Rational* KRYLOV *Algorithm* [GAB08]. The method is suitable for large systems. The necessary optimality conditions are interpolation conditions for the transfer function. This fits well with the interpolation setting we will establish over the parameter space as we have interpolation conditions over the frequency and the parameter space (see subsection 4.3.3).

Before we turn to the $\mathcal{H}_2$-optimal Iterative Rational KRYLOV Algorithm we introduce the more general class of rational KRYLOV methods, which are interpolatory model order reduction methods. A thorough introduction to interpolatory model order reduction methods can be found in [ABG10].

Rational Krylov Methods

To construct an approximation $\mathbf{H}_r$ of $\mathbf{H}$ we demand that $\mathbf{H}_r$ interpolates $\mathbf{H}$ (and some of its derivatives) at some frequency points.

This interpolation in frequency space should not be confused with the interpolation in parameter space later in this thesis. Here we have no parameters, only model order reduction.

Since transfer functions are in general matrix-valued, we impose tangential interpolation conditions of the form

$$\frac{\mathrm{d}^\ell}{\mathrm{d}s^\ell}\mathbf{H}(\mu)\ \mathbf{b} = \frac{\mathrm{d}^\ell}{\mathrm{d}s^\ell}\mathbf{H}_r(\mu)\ \mathbf{b}, \qquad \mathbf{c}^T\frac{\mathrm{d}^\ell}{\mathrm{d}s^\ell}\mathbf{H}(\nu) = \mathbf{c}^T\frac{\mathrm{d}^\ell}{\mathrm{d}s^\ell}\mathbf{H}_r(\nu),$$

where $\ell \geq 0$, with the right interpolation point μ and tangent direction $\mathbf{b} \in \mathbb{C}^m$ and the left interpolation point ν and tangent direction $\mathbf{c} \in \mathbb{C}^q$, which need to be chosen appropriately. In the next section we will see how they can be selected such that the reduced order model is $\mathcal{H}_2$-optimal. There are other choices which preserve a specific system property or structure (see [ABG10] for details).

Interpolation conditions for the transfer function are linked with conditions for the projection matrices $\mathbf{V}$ and $\mathbf{W}$ (cf. subsection 2.2.2).

Theorem 2.12 ([ABG10, Theorem 1.1.])**.** *Let $\mu, \nu \in \mathbb{C}$ be such that $s\mathbf{E} - \mathbf{A}$ and $s\mathbf{E}_r - \mathbf{A}_r$ are invertible for $s = \mu, \nu$. If $\mathbf{b} \in \mathbb{C}^m$ and $\mathbf{c} \in \mathbb{C}^q$ are fixed non-zero vectors then*

1. *if $(\mu\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathbf{b} \in \operatorname{range}(\mathbf{V})$, then $\mathbf{H}(\mu)\mathbf{b} = \mathbf{H}_r(\mu)\mathbf{b}$;*
2. *if $(\nu\mathbf{E} - \mathbf{A})^{-T}\mathbf{C}^T\mathbf{c} \in \operatorname{range}(\mathbf{W})$, then $\mathbf{c}^T\mathbf{H}(\nu) = \mathbf{c}^T\mathbf{H}_r(\nu)$;*
3. *if both 1 and 2 hold, and $\mu = \nu$, then*

$$\mathbf{c}^T\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{H}(\mu)\ \mathbf{b} = \mathbf{c}^T\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{H}_r(\mu)\ \mathbf{b}.$$

Similar conditions for higher derivatives are stated in [ABG10, Theorem 1.2.].

There are many possible combinations of how many derivatives to match at what interpolation points.

Well-known methods that belong to this class are, e.g., *moment matching* and partial realization. However, those methods are mainly for single-input, single-output systems. In the case of multi-input, multi-output, they involve no tangent directions, so one interpolation condition results in m or q instead of one column-conditions for the projection matrices. Another problem is that it is not obvious how to choose good left and right interpolation points. Mostly, greedy strategies are used.

In the next section we see that the necessary conditions for $\mathcal{H}_2$-optimality are Hermite interpolation conditions at the poles of the reduced order model mirrored at the imaginary axis and fixed point iteration leads to a suitable algorithm.

The Iterative Rational Krylov Algorithm (IRKA)—an $\mathcal{H}_2$-optimal Method

We want to select interpolation points and tangent directions such that

$$\mathbf{H}_r = \underset{\check{\mathbf{H}}_r\text{: stable}}{\operatorname{argmin}} \|\mathbf{H} - \check{\mathbf{H}}_r\|_{\mathcal{H}_2}. \tag{2.39}$$

Theorem 2.13 ([BBBG11, Theorem 3.2.]). *Suppose that the transfer function $\mathbf{H}_r(s) = \mathbf{C}_r\left(s\mathbf{E}_r - \mathbf{A}_r\right)^{-1}\mathbf{B}_r$ solves (2.39) and the associated reduced-order pencil $s\mathbf{E}_r - \mathbf{A}_r$ has distinct eigenvalues λ_i, $i = 1, \ldots, r$. Let $\mathbf{x}_i$ and $\mathbf{y}_i^H$ denote right and left eigenvectors associated with λ_i so that $\mathbf{A}_r\mathbf{x}_i = \lambda_i\mathbf{E}_r\mathbf{x}_i$, $\mathbf{y}_i^H\mathbf{A}_r = \lambda_i\mathbf{y}_i^H\mathbf{E}_r$ and $\mathbf{y}_i^H\mathbf{E}_r\mathbf{x}_j = \delta_{ij}$. Define $\mathbf{b}_i^T = \mathbf{y}_i^H\mathbf{B}_r$ and $\mathbf{c}_i = \mathbf{C}_r\mathbf{x}_i$.*
Then, for $i = 1, \ldots, r$,

$$\mathbf{H}(-\lambda_i)\ \mathbf{b}_i = \mathbf{H}_r(-\lambda_i)\ \mathbf{b}_i, \tag{2.40a}$$

$$\mathbf{c}_i^T\ \mathbf{H}(-\lambda_i) = \mathbf{c}_i^T\ \mathbf{H}_r(-\lambda_i), \qquad \textit{and} \tag{2.40b}$$

$$\mathbf{c}_i^T\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{H}(-\lambda_i)\ \mathbf{b}_i = \mathbf{c}_i^T\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{H}_r(-\lambda_i)\ \mathbf{b}_i. \tag{2.40c}$$

That is, necessary conditions for $\mathcal{H}_2$-optimality are bi-tangential Hermite interpolation conditions.

If we construct the projection matrices $\mathbf{V}$ and $\mathbf{W}$ such that

$$\begin{aligned}\operatorname{range}(\mathbf{V}) &\supseteq \left\{(\mu_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathbf{b}_1, \ldots, (\mu_r\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathbf{b}_r\right\},\\ \operatorname{range}(\mathbf{W}) &\supseteq \left\{(\mu_1\mathbf{E}^T - \mathbf{A}^T)^{-1}\mathbf{C}^T\mathbf{c}_1, \ldots, (\mu_r\mathbf{E}^T - \mathbf{A}^T)^{-1}\mathbf{C}^T\mathbf{c}_r\right\},\end{aligned}$$

with $\mu_i = -\lambda_i$, $\mathbf{b}_i^T = \mathbf{y}_i^H\mathbf{B}_r$ and $\mathbf{c}_i = \mathbf{C}_r\mathbf{x}_i$ for $i = 1, \ldots, r$, then Theorem 2.12 implies that we fulfill the interpolation conditions (2.40) of Theorem 2.13.

The optimal interpolation points are the mirror images of the reduced-order poles. So when we want to compute $\mathbf{V}$ and $\mathbf{W}$, we need knowledge of the reduced order model, which is the result of the full order model with projection by $\mathbf{V}$ and $\mathbf{W}$. So we are going round in circles. A remedy is to use an iterative algorithm.

The resulting *Iterative Rational* Krylov *Algorithm (IRKA)* from [GAB08] was mainly designed for SISO systems, even if the $\mathcal{H}_2$-optimality conditions for MIMO systems (2.40) have also been mentioned (see also [VGA08, BKVW10]). The more general IRKA version from [ABG10, BBBG11] includes $\mathbf{E} \neq \mathbf{I}$ and is stated for MIMO systems (see Algorithm 2.1).

By $z_1, \ldots, z_r$ *closed under conjugation* we mean that viewed as sets $\{z_1, \ldots, z_r\} = \{\overline{z_1}, \ldots, \overline{z_r}\}$. Since everything is closed under conjugation, it can be shown that $\mathbf{V}$ and $\mathbf{W}$ can be chosen real [GAB08, Corollary 2.2].

Convergence of this fixed-point iteration algorithm for a special class of SISO systems is proven in [FBG12]. When the number of both inputs and outputs is large, a residue correction step can improve the convergence of the algorithm (see [BG12]).

Algorithm 2.1 Iterative Rational Krylov Algorithm.

1. Initial r-fold shift selection:
   Choose frequencies $\mu_1, \ldots, \mu_r$ (with $\mathrm{Re}(\mu_i) > 0$) and initial tangent directions $\mathbf{b}_1, \ldots, \mathbf{b}_r$ and $\mathbf{c}_1, \ldots, \mathbf{c}_r$ that are closed under conjugation (see Algorithm 7.1).
2. while not converged:
   a) Set $\mathbf{V} = \left[(\mu_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathbf{b}_1, \ldots, (\mu_r\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathbf{b}_r\right]$
   and $\mathbf{W} = \left[(\mu_1\mathbf{E}^T - \mathbf{A}^T)^{-1}\mathbf{C}^T\mathbf{c}_1, \ldots, (\mu_r\mathbf{E}^T - \mathbf{A}^T)^{-1}\mathbf{C}^T\mathbf{c}_r\right]$.
   b) Set $\mathbf{A}_r = \mathbf{W}^T\mathbf{A}\mathbf{V}$, $\mathbf{E}_r = \mathbf{W}^T\mathbf{E}\mathbf{V}$, $\mathbf{B}_r = \mathbf{W}^T\mathbf{B}$ and $\mathbf{C}_r = \mathbf{C}\mathbf{V}$.
   c) Compute eigenvalues λ_i and associated right and left eigenvectors $\mathbf{x}_i$ and $\mathbf{y}_i^H$ for which $\mathbf{A}_r\mathbf{x}_i = \lambda_i\mathbf{E}_r\mathbf{x}_i$, $\mathbf{y}_i^H\mathbf{A}_r = \lambda_i\mathbf{y}_i^H\mathbf{E}_r$ and $\mathbf{y}_i^H\mathbf{E}_r\mathbf{x}_j = \delta_{ij}$.
   d) Set $\mu_i = -\lambda_i$, $\mathbf{b}_i^T = \mathbf{y}_i^H\mathbf{B}_r$ and $\mathbf{c}_i = \mathbf{C}_r\mathbf{x}_i$ for $i = 1, \ldots, r$.
   e) Set $\mu_i \leftarrow |\mathrm{Re}(\mu_i)| + \imath\,\mathrm{Im}(\mu_i)$.

The IRKA variant of [ZCL13] guarantees stability of the reduced order model. The idea is to mirror the μ_i along the imaginary axis if they have a negative real part. So in Algorithm 2.1 we inserted $\mu_i \leftarrow |\mathrm{Re}(\mu_i)| + \imath\,\mathrm{Im}(\mu_i)$ as step 2e.

If only the transfer function (TF) but no system matrices are available, a combination with the Loewner-matrix framework leads to the TF-IRKA [BG12].

An alternative to the fixed-point iteration algorithm IRKA is a purely optimization-based descent algorithm: Beattie and Gugercin [BG09] proposed an algorithm for $\mathcal{H}_2$-optimal model order reduction, based on the *trust-region* approach. It results in a sequence of reduced order models, whose $\mathcal{H}_2$-error norms are improving monotonically. Furthermore, it is globally convergent to a reduced order model that satisfies the necessary optimality conditions (2.40).

Even though the presented method is $\mathcal{H}_2$-optimal the $\mathcal{H}_\infty$-error of the resulting reduced order model is usually also quite small.

2.3 Interpolation on Sparse Grids

In this section we introduce interpolation on sparse grids.

In contrast to full grids, which suffer from the so-called "curse of dimensionality", sparse grids enable to deal with higher-dimensional spaces because of their weaker (albeit still exponential) dependence on the grid dimension. Some degree of smoothness of the function is essential for efficient interpolation on sparse grids.

We apply sparse grid interpolation with two different types of basis functions. Firstly, we use piece-wise linear basis functions to interpolate functions from the Sobolev space with bounded mixed derivatives for $R = 2$ (see section 6.1). Secondly, we use global polynomial basis functions to interpolate functions from the space of continuously differentiable functions for any $R \geq 0$ (see section 6.2).

The theory in standard sparse grid literature ([BG04, Gar13, BNR00]) is for scalar-valued functions. We deal with matrix-valued functions by treating every entry of the matrix as a scalar function. Later we will see that those functions can be interpolated simultaneously (see subsection 2.3.4), since we can use the same basis functions with matrix coefficients instead of scalar coefficients.

2.3.1 Piece-wise Linear Basis Functions

In this section we follow the lines of [BG04, Section 3] for interpolation of scalar-valued functions on sparse grids without boundary points. For sparse grids with boundary points in the last part of this section we mainly refer to [Feu10, Section 2] and [Gar13, Section 2].

We consider multi-variate functions $f : \Omega \to \mathcal{B}$ for $\Omega = [0,1]^D$ and $\mathcal{B} = \mathbb{R}$ from the SOBOLEV *space with bounded second mixed derivatives* $\mathcal{W}_p^2([0,1]^D, \mathbb{R})$ for $p \in \{2, \infty\}$ (cf. definition (2.12)). So in this section the domain Ω is the D-dimensional unit cube and the BANACH space $\mathcal{B}$ is the set of real numbers. The smoothness parameter is $R = 2$ since we interpolate with piece-wise linear basis functions.

For simplicity we restrict ourselves to functions which are zero on the boundary, i.e., $f \in \overset{0}{\mathcal{W}}{}_p^2([0,1]^D, \mathbb{R})$ (cf. definition (2.17)). At the end of this section we treat non-zero boundary conditions.

The main ingredients for interpolation on sparse grids are the hierarchical basis and the corresponding multi-level splitting of the subspace we use for interpolation. Then the choice of the index set for the levels makes the difference between full and sparse grids.

Hierarchical Multi-level Subspace Splitting

The multi-index $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_D) \in \mathbb{N}^D$ represents the multi-variate level of a grid, a point or a basis function.

We want to apply piece-wise linear interpolation and hence use a hat function basis, which is piece-wise D-linear, i.e., piece-wise linear in each dimension.

Firstly, we introduce the grids for the centers of the hat functions.

We consider the set of rectangular grids

$$\left\{ \Omega_{\boldsymbol{\ell}} : \boldsymbol{\ell} \in \mathbb{N}^D \right\}$$

on Ω with mesh size

$$\mathbf{h}_{\boldsymbol{\ell}} := (h_{\ell_1}, \ldots, h_{\ell_D}) := 2^{-\boldsymbol{\ell}} = (2^{-\ell_1}, \ldots, 2^{-\ell_D})$$

and the grid points

$$\mathbf{x}_{\boldsymbol{\ell},\mathbf{i}} := (x_{\ell_1,i_1}, \ldots, x_{\ell_D,i_D}) := \mathbf{i} \cdot \mathbf{h}_{\boldsymbol{\ell}} = (i_1 h_{\ell_1}, \ldots, i_D h_{\ell_D}), \quad \mathbf{0} \leq \mathbf{i} \leq 2^{\boldsymbol{\ell}},$$

where the multi-index $\mathbf{i}$ indicates the location of a grid point $\mathbf{x}_{\boldsymbol{\ell},\mathbf{i}}$ in the corresponding grid $\Omega_{\boldsymbol{\ell}}$.

Secondly, we define the hat function basis, which spans the space of piece-wise D-linear functions.

The standard one-dimensional *hat function*

$$\phi(x) := \begin{cases} 1 - |x| & \text{for } x \in [-1, 1], \\ 0 & \text{otherwise,} \end{cases}$$

is used to construct arbitrary one-dimensional hat functions for dimension d

$$\phi_{\ell_d, i_d}(x_d) := \phi\left(\frac{x_d - x_{\ell_d, i_d}}{h_{\ell_d}}\right)$$

with support $[x_{\ell_d,i_d} - h_{\ell_d}, x_{\ell_d,i_d} + h_{\ell_d}]$. Now we tensorize the one-dimensional hat functions to generate D-dimensional hat functions (see Figure 2.2 for a two-dimensional example)

$$\phi_{\boldsymbol{\ell},\mathbf{i}}(x_1, \ldots, x_D) := \prod_{d=1}^{D} \phi_{\ell_d,i_d}(x_d).$$

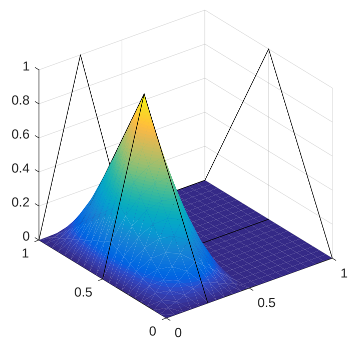


Figure 2.2: The two-dimensional bi-linear basis function $\phi_{(2,1),(1,1)} = \phi_{2,1} \cdot \phi_{1,1}$.

We first consider functions which are zero on the boundary, arbitrary functions will be dealt with later. Hence, we only use hat functions that belong to inner grid points to define the space

$$\mathcal{V}_{\boldsymbol{\ell}} := \operatorname{span}\left\{\phi_{\boldsymbol{\ell},\mathbf{i}} : \mathbf{i} \in \mathcal{I}_{\boldsymbol{\ell}}^{\text{all}}\right\},$$

with the index set

$$\mathcal{I}_{\boldsymbol{\ell}}^{\text{all}} := \left\{\mathbf{i} \in \mathbb{N}^D : \mathbf{1} \leq \mathbf{i} \leq 2^{\boldsymbol{\ell}} - \mathbf{1}\right\}.$$

The set of hat functions $\{\phi_{\boldsymbol{\ell},\mathbf{i}} : \mathbf{i} \in \mathcal{I}_{\boldsymbol{\ell}}^{\text{all}}\}$ is the so-called *nodal basis* of the finite-dimensional space $\mathcal{V}_{\boldsymbol{\ell}}$.

We define the index set

$$\mathcal{I}_{\boldsymbol{\ell}}^{\text{odd}} := \left\{\mathbf{i} \in \mathcal{I}_{\boldsymbol{\ell}}^{\text{all}} : i_d \text{ odd for } d = 1, \dots, D\right\} \tag{2.41}$$

$$= \left\{\mathbf{i} \in \mathbb{N}^D : \mathbf{1} \le \mathbf{i} \le 2^{\boldsymbol{\ell}} - \mathbf{1}, i_d \text{ odd for } d = 1, \dots, D\right\}, \tag{2.42}$$

so the supports of the basis functions $\{\phi_{\boldsymbol{\ell},\mathbf{i}} : \mathbf{i} \in \mathcal{I}_{\boldsymbol{\ell}}^{\text{odd}}\}$ are pair-wise disjoint.

Now, we can split the space

$$\mathcal{V}_{\boldsymbol{\ell}} = \bigoplus_{\mathbf{1} \le \mathbf{k} \le \boldsymbol{\ell}} \mathcal{W}_{\mathbf{k}}$$

into *hierarchical difference spaces* (also called *hierarchical increments*)

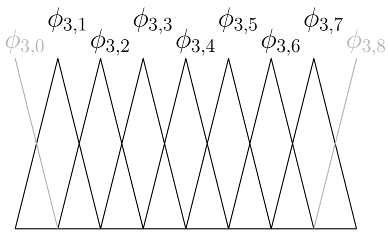
$$\mathcal{W}_{\boldsymbol{\ell}} := \operatorname{span}\left\{\phi_{\boldsymbol{\ell},\mathbf{i}} : \mathbf{i} \in \mathcal{I}_{\boldsymbol{\ell}}^{\text{odd}}\right\}.$$

The set of hat functions

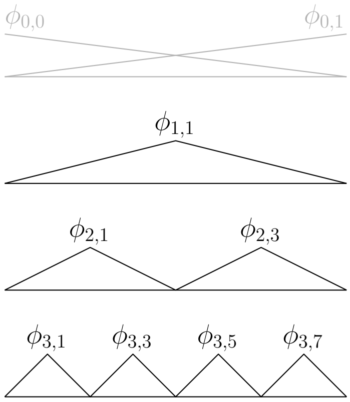
$$\left\{\phi_{\mathbf{k},\mathbf{i}} : \mathbf{1} \le \mathbf{k} \le \boldsymbol{\ell}, \mathbf{i} \in \mathcal{I}_{\mathbf{k}}^{\text{odd}}\right\}$$

is called the *hierarchical basis* of the finite-dimensional space $\mathcal{V}_{\boldsymbol{\ell}}$.

The one-dimensional nodal and hierarchical basis of level 3 are depicted in Figure 2.3.



(a) Nodal basis of $\mathcal{V}_3$.



(b) Hierarchical basis of $\mathcal{V}_3$.

Figure 2.3: One-dimensional piece-wise linear bases of level 3 with boundary.

We define the space

$$\mathcal{V} := \bigoplus_{\boldsymbol{\ell} \in \mathbb{N}^D} \mathcal{W}_{\boldsymbol{\ell}}$$

with its hierarchical basis

$$\left\{\phi_{\boldsymbol{\ell},\mathbf{i}} : \boldsymbol{\ell} \in \mathbb{N}^D, \mathbf{i} \in \mathcal{I}_{\boldsymbol{\ell}}^{\text{odd}}\right\}.$$

The completion of $\mathcal{V}$ with respect to the $\mathcal{S}_2^1$-norm is the standard Sobolev *space* $\overset{0}{\mathcal{S}}{}_2^1$.

Any function $f \in \overset{0}{\mathcal{S}}{}_2^1 \supseteq \overset{0}{\mathcal{W}}{}_2^1 \supseteq \overset{0}{\mathcal{W}}{}_2^2$ and hence any $f \in \overset{0}{\mathcal{W}}{}_p^2$ with $p \in \{2, \infty\}$ can be uniquely split

$$f(\mathbf{x}) = \sum_{\boldsymbol{\ell} \in \mathbb{N}^D} f_{\boldsymbol{\ell}}(\mathbf{x}), \qquad f_{\boldsymbol{\ell}}(\mathbf{x}) = \sum_{\mathbf{i} \in \mathcal{I}_{\boldsymbol{\ell}}^{\text{odd}}} \gamma_{\boldsymbol{\ell},\mathbf{i}} \cdot \phi_{\boldsymbol{\ell},\mathbf{i}}(\mathbf{x}) \in \mathcal{W}_{\boldsymbol{\ell}}, \tag{2.43}$$

where $\gamma_{\boldsymbol{\ell},\mathbf{i}} \in \mathbb{R}$ are the coefficients of the hierarchical basis representation of f, also called *hierarchical surplus*.

Interpolation in Finite-dimensional Spaces

We want to construct a finite-dimensional approximation space $\mathcal{U}$ of $\mathcal{V}$. The possibilities differ in the choice of the subspaces, i.e., the choice of the index set for the levels $\mathcal{L} \subset \mathbb{N}^D$. We define

$$\mathcal{U} := \bigoplus_{\boldsymbol{\ell} \in \mathcal{L}} \mathcal{W}_{\boldsymbol{\ell}} \tag{2.44}$$

with the corresponding interpolant of a function $f \in \mathcal{V}$ or $f \in \overset{0}{\mathcal{W}}{}_p^2$

$$\widetilde{f}_{\mathcal{U}}(\mathbf{x}) := \sum_{\boldsymbol{\ell} \in \mathcal{L}} f_{\boldsymbol{\ell}}(\mathbf{x}), \qquad f_{\boldsymbol{\ell}}(\mathbf{x}) := \sum_{\mathbf{i} \in \mathcal{I}_{\boldsymbol{\ell}}^{\text{odd}}} \gamma_{\boldsymbol{\ell},\mathbf{i}} \cdot \phi_{\boldsymbol{\ell},\mathbf{i}}(\mathbf{x}) \in \mathcal{W}_{\boldsymbol{\ell}}, \tag{2.45}$$

For a given integer $n \in \mathbb{N}$ we consider level index sets $\mathcal{L} \subseteq \{1, \dots, n\}^D$, which are finite.

One possible choice of $\mathcal{L}$ leads to the *full grid spaces*

$$\mathcal{V}_n^{(\infty)} := \bigoplus_{1 \leq |\boldsymbol{\ell}|_\infty \leq n} \mathcal{W}_{\boldsymbol{\ell}}. \tag{2.46}$$

Their *dimension*, i.e., the number of inner grid points of the corresponding grid, is

$$|\mathcal{V}_n^{(\infty)}| = O(h_n^{-D}) = O(2^{nD}), \tag{2.47}$$

where $h_n = 2^{-n}$ is the mesh size in every direction of the corresponding grid. The interpolation error in dependence on the mesh size is as follows.

Lemma 2.14 ([BG04, Lemma 3.5.])**.** *Let $f \in \overset{0}{\mathcal{W}}{}_p^2([0,1]^D, \mathbb{R})$ with $p \in \{2, \infty\}$. Then for the interpolant $\widetilde{f}_n^{(\infty)} \in \mathcal{V}_n^{(\infty)}$ and $q \in \{2, \infty\}$ it holds*

$$\|f - \widetilde{f}_n^{(\infty)}\|_{\mathcal{L}_q} = O(h_n^2).$$

Another choice of $\mathcal{L}$ leads to the *sparse grid spaces* (introduced by [Zen91] for $D = 2$)

$$\mathcal{V}_n^{(1)} := \bigoplus_{1 \le |\boldsymbol{\ell}|_1 - D + 1 \le n} \mathcal{W}_{\boldsymbol{\ell}} \tag{2.48}$$

$$= \bigoplus_{1 \le n(\boldsymbol{\ell}) \le n} \mathcal{W}_{\boldsymbol{\ell}} \tag{2.49}$$

where

$$n(\boldsymbol{\ell}) := |\boldsymbol{\ell}|_1 - D + 1. \tag{2.50}$$

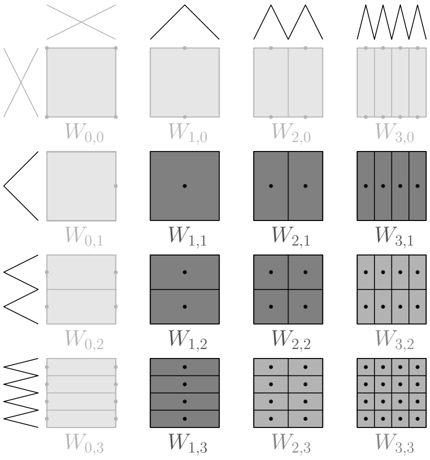


Figure 2.4: Subspaces belonging to grids without boundary for $D = 2$ and $n = 3$ (dark gray: sparse grid; dark gray&gray: full grid) and the boundary (light gray).

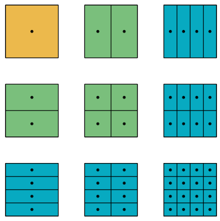


Figure 2.5: Subspaces belonging to full grids without boundary for $D = 2$ up to $n = 3$.

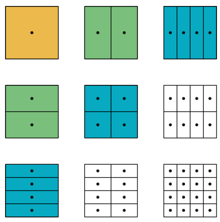


Figure 2.6: Subspaces belonging to sparse grids without boundary for $D = 2$ up to $n = 3$.

For $D = 1$ sparse and full grids are exactly the same. For $D \ge 2$, however, the number of sparse grid points is much smaller than that of a full grid with the same mesh size (see Figure 2.7 and Figure 2.8 for grids with $D = 2$).

Lemma 2.15 ([BG04, Lemma 3.6.]). *It holds*

$$|\mathcal{V}_n^{(1)}| = O(h_n^{-1} \cdot |\log_2 h_n|^{D-1}) = O(2^n \cdot n^{D-1}).$$

However, the interpolation error in dependence on the mesh size is comparably good.

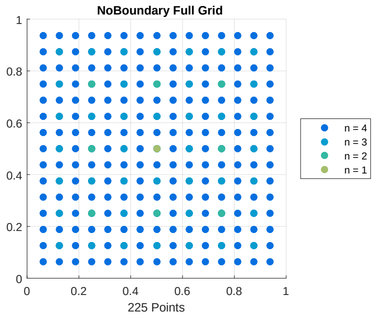


Figure 2.7: Points of full grids without boundary for $D = 2$ up to $n = 4$.

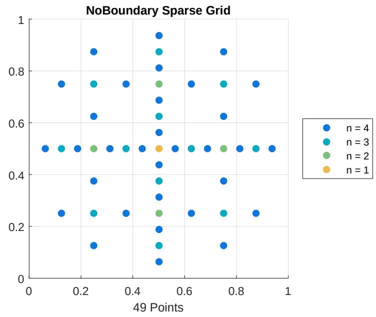


Figure 2.8: Points of sparse grids without boundary for $D = 2$ up to $n = 4$.

Lemma 2.16 ([BG04, Lemma 3.8.]). *Let $f \in \overset{0}{\mathcal{W}}{}^2_p([0,1]^D, \mathbb{R})$ with $p \in \{2, \infty\}$. Then for the interpolant $\widetilde{f}^{(1)}_n \in \mathcal{V}^{(1)}_n$ and $q \in \{2, \infty\}$ it holds*

$$\|f - \widetilde{f}^{(1)}_n\|_{\mathcal{L}_q} = O(h_n^2 \cdot n^{D-1}).$$

Later in this thesis complexity is measured by the number of interpolation points. So we turn to the interpolation errors in dependence on the number of grid points.

For interpolation on full grids we have the following.

Lemma 2.17 ([BG04]). *Let $f \in \overset{0}{\mathcal{W}}{}^2_p([0,1]^D, \mathbb{R})$ with $p \in \{2, \infty\}$. Then for the interpolant $\widetilde{f}^{(N)} \in \mathcal{V}^{(\infty)}_n$ with $N := |\mathcal{V}^{(\infty)}_n|$ and $q \in \{2, \infty\}$ it holds*

$$\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q} \leq c \cdot N^{-2/D} \cdot |f|_{\mathbf{2},q}.$$

The constant c depends on the dimension D.

For interpolation on sparse grids we have the following.

Lemma 2.18 ([BG04]). *Let $f \in \overset{0}{\mathcal{W}}{}^2_p([0,1]^D, \mathbb{R})$ with $p \in \{2, \infty\}$. Then for the interpolant $\widetilde{f}^{(N)} \in \mathcal{V}^{(1)}_n$ with $N := |\mathcal{V}^{(1)}_n|$ and $q \in \{2, \infty\}$ it holds*

$$\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q} \leq c \cdot N^{-2} \cdot (\log N)^{3 \cdot (D-1)} \cdot |f|_{\mathbf{2},q}.$$

The constant c depends on the dimension D.

We summarize the interpolation errors in dependence on the number of grid points in Table 2.1.

The sparse grid space $\mathcal{V}^{(1)}_n$ is the standard sparse grid space. It is cost–benefit (number of points vs. error) optimal in the $\mathcal{L}_\infty$-norm (and $\mathcal{L}_2$-norm).

Another sparse grid type contained in [BG04] are sparse grids which are optimal in the so-called energy norm, which is defined as the $\mathcal{L}_2$-norm of the gradient. For the Laplacian this is indeed the energy norm in finite element terminology. Those grids completely overcome the curse of dimensionality, which might be important for really high dimensions. We do not consider this type here.

Table 2.1: Interpolation error of interpolation on full and sparse grids in dependence on the number of grid points.

for $f \in \overset{0}{\mathcal{W}}{}_p^2([0,1]^D, \mathbb{R})$, $p \in \{2, \infty\}$		$\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q}$, $q \in \{2, \infty\}$
full grid	$\mathcal{V}_n^{(\infty)}$ with $\|\mathcal{V}_n^{(\infty)}\| = N$	$O(N^{-2/D})$
sparse grid	$\mathcal{V}_n^{(1)}$ with $\|\mathcal{V}_n^{(1)}\| = N$	$O(N^{-2} \cdot \|\log_2 N\|^{3 \cdot (D-1)})$

Treatment of the Boundary

So far we only considered functions which are zero on the boundary. This simplification is often applicable when the function is the solution of a partial differential equation, but we want to interpolate arbitrary functions.

We define the boundaries of a grid recursively as inner grids of lower dimension, but with the same resolution.

The full grid space with boundary

$$\underline{\mathcal{V}}_n^{(\infty)} := \bigoplus_{0 \leq |\boldsymbol{\ell}|_\infty \leq n} \mathcal{W}_{\boldsymbol{\ell}} \tag{2.51}$$

is defined as before $\mathcal{V}_n^{(\infty)}$ (see equation (2.46)), but note that here $\boldsymbol{\ell} \geq \mathbf{0}$ instead of $\boldsymbol{\ell} \geq \mathbf{1}$.

To include basis functions on the boundary, we allow $\boldsymbol{\ell} \in \mathbb{N}_0^D$ and $\mathbf{i} \in \mathbb{N}_0^D$, so we have an extra level $\ell_d = 0$ with indices $i_d \in \{0, 1\}$. Hat functions $\phi_{\boldsymbol{\ell},\mathbf{i}}$ belonging to boundary points are cut off at the boundary (as in Figure 2.3). The index sets for the hierarchical difference spaces $\underline{\mathcal{W}}_{\boldsymbol{\ell}} = \operatorname{span}\{\phi_{\boldsymbol{\ell},\mathbf{i}} : \mathbf{i} \in \underline{\mathcal{I}}_{\boldsymbol{\ell}}^{\text{odd}}\}$ become (cf. equation (2.42); see Figure 2.4)

$$\underline{\mathcal{I}}_{\boldsymbol{\ell}}^{\text{odd}} := \left\{ \mathbf{i} \in \mathbb{N}_0^D : \begin{array}{ll} 1 \leq i_d \leq 2^{\ell_d} - 1,\ i_d \text{ odd} & \text{if } \ell_d > 0 \\ 0 \leq i_d \leq 1 & \text{if } \ell_d = 0 \end{array} \text{ for } d = 1, \ldots, D \right\}. \tag{2.52}$$

The sparse grid space with boundary

$$\underline{\mathcal{V}}_n^{(1)} := \bigoplus_{0 \leq \underline{n}(\boldsymbol{\ell}) \leq n} \underline{\mathcal{W}}_{\boldsymbol{\ell}} \tag{2.53}$$

where

$$\underline{n}(\boldsymbol{\ell}) := \sum_{\substack{d = 1:\, D \\ d > 0}} (\ell_d - 1) + 1 \tag{2.54}$$

$$= |\boldsymbol{\ell}|_1 - (D - z(\boldsymbol{\ell})) + 1 \tag{2.55}$$

is defined very similar as before (see equations (2.49) and (2.50)), but with an extra quantity

$$z(\boldsymbol{\ell}) := \left| \left\{ d \in \{1, \ldots, D\} : \ell_d = 0 \right\} \right| \tag{2.56}$$

that ensures the resolution on the boundary is the same as in the interior. We set

$$\underline{n}(\mathbf{0}) := 0. \tag{2.57}$$

In the interior, i.e., $\boldsymbol{\ell} \geq \mathbf{1}$, we have $z(\boldsymbol{\ell}) = 0$ and hence $\underline{n}(\boldsymbol{\ell}) = n(\boldsymbol{\ell})$, so we recover (2.49). On the boundary we recover (2.49) on the lower dimensional boundary manifold, i.e., when we remove all dimensions $\{d : \ell_d = 0\}$ that contribute to z.

So far the interior hierarchical basis was extended by the two nodal basis functions $\phi_{0,0}(x) = 1 - x$ and $\phi_{0,1}(x) = x$ on level $\ell = 0$. To analyze the interpolation error we need another treatment of the boundary. Now we extend the interior hierarchical basis by the two hierarchical basis functions $\phi_{-1,0}(x) = 1$ and $\phi_{0,1}(x) = x$ on the levels $\ell = -1$ and $\ell = 0$ (see Figure 2.10). The level $\ell = -1$ is an artificial level and $\phi_{-1,0}$ is attached to $x_{-1,0} = x_{0,0} = 0$. On level $\ell = 0$ the basis function $\phi_{0,1}$ is attached to $x_{0,1} = 1$. We have $\boldsymbol{\ell} \in \{\mathbb{N}_0 \cup \{-1\}\}^D$.

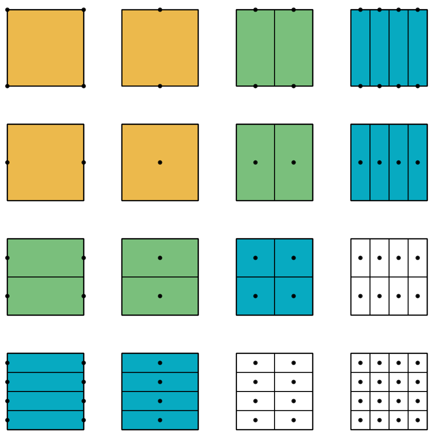


Figure 2.9: Subspaces belonging to sparse grids with nodal boundary for $D = 2$ up to $n = 3$.

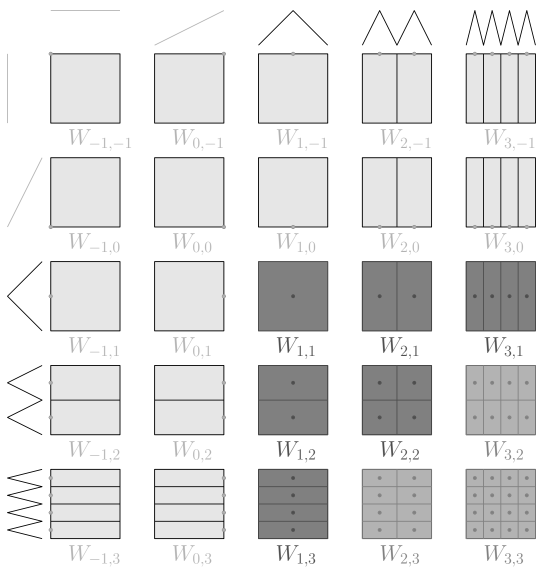


Figure 2.10: Subspaces belonging to grids with hierarchical boundary for $D = 2$ for $n = 3$ (dark gray: sparse grid; dark gray&gray: full grid) and the boundary (light gray).

The definition of the index set for the hierarchical difference spaces is straightforward (cf. equation (2.52))

$$\underline{\mathcal{I}}_{\boldsymbol{\ell}}^{\text{odd}} := \left\{ \mathbf{i} \in \mathbb{N}_0^D : \begin{array}{ll} 1 \leq i_d \leq 2^{\ell_d} - 1, \; i_d \text{ odd} & \text{if } \ell_d > 0 \\ 1 = i_d & \text{if } \ell_d = 0 \\ 0 = i_d & \text{if } \ell_d = -1 \end{array} \quad \text{for } d = 1, \ldots, D \right\}.$$

For the definition of $\underline{n}(\boldsymbol{\ell})$ (see equations (2.54)–(2.55)) we now have

$$z(\boldsymbol{\ell}) := \left|\left\{d \in \{1,\ldots,D\} : \ell_d \in \{-1,0\}\right\}\right| \tag{2.58}$$

and

$$\underline{n}(\boldsymbol{\ell}) := 0 \quad \text{for } \boldsymbol{\ell} \in \{-1,0\}^D \tag{2.59}$$

instead of equations (2.56) and (2.57).

With the above definitions the two representations of the boundary lead to the same grids and spaces.

We define

$$\underline{\mathcal{V}} := \bigoplus_{\boldsymbol{\ell} \in \{\mathbb{N}_0 \cup \{-1\}\}^D} \underline{\mathcal{W}_{\boldsymbol{\ell}}}.$$

The completion of $\underline{\mathcal{V}}$ with respect to the $\mathcal{S}_2^1$-norm is the standard SOBOLEV *space* $\mathcal{S}_2^1$.

Before we can analyze he interpolation error, we need to introduce a dimension-wise decomposition into the three cases $\ell_d = -1$, $\ell_d = 0$ and $\ell_d > 0$ (denoted $\ell_d = *$) for $d = 1,\ldots,D$. We define $\boldsymbol{\ell}^{(-1)} := \left\{d \in \{1,\ldots,D\} : \ell_d = -1\right\}$, $\boldsymbol{\ell}^{(0)} := \left\{d \in \{1,\ldots,D\} : \ell_d = 0\right\}$ and $\boldsymbol{\ell}^{(*)} := \left\{d \in \{1,\ldots,D\} : \ell_d = *\right\}$.

Lemma 2.19 ([Feu10, Lemma 2.1.13.])**.** *Any function $f \in \mathcal{S}_2^1([0,1]^D,\mathbb{R})$ can be split uniquely into*

$$f(x) = \sum_{\boldsymbol{\ell} \in \{-1,0,*\}^D} \underline{f_{\boldsymbol{\ell}}}(x), \quad \underline{f_{\boldsymbol{\ell}}} \in \underline{\mathcal{W}_{\boldsymbol{\ell}}} \tag{2.60}$$

such that each element consists of inner *and* linear *contributions*

$$\underline{f_{\boldsymbol{\ell}}}(x) = f_{\boldsymbol{\ell},*}\left(x_{\boldsymbol{\ell}^{(*)}}\right) \cdot \prod_{d \in \boldsymbol{\ell}^{(0)}} x_d \tag{2.61}$$

Here, the uniquely determined $f_{\boldsymbol{\ell},*}$ is a $|\boldsymbol{\ell}^{(*)}|$-dimensional function satisfying the homogeneous boundary conditions $x_d \in \{0,1\} \Rightarrow f_{\boldsymbol{\ell},*}\left(x_{\boldsymbol{\ell}^{(*)}}\right) = 0$ for any $d \in \boldsymbol{\ell}^{(*)}$.

Thus, $\underline{f_{\boldsymbol{\ell}}}$ does not depend on x_d for $d \in \boldsymbol{\ell}^{(-1)}$ and it is linear in directions $d \in \boldsymbol{\ell}^{(0)}$ whereas the dynamics in the remaining directions is governed by $f_{\boldsymbol{\ell},*} : [0,1]^{|\boldsymbol{\ell}^{(*)}|} \to \mathbb{R}$.

Note again that $\mathcal{S}_2^1 \supseteq \mathcal{W}_2^1 \supseteq \mathcal{W}_2^2 \supseteq \mathcal{W}_\infty^2$ on a bounded domain.

For interpolation on sparse grids with boundary we have the following.

Lemma 2.20 ([Feu10, Lemma 2.1.16])**.** *Let $f \in \mathcal{W}_p^2([0,1]^D,\mathbb{R})$ with $p \in \{2,\infty\}$ be given such that*

$$|||f|||_{\mathbf{2},q} := \max_{\substack{\boldsymbol{\ell} \in \{-1,0,*\}^D \\ |\boldsymbol{\ell}^{(*)}| > 0}} \{|f_{\boldsymbol{\ell},*}|_{\mathbf{2},q}\} < \infty \tag{2.62}$$

for $q \in \{2, \infty\}$ *and* $f = \sum_{\boldsymbol{\ell} \in \{-1,0,*\}^D} \underline{f}_{\boldsymbol{\ell}}$ *its unique boundary splitting according to* (2.60).

Then for the interpolant $\widetilde{f}^{(N)} \in \underline{\mathcal{V}}_n^{(1)}$ *with* $N := |\underline{\mathcal{V}}_n^{(1)}|$ *it holds*

$$\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q} \leq c \cdot N^{-2} \cdot (\log N)^{3\cdot(D-1)} \cdot |||f|||_{\mathbf{2},q}.$$

The constant c *depends on the dimension* D*.*

2.3.2 Global Polynomial Basis Functions

Global (Lagrange) polynomials [BNR00] instead of piece-wise linear basis functions (hat functions) are suitable to interpolate smooth functions which are continuously differentiable up to some order R. They give an interpolation error convergence rate of R (see Lemma 2.21) compared to a rate of 2 (see Lemma 2.20).

We consider multi-variate functions from $\mathcal{C}_\infty^R(\Omega, \mathcal{B})$ for the domain $\Omega = [-1, 1]^D$ and the Banach space $\mathcal{B} = \mathbb{R}$ (cf. definition (2.19)).

In contrast to piece-wise linear basis functions, the interpolation points in one dimension are not chosen equidistantly, but as the m *extrema of the* Chebyshev *polynomials* of degree $m - 1$, which are given as $x_1 = 0$ for $m = 1$ and

$$x_i = -\cos\left(\frac{\pi \cdot (i-1)}{m-1}\right), \quad i = 1, \ldots, m,$$

for any $m \geq 2$.

The choice $m_1 = 1$ and $m_\ell = 2^{\ell-1} + 1$ for $\ell \geq 2$ leads to nested sets of points.

Table 2.2: Extrema of Chebyshev polynomials.

ℓ	m_ℓ	$x_1, \ldots, x_{m_\ell}$
1	1	0
2	3	-1; 0; 1
3	5	-1; $-\frac{\sqrt{2}}{2}$; 0; $\frac{\sqrt{2}}{2}$; 1

The same tensor product construction as for piece-wise linear sparse grids then leads to higher-dimensional sparse grids.

Lemma 2.21 ([BNR00, Theorem 8.])**.** *Let* $f \in \mathcal{C}_\infty^R([-1, 1]^D, \mathbb{R})$*. Then it holds*

$$\|f - \widetilde{f}^{(N)}\|_{\mathcal{C}_\infty^0} \leq c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot \|f\|_{\mathcal{C}_\infty^R},$$

where N *is the number of grid points. The constant* c *depends on the dimension* D *and the smoothness* R*.*

2.3.3 Basis Representations

In this section we develop a representation of the interpolant of a function $f : \mathcal{P} \to \mathbb{R}$ whose coefficients are point evaluations of the function f at the interpolation grid.

For full grids this is already true for the nodal basis representation. For sparse grids we need to modify the nodal basis to achieve this goal.

We consider a fixed refinement n. To simplify notation we enumerate the grid points consecutively, such that the points with indices $\ell = 1, \dots, N_{\mathrm{S}}$ are the sparse grid points and the points with indices $\ell = 1, \dots, N_{\mathrm{F}}$ are the full grid points. It holds $N_{\mathrm{S}} \leq N_{\mathrm{F}}$. Similarly we enumerate the basis functions consecutively such that the hierarchical basis is $\{\phi^{(\ell)} \mid \ell \in \{1, \dots, N_{\mathrm{S}}, \dots, N_{\mathrm{F}}\}\}$ and the nodal basis is $\{\xi^{(\ell)} \mid \ell \in \{1, \dots, N_{\mathrm{F}}\}\}$. We put the basis functions into vectors $\boldsymbol{\xi}(\mathbf{p}) := [\xi^{(j)}(\mathbf{p})]_{j=1}^{N_{\mathrm{F}}}$ and $\boldsymbol{\phi}(\mathbf{p}) := [\phi^{(i)}(\mathbf{p})]_{i=1}^{N_{\mathrm{F}}}$. Also the values of f at the full grid points, the so-called nodal values, and the hierarchical coefficients are stored in vectors $\mathbf{f} := [f(\mathbf{p}^{(j)})]_{j=1}^{N_{\mathrm{F}}}$ and $\boldsymbol{\gamma} := [\gamma^{(i)}]_{i=1}^{N_{\mathrm{F}}}$, respectively.

Full Grid Interpolant

The full grid interpolant $\widetilde{f}^{(N_{\mathrm{F}})}$ of f fulfills the interpolation conditions

$$\widetilde{f}^{(N_{\mathrm{F}})}(\mathbf{p}^{(\ell)}) = f(\mathbf{p}^{(\ell)}) \quad \text{for } \ell = 1, \dots, N_{\mathrm{F}} \tag{2.63}$$

and lies in $\underline{\mathcal{V}}_n^{(\infty)} = \operatorname{span}\{\xi^{(j)} \mid j = 1, \dots, N_{\mathrm{F}}\} = \operatorname{span}\{\phi^{(i)} \mid i = 1, \dots, N_{\mathrm{F}}\}$.

The nodal basis representation is therefore

$$\widetilde{f}^{(N_{\mathrm{F}})}(\mathbf{p}) = \mathbf{f}^T \boldsymbol{\xi}(\mathbf{p}) = \sum_{j=1}^{N_{\mathrm{F}}} f(\mathbf{p}^{(j)}) \cdot \xi^{(j)}(\mathbf{p}) \tag{2.64}$$

since $\xi^{(j)}(\mathbf{p}^{(\ell)}) = \delta_{j\ell}$.

The hierarchical basis representation is

$$\widetilde{f}^{(N_{\mathrm{F}})}(\mathbf{p}) = \boldsymbol{\gamma}^T \boldsymbol{\phi}(\mathbf{p}) = \sum_{i=1}^{N_{\mathrm{F}}} \gamma^{(i)} \cdot \phi^{(i)}(\mathbf{p}). \tag{2.65}$$

The hierarchical coefficients $\gamma^{(i)}$ are not simply point evaluations of the function f as the nodal coefficients in the nodal basis representation (2.64). But one can convert the nodal basis and coefficients into the hierarchical basis and coefficients, respectively, (*hierarchization*) and the other way round (*de-hierarchization*) (see also [Gri98, pages 159-160]). There exists an invertible upper triangular matrix $\boldsymbol{\Theta} = [\theta_{ij}]_{i,j=1}^{N_{\mathrm{F}}}$ [Feu10] with

$$\boldsymbol{\phi}(\mathbf{p}) = \boldsymbol{\Theta} \cdot \boldsymbol{\xi}(\mathbf{p}). \tag{2.66}$$

Hence,

$$\begin{aligned}
\widetilde{f}^{(N_{\mathrm{F}})}(\mathbf{p}) &\overset{(2.64)}{=} \mathbf{f}^T \boldsymbol{\xi}(\mathbf{p}) \\
&\overset{(2.66)}{=} \mathbf{f}^T \boldsymbol{\Theta}^{-1} \boldsymbol{\phi}(\mathbf{p}) \\
&= (\boldsymbol{\Theta}^{-T} \mathbf{f})^T \boldsymbol{\phi}(\mathbf{p}),
\end{aligned}$$

i.e., the hierarchical coefficient vector $\boldsymbol{\gamma}$ (cf. (2.65)) is

$$\boldsymbol{\gamma} = \boldsymbol{\Theta}^{-T}\mathbf{f}. \tag{2.67}$$

We partition $\boldsymbol{\Theta}$ and $\boldsymbol{\Theta}^{-1}$ as follows: the upper left block is of size $N_\mathrm{S} \times N_\mathrm{S}$ and implies the sizes of the other blocks. We get

$$\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{\Theta}_{11} & \boldsymbol{\Theta}_{12} \\ \boldsymbol{\Theta}_{21} & \boldsymbol{\Theta}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Theta}_{11} & \boldsymbol{\Theta}_{12} \\ \mathbf{O} & \boldsymbol{\Theta}_{22} \end{bmatrix}$$

since $\boldsymbol{\Theta}$ is an upper triangular matrix. Similarly we get

$$\boldsymbol{\Theta}^{-1} = \begin{bmatrix} (\boldsymbol{\Theta}^{-1})_{11} & (\boldsymbol{\Theta}^{-1})_{12} \\ (\boldsymbol{\Theta}^{-1})_{21} & (\boldsymbol{\Theta}^{-1})_{22} \end{bmatrix} = \begin{bmatrix} (\boldsymbol{\Theta}^{-1})_{11} & (\boldsymbol{\Theta}^{-1})_{12} \\ \mathbf{O} & (\boldsymbol{\Theta}^{-1})_{22} \end{bmatrix}$$

since $\boldsymbol{\Theta}^{-1}$ is an upper triangular matrix as well.

We split a vector $\mathbf{v} \in \mathbb{R}^{N_\mathrm{F}}$ into $\mathbf{v}_S := \mathbf{v}(1\colon N_\mathrm{S}) \in \mathbb{R}^{N_\mathrm{S}}$ and $\mathbf{v}_{F-S} := \mathbf{v}(N_\mathrm{S} + 1\colon N_\mathrm{F}) \in \mathbb{R}^{N_\mathrm{F}-N_\mathrm{S}}$.

Then (2.67) can be written as

$$\begin{aligned} \begin{bmatrix} \boldsymbol{\gamma}_S \\ \boldsymbol{\gamma}_{F-S} \end{bmatrix} &= \begin{bmatrix} (\boldsymbol{\Theta}^{-1})_{11} & (\boldsymbol{\Theta}^{-1})_{12} \\ \mathbf{O} & (\boldsymbol{\Theta}^{-1})_{22} \end{bmatrix}^T \cdot \begin{bmatrix} \mathbf{f}_S \\ \mathbf{f}_{F-S} \end{bmatrix} \\ &= \begin{bmatrix} (\boldsymbol{\Theta}^{-1})_{11}^T & \mathbf{O} \\ (\boldsymbol{\Theta}^{-1})_{12}^T & (\boldsymbol{\Theta}^{-1})_{22}^T \end{bmatrix} \cdot \begin{bmatrix} \mathbf{f}_S \\ \mathbf{f}_{F-S} \end{bmatrix}, \end{aligned}$$

so

$$\boldsymbol{\gamma}_S = (\boldsymbol{\Theta}^{-1})_{11}^T \cdot \mathbf{f}_S. \tag{2.68}$$

Sparse Grid Interpolant

For $D = 1$ sparse grids are identical to full grids.

The sparse grid interpolant $\widetilde{f}^{(N_\mathrm{S})}$ of f is defined in hierarchical basis representation as

$$\widetilde{f}^{(N_\mathrm{S})}(\mathbf{p}) = \boldsymbol{\gamma}_S^T \cdot \boldsymbol{\phi}_S(\mathbf{p}). \tag{2.69}$$

We want to represent $\widetilde{f}^{(N_\mathrm{S})}$ in a basis $\{\widehat{\xi}^{(j)} \mid j = 1, \ldots, N_\mathrm{S}\}$ that does not depend on the function f and in which the coefficients of $\widetilde{f}^{(N_\mathrm{S})}$ are $f(\mathbf{p}^{(1)}), \ldots, f(\mathbf{p}^{(N_\mathrm{S})})$ (cf. (2.64) for full grids).

It holds

$$\begin{aligned} \widetilde{f}^{(N_\mathrm{S})}(\mathbf{p}) &\overset{(2.69)}{=} \boldsymbol{\gamma}_S^T \cdot \boldsymbol{\phi}_S(\mathbf{p}) \\ &\overset{(2.68)}{=} \left((\boldsymbol{\Theta}^{-1})_{11}^T \cdot \mathbf{f}_S\right)^T \cdot \boldsymbol{\phi}_S(\mathbf{p}) \\ &= \mathbf{f}_S^T \cdot (\boldsymbol{\Theta}^{-1})_{11} \cdot \boldsymbol{\phi}_S(\mathbf{p}) \\ &= \mathbf{f}_S^T \cdot \widehat{\boldsymbol{\xi}}(\mathbf{p}) \end{aligned} \tag{2.70}$$

with the basis we were looking for

$$\widehat{\boldsymbol{\xi}}(\mathbf{p}) := (\boldsymbol{\Theta}^{-1})_{11} \cdot \boldsymbol{\phi}_S(\mathbf{p}). \tag{2.71}$$

Alternatively, we can write

$$\begin{aligned}\widehat{\boldsymbol{\xi}}(\mathbf{p}) &\overset{(2.71)}{=} (\boldsymbol{\Theta}^{-1})_{11} \cdot \boldsymbol{\phi}_S(\mathbf{p}) \\ &= \begin{bmatrix} (\boldsymbol{\Theta}^{-1})_{11} & \mathbf{O} \end{bmatrix} \cdot \boldsymbol{\phi}(\mathbf{p}) \\ &\overset{(2.66)}{=} \begin{bmatrix} (\boldsymbol{\Theta}^{-1})_{11} & \mathbf{O} \end{bmatrix} \cdot \boldsymbol{\Theta} \cdot \boldsymbol{\xi}(\mathbf{p}). \end{aligned} \tag{2.72}$$

It holds

$$\begin{aligned}\begin{bmatrix} (\boldsymbol{\Theta}^{-1})_{11} & \mathbf{O} \end{bmatrix} \cdot \boldsymbol{\Theta} &= \begin{bmatrix} (\boldsymbol{\Theta}^{-1})_{11} & \mathbf{O} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{\Theta}_{11} & \boldsymbol{\Theta}_{12} \\ \mathbf{O} & \boldsymbol{\Theta}_{22} \end{bmatrix} \\ &= \begin{bmatrix} (\boldsymbol{\Theta}^{-1})_{11} \cdot \boldsymbol{\Theta}_{11} & (\boldsymbol{\Theta}^{-1})_{11} \cdot \boldsymbol{\Theta}_{12} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I} & (\boldsymbol{\Theta}^{-1})_{11} \cdot \boldsymbol{\Theta}_{12} \end{bmatrix} \end{aligned} \tag{2.73}$$

since $\boldsymbol{\Theta}^{-1} \cdot \boldsymbol{\Theta} = \mathbf{I}$, and $\boldsymbol{\Theta}$ and $\boldsymbol{\Theta}^{-1}$ are upper triangular matrices.

So we obtain

$$\begin{aligned}\widehat{\boldsymbol{\xi}}(\mathbf{p}) &\overset{(2.72)}{=} \begin{bmatrix} (\boldsymbol{\Theta}^{-1})_{11} & \mathbf{O} \end{bmatrix} \cdot \boldsymbol{\Theta} \cdot \boldsymbol{\xi}(\mathbf{p}) \\ &\overset{(2.73)}{=} \begin{bmatrix} \mathbf{I} & (\boldsymbol{\Theta}^{-1})_{11} \cdot \boldsymbol{\Theta}_{12} \end{bmatrix} \cdot \boldsymbol{\xi}(\mathbf{p}) \\ &= \boldsymbol{\xi}_S(\mathbf{p}) + (\boldsymbol{\Theta}^{-1})_{11} \cdot \boldsymbol{\Theta}_{12} \cdot \boldsymbol{\xi}_{F-S}(\mathbf{p}) \end{aligned} \tag{2.74}$$

and finally get

$$\begin{aligned}\widetilde{f}^{(N_\mathrm{S})}(\mathbf{p}) &\overset{(2.70)}{=} \mathbf{f}_S^T \cdot \widehat{\boldsymbol{\xi}}(\mathbf{p}) \\ &\overset{(2.74)}{=} \mathbf{f}_S^T \cdot \left(\boldsymbol{\xi}_S(\mathbf{p}) + (\boldsymbol{\Theta}^{-1})_{11} \cdot \boldsymbol{\Theta}_{12} \cdot \boldsymbol{\xi}_{F-S}(\mathbf{p})\right) \\ &= \sum_{j=1}^{N_\mathrm{S}} \mathbf{f}(\mathbf{p}^{(j)}) \cdot \widehat{\xi}^{(j)}(\mathbf{p}) \end{aligned} \tag{2.75}$$

with

$$\widehat{\xi}^{(j)}(\mathbf{p}) \;= \xi^{(j)}(\mathbf{p}) + \sum_{k=N_\mathrm{S}+1}^{N_\mathrm{F}} [(\boldsymbol{\Theta}^{-1})_{11} \cdot \boldsymbol{\Theta}_{12}]_{jk} \cdot \xi^{(k)}(\mathbf{p}) \quad \text{for } j = 1, \ldots, N_\mathrm{S}. \tag{2.76}$$

It follows that the sparse grid interpolant fulfills the interpolation conditions

$$\widetilde{f}^{(N_\mathrm{S})}(\mathbf{p}^{(\ell)}) = f(\mathbf{p}^{(\ell)}) \quad \text{for } \ell = 1, \ldots, N_\mathrm{S} \tag{2.77}$$

since for $j = 1, \ldots, N_\mathrm{S}$

$$\begin{aligned}\widehat{\xi}^{(j)}(\mathbf{p}^{(\ell)}) &= \xi^{(j)}(\mathbf{p}^{(\ell)}) + \sum_{k=N_\mathrm{S}+1}^{N_\mathrm{F}} [(\mathbf{\Theta}^{-1})_{11} \cdot \mathbf{\Theta}_{12}]_{jk} \cdot \xi^{(k)}(\mathbf{p}^{(\ell)}) \\ &= \delta_{j\ell} + \sum_{k=N_\mathrm{S}+1}^{N_\mathrm{F}} [(\mathbf{\Theta}^{-1})_{11} \cdot \mathbf{\Theta}_{12}]_{jk} \cdot \delta_{k\ell} \\ &= \delta_{j\ell}.\end{aligned}$$

We have $\|\xi^{(\ell)}\|_{\mathcal{L}_\infty} = 1$, $\ell \in \{1, \ldots, N_F\}$, so for $j = 1, \ldots, N_\mathrm{S}$

$$\|\widehat{\xi}^{(j)}\|_{\mathcal{L}_\infty} < \infty, \tag{2.78}$$

since $\widehat{\xi}^{(j)}$ is a finite linear combination of functions $\xi^{(\ell)}$.

Note that the full grid interpolant of a function which has only positive values has only positive values. This is generally not the case for the sparse grid interpolant of such a function. So even for scalar-valued functions direct interpolation on sparse grids does not preserve positivity. As we want to preserve the symmetric positive (negative) definiteness of the system matrices later in this thesis we will interpolate them on matrix manifolds.

Further note that equation (2.75) particularly implies that, if the sparse grid interpolant of a positive function is negative at some point $\mathbf{p}$ in parameter space, then there exists a basis function $\widehat{\xi}^{(j)}$ for which $\widehat{\xi}^{(j)}(\mathbf{p}) < 0$. Hence, we cannot combine interpolation on sparse grids with methods that require positive weighting functions.

2.3.4 Entry-wise Interpolation of Matrix-valued Functions

We treat each entry of a matrix as a scalar and interpolate the matrix-valued function entry-wise ($\widetilde{\mathbf{F}}^{(N)}(\mathbf{p}) = [\widetilde{f}_{ij}^{(N)}(\mathbf{p})]_{ij}$).

We perform simultaneous entry-wise interpolation, which means that we interpolate each matrix entry in the same space with the same set of basis functions independent of the indices i and j. Therefore we can write the sparse grid interpolant of the matrix-valued function $\mathbf{F}$ as

$$\widetilde{\mathbf{F}}^{(N_S)}(\mathbf{p}) = \sum_{k=1}^{N_S} \mathbf{F}(\mathbf{p}^{(k)}) \cdot \widehat{\xi}^{(k)}(\mathbf{p}) \tag{2.79}$$

(cf. (2.75) for scalar-valued functions) with $\widehat{\xi}^{(k)}(\mathbf{p})$ from the scalar case (2.76).

3 Parametric Model Order Reduction

3.1 Parametric Model

The *parametric full order model* is a multi-input, multi-output linear time-invariant dynamical system in (generalized) state space form parametrized with D parameters $\mathbf{p} = [p_1, \ldots, p_D]^T \in \mathcal{P} = [0,1]^D$

$$\begin{aligned} \mathbf{E}(\mathbf{p}) \cdot \dot{\mathbf{x}}(t;\mathbf{p}) &= \mathbf{A}(\mathbf{p}) \cdot \mathbf{x}(t;\mathbf{p}) + \mathbf{B}(\mathbf{p}) \cdot \mathbf{u}(t), \\ \mathbf{y}(t;\mathbf{p}) &= \mathbf{C}(\mathbf{p}) \cdot \mathbf{x}(t;\mathbf{p}), \end{aligned} \tag{3.1}$$

where $\mathbf{A}(\mathbf{p}), \mathbf{E}(\mathbf{p}) \in \mathbb{R}^{n\times n}$, $\mathbf{B}(\mathbf{p}) \in \mathbb{R}^{n\times m}$ and $\mathbf{C}(\mathbf{p}) \in \mathbb{R}^{q\times n}$.

When the parameter space $\mathcal{P}$ is not the unit cube but a cuboid, it can be scaled and translated to fit the unit cube.

The *transfer function of the parametric full order model* (3.1) is

$$\mathbf{H}(s;\mathbf{p}) = \mathbf{C}(\mathbf{p}) \cdot (s\mathbf{E}(\mathbf{p}) - \mathbf{A}(\mathbf{p}))^{-1} \cdot \mathbf{B}(\mathbf{p}). \tag{3.2}$$

3.1.1 Requirements

We demand the following.

Requirement 3.1. *Let* $\mathbf{A}$*,* $\mathbf{A}^{-1}$*,* $\mathbf{E}$ *and* $\mathbf{E}^{-1}$ *be in* $\mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{n\times n})$*.*

For symmetric systems we can equivalently say the spectra of $\mathbf{A}$ and $\mathbf{E}$ should be bounded over the whole parameter space and stay away from zero (Lemma 4.3), which is a reasonable requirement.

Reasonable requirements for the parametric input- and output-map are the following.

Requirement 3.2. *Let* $\mathbf{B} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{n\times m})$ *and* $\mathbf{C} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{q\times n})$*.*

With the above requirements the norms we introduce in section 3.3 are finite.

Later we will require the following to ensure the stability of the reduced order models (see subsection 4.3.1).

Requirement 3.3. *Let* $-\mathbf{A}(\mathbf{p})$ *and* $\mathbf{E}(\mathbf{p})$ *be symmetric positive definite for all* $\mathbf{p}$ *in the parameter space* $\mathcal{P}$*.*

3.2 Parametric Model Order Reduction

We seek a *parametric reduced order model*

$$\begin{aligned} \mathbf{E}_r(\mathbf{p}) \cdot \dot{\mathbf{x}}_r(t;\mathbf{p}) &= \mathbf{A}_r(\mathbf{p}) \cdot \mathbf{x}_r(t;\mathbf{p}) + \mathbf{B}_r(\mathbf{p}) \cdot \mathbf{u}(t), \\ \mathbf{y}_r(t;\mathbf{p}) &= \mathbf{C}_r(\mathbf{p}) \cdot \mathbf{x}_r(t;\mathbf{p}), \end{aligned} \tag{3.3}$$

where $\mathbf{A}_r(\mathbf{p})$, $\mathbf{E}_r(\mathbf{p}) \in \mathbb{R}^{r\times r}$, $\mathbf{B}_r(\mathbf{p}) \in \mathbb{R}^{r\times m}$ and $\mathbf{C}_r(\mathbf{p}) \in \mathbb{R}^{q\times r}$ with $r \ll n$.
The *transfer function of the parametric reduced order model* (3.3) is

$$\mathbf{H}_r(s;\mathbf{p}) = \mathbf{C}_r(\mathbf{p}) \cdot (s\mathbf{E}_r(\mathbf{p}) - \mathbf{A}_r(\mathbf{p}))^{-1} \cdot \mathbf{B}_r(\mathbf{p}). \tag{3.4}$$

The following requirements ensure that the norms we introduce in section 3.3 are finite and the error measures we introduce in section 3.3 are well-defined.

Requirement 3.4. *Let* $\mathbf{A}_r$, $\mathbf{A}_r^{-1}$, $\mathbf{E}_r$ *and* $\mathbf{E}_r^{-1}$ *be in* $\mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$.

Requirement 3.5. *Let* $\mathbf{B}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$ *and* $\mathbf{C}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$.

The method we propose in chapter 4 gives parametric reduced order models of this class (see subsection 4.3.2).

3.3 Error Measures

Firstly, we consider error measures at a fixed parameter point $\widehat{\mathbf{p}}$ and are hence in the case of error measures for non-parametric models (see subsection 2.2.3). Secondly, we consider error measures for the whole parameter space.

To measure the transfer function error (in the frequency domain) for a given parameter point $\widehat{\mathbf{p}}$ we use the $\mathcal{H}_2$*-norm*

$$\|\mathbf{H}(\cdot\,;\widehat{\mathbf{p}}) - \mathbf{H}_r(\cdot\,;\widehat{\mathbf{p}})\|_{\mathcal{H}_2} = \left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\mathbf{H}(s;\widehat{\mathbf{p}}) - \mathbf{H}_r(s;\widehat{\mathbf{p}})\|_F^2 \;\mathrm{d}s\right)^{1/2} \tag{3.5}$$

or the $\mathcal{H}_\infty$*-norm*

$$\|\mathbf{H}(\cdot\,;\widehat{\mathbf{p}}) - \mathbf{H}_r(\cdot\,;\widehat{\mathbf{p}})\|_{\mathcal{H}_\infty} = \sup_{s\in\imath\mathbb{R}} \|\mathbf{H}(s;\widehat{\mathbf{p}}) - \mathbf{H}_r(s;\widehat{\mathbf{p}})\|_2. \tag{3.6}$$

For the whole parameter space we define the norms

$$\|\mathbf{H}\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)} := \left(\frac{1}{2\pi}\int_{\mathcal{P}}\int_{\imath\mathbb{R}} \|\mathbf{H}(s;\mathbf{p})\|_F^2 \;\mathrm{d}s\;\mathrm{d}\mathbf{p}\right)^{1/2} \tag{3.7}$$

and

$$\|\mathbf{H}\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} := \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \sup_{s\in\imath\mathbb{R}} \|\mathbf{H}(s;\mathbf{p})\|_2. \tag{3.8}$$

So we can measure the transfer function error (in the frequency domain) with

$$\|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)} = \left(\frac{1}{2\pi}\int_{\mathcal{P}}\int_{\imath\mathbb{R}} \|\mathbf{H}(s;\mathbf{p}) - \mathbf{H}_r(s;\mathbf{p})\|_F^2 \;\mathrm{d}s\;\mathrm{d}\mathbf{p}\right)^{1/2} \tag{3.9}$$

or

$$\|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} = \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \sup_{s\in\imath\mathbb{R}} \|\mathbf{H}(s;\mathbf{p}) - \mathbf{H}_r(s;\mathbf{p})\|_2. \tag{3.10}$$

These two error measures are very similar to the error measures in [BGW15]. However, we write $\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)$ instead of $\mathcal{H}_2 \otimes \mathcal{L}_2$ introduced by [BBBG11] (which is equivalent since $\mathcal{L}_2$ and $\mathcal{H}_2$ are Hilbert spaces) and $\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)$ instead of $\mathcal{H}_\infty \otimes \mathcal{L}_\infty$ (which is only a notation) with ess sup instead of sup (which should make no difference in practice).

We now prove that the requirements for the parametric full order model from subsection 3.1.1 and the requirements for the parametric reduced order model from section 3.2, respectively, ensure that the transfer functions are in $\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)$ and $\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)$. So the error measures introduced in this section are well-defined.

Everything is based on the $\mathcal{L}_\infty$-norm of the matrix-valued functions—not only for $\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)$, but also for $\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)$. The $\mathcal{L}_2$-norm is not suitable, since we need to deal with terms of product structure.

Lemma 3.6. *Let $\mathbf{A}^{-1} \in \mathcal{L}_\infty(\mathcal{P},\mathbb{R}^{n\times n})$, $\mathbf{B} \in \mathcal{L}_\infty(\mathcal{P},\mathbb{R}^{n\times m})$ and $\mathbf{C} \in \mathcal{L}_\infty(\mathcal{P},\mathbb{R}^{q\times n})$. Then $\mathbf{H} \in \mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)$.*

Proof. We can bound the $\mathcal{H}_\infty$-norm of the transfer function by the spectral norm of $\mathbf{A}^{-1}$, $\mathbf{B}$ and $\mathbf{C}$ at any parameter point $\mathbf{p}$. We have

$$\begin{aligned}
\|\mathbf{H}(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty} &= \sup_{s\in\imath\mathbb{R}} \|\mathbf{H}(s;\mathbf{p})\|_2 \\
&= \sup_{s\in\imath\mathbb{R}} \|\mathbf{C}(\mathbf{p}) \cdot (s\mathbf{E}(\mathbf{p}) - \mathbf{A}(\mathbf{p}))^{-1} \cdot \mathbf{B}(\mathbf{p})\|_2 \\
&\le \sup_{s\in\imath\mathbb{R}} \left\{\|\mathbf{C}(\mathbf{p})\|_2 \cdot \|(s\mathbf{E}(\mathbf{p}) - \mathbf{A}(\mathbf{p}))^{-1}\|_2 \cdot \|\mathbf{B}(\mathbf{p})\|_2\right\} \\
&= \|\mathbf{C}(\mathbf{p})\|_2 \cdot \sup_{s\in\imath\mathbb{R}} \|(s\mathbf{E}(\mathbf{p}) - \mathbf{A}(\mathbf{p}))^{-1}\|_2 \cdot \|\mathbf{B}(\mathbf{p})\|_2 \\
&= \|\mathbf{C}(\mathbf{p})\|_2 \cdot \|\mathbf{A}(\mathbf{p})^{-1}\|_2 \cdot \|\mathbf{B}(\mathbf{p})\|_2
\end{aligned}$$

using Lemma B.4 for the last equality.

So we can bound the $\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)$-norm by the $\mathcal{L}_\infty$-norm of $\mathbf{A}^{-1}$, $\mathbf{B}$ and $\mathbf{C}$. It holds

$$\begin{aligned}
\|\mathbf{H}\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{H}(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty} \\
&\le \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \left\{\|\mathbf{C}(\mathbf{p})\|_2 \cdot \|\mathbf{A}(\mathbf{p})^{-1}\|_2 \cdot \|\mathbf{B}(\mathbf{p})\|_2\right\} \\
&\le \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{C}(\mathbf{p})\|_2 \cdot \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{A}(\mathbf{p})^{-1}\|_2 \cdot \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{B}(\mathbf{p})\|_2 \\
&= \|\mathbf{C}\|_{\mathcal{L}_\infty} \cdot \|\mathbf{A}^{-1}\|_{\mathcal{L}_\infty} \cdot \|\mathbf{B}\|_{\mathcal{L}_\infty}.
\end{aligned}$$

Due to the assumptions all factors are finite. □

Lemma 3.7 (cf. Lemma 3.6)**.** *Let* $\mathbf{A}$*,* $\mathbf{A}^{-1}$ *and* $\mathbf{E}^{-1}$ *be in* $\mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{n\times n})$*. Let* $\mathbf{B} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{n\times m})$ *and* $\mathbf{C} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{q\times n})$*. Then* $\mathbf{H} \in \mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$*.*

Proof. We can bound the $\mathcal{H}_2$-norm of the transfer function by the spectral norms of $\mathbf{A}^{-1}$, $\mathbf{A}$, $\mathbf{E}^{-1}$, $\mathbf{B}$ and $\mathbf{C}$ at any parameter point $\mathbf{p}$. We have

$$\begin{aligned}
\|\mathbf{H}(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2}^2 &= \frac{1}{2\pi}\int_{\imath\mathbb{R}} \left\|\mathbf{C}(\mathbf{p})\cdot(s\mathbf{E}(\mathbf{p})-\mathbf{A}(\mathbf{p}))^{-1}\cdot\mathbf{B}(\mathbf{p})\right\|_F^2 \,\mathrm{d}s\\
&\le \frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\mathbf{C}(\mathbf{p})\|_F^2\cdot\left\|(s\mathbf{E}(\mathbf{p})-\mathbf{A}(\mathbf{p}))^{-1}\right\|_F^2\cdot\|\mathbf{B}(\mathbf{p})\|_F^2 \,\mathrm{d}s\\
&= \frac{1}{2\pi}\cdot\|\mathbf{C}(\mathbf{p})\|_F^2\cdot\left(\int_{\imath\mathbb{R}}\left\|(s\mathbf{E}(\mathbf{p})-\mathbf{A}(\mathbf{p}))^{-1}\right\|_F^2\,\mathrm{d}s\right)\cdot\|\mathbf{B}(\mathbf{p})\|_F^2\\
&\le \frac{1}{2\pi}\cdot\|\mathbf{C}(\mathbf{p})\|_F^2\cdot\pi\cdot n\cdot\|\mathbf{A}(\mathbf{p})^{-1}\|_2^2\cdot\|\mathbf{A}(\mathbf{p})\|_2\cdot\|\mathbf{E}(\mathbf{p})^{-1}\|_2\cdot\|\mathbf{B}(\mathbf{p})\|_F^2\\
&= \frac{n}{2}\cdot\|\mathbf{C}(\mathbf{p})\|_F^2\cdot\|\mathbf{A}(\mathbf{p})^{-1}\|_2^2\cdot\|\mathbf{A}(\mathbf{p})\|_2\cdot\|\mathbf{E}(\mathbf{p})^{-1}\|_2\cdot\|\mathbf{B}(\mathbf{p})\|_F^2
\end{aligned}$$

using submultiplicativity twice for the first inequality and Lemma B.5 for the last inequality.

So we can bound the $\mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$-norm by the $\mathcal{L}_\infty$-norm of $\mathbf{A}^{-1}$, $\mathbf{A}$, $\mathbf{E}^{-1}$, $\mathbf{B}$ and $\mathbf{C}$. To go from Frobenius norm to spectral norm we use the matrix norm inequality (2.5). It holds

$$\begin{aligned}
&\|\mathbf{H}\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)}^2\\
&= \int_{\mathcal{P}} \|\mathbf{H}(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2}^2\,\mathrm{d}\mathbf{p}\\
&\le \int_{\mathcal{P}} \left\{\frac{n}{2}\cdot\|\mathbf{C}(\mathbf{p})\|_F^2\cdot\|\mathbf{A}(\mathbf{p})^{-1}\|_2^2\cdot\|\mathbf{A}(\mathbf{p})\|_2\cdot\|\mathbf{E}(\mathbf{p})^{-1}\|_2\cdot\|\mathbf{B}(\mathbf{p})\|_F^2\right\}\,\mathrm{d}\mathbf{p}\\
&\overset{(2.5)}{\le} \int_{\mathcal{P}} \left\{\frac{n}{2}\cdot n\cdot\|\mathbf{C}(\mathbf{p})\|_2^2\cdot\|\mathbf{A}(\mathbf{p})^{-1}\|_2^2\cdot\|\mathbf{A}(\mathbf{p})\|_2\cdot\|\mathbf{E}(\mathbf{p})^{-1}\|_2\cdot n\cdot\|\mathbf{B}(\mathbf{p})\|_2^2\right\}\,\mathrm{d}\mathbf{p}\\
&= \frac{n^3}{2}\cdot\int_{\mathcal{P}} \left\{\|\mathbf{C}(\mathbf{p})\|_2^2\cdot\|\mathbf{A}(\mathbf{p})^{-1}\|_2^2\cdot\|\mathbf{A}(\mathbf{p})\|_2\cdot\|\mathbf{E}(\mathbf{p})^{-1}\|_2\cdot\|\mathbf{B}(\mathbf{p})\|_2^2\right\}\,\mathrm{d}\mathbf{p}\\
&\le \frac{n^3}{2}\cdot\|\mathbf{C}\|_{\mathcal{L}_\infty}^2\cdot\|\mathbf{A}^{-1}\|_{\mathcal{L}_\infty}^2\cdot\|\mathbf{A}\|_{\mathcal{L}_\infty}\cdot\|\mathbf{E}^{-1}\|_{\mathcal{L}_\infty}\cdot\|\mathbf{B}\|_{\mathcal{L}_\infty}^2\cdot\int_{\mathcal{P}} 1\,\mathrm{d}\mathbf{p}.
\end{aligned}$$

Since $\mathcal{P}$ is bounded, it has finite measure. Due to the assumptions the other factors are finite, too. □

Lemma 3.8. *It holds* $\mathbf{H} \in \mathcal{L}_\infty(\mathcal{P}, \mathcal{H}_\infty)$ *and* $\mathbf{H} \in \mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$*.*

Proof. We apply Lemma 3.6 and Lemma 3.7 using Requirement 3.1 and Requirement 3.2. □

Lemma 3.9. *It holds* $\mathbf{H}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathcal{H}_\infty)$ *and* $\mathbf{H}_r \in \mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$*.*

Proof. We apply Lemma 3.6 and Lemma 3.7 using Requirement 3.4 and Requirement 3.5. □

3.4 Methods

In parametric model order reduction, the reduced order model is parameter dependent. The question is how to efficiently deal with this parameter dependence.

We divide a parametric model order reduction method into an offline pre-computation phase and an online phase, which is applied every time the parametric reduced order model is evaluated at an arbitrary point in parameter space.

One possibility to compute the system matrices of the parametric reduced order model is by projection

$$\begin{aligned}\mathbf{A}_r(\mathbf{p}) &= \mathbf{W}(\mathbf{p})^T \cdot \mathbf{A}(\mathbf{p}) \cdot \mathbf{V}(\mathbf{p}),\\ \mathbf{E}_r(\mathbf{p}) &= \mathbf{W}(\mathbf{p})^T \cdot \mathbf{E}(\mathbf{p}) \cdot \mathbf{V}(\mathbf{p}),\\ \mathbf{B}_r(\mathbf{p}) &= \mathbf{W}(\mathbf{p})^T \cdot \mathbf{B}(\mathbf{p}) \quad \text{and}\\ \mathbf{C}_r(\mathbf{p}) &= \mathbf{C}(\mathbf{p}) \cdot \mathbf{V}(\mathbf{p}).\end{aligned} \tag{3.11}$$

Constant (global) projection matrices $\mathbf{W}(\mathbf{p}) = \mathbf{W}$ and $\mathbf{V}(\mathbf{p}) = \mathbf{V}$ are proposed e.g., in [BBBG11, BF14]. This approach might lead to a large reduction order r, especially when the system matrices of the parametric full order model vary substantially over the parameter space. Only if the parameter dependence of the full order system matrices is or is made *affine*, i.e., a linear combination of scalar-valued functions and constant matrices, the representation (3.11) can be simplified and all computations that depend on the full order n can be done in the offline phase of the method.

In contrast to that, in [AF08] local reduced order bases and interpolation of the projection matrices $\mathbf{W}(\cdot)$ and $\mathbf{V}(\cdot)$ (on a manifold) over the parameter space are proposed. The resulting reduction order r is smaller than for global reduced order bases. However, for every evaluation of the parametric reduced order model at a specific point in parameter space the parametric full order model needs to be evaluated. That is, the online phase of the method depends on the full order n. Therefore the method should only be applied when model order reduction at an arbitrary point in parameter space is far more expensive than the online phase.

This weakness is overcome by computing local reduced order models and interpolating the reduced system matrices $\mathbf{A}_r(\cdot)$, $\mathbf{E}_r(\cdot)$, $\mathbf{B}_r(\cdot)$, and $\mathbf{C}_r(\cdot)$ (on manifolds) instead [DVW10, PMEL10, AF11, Maz14]. Note that for this approach the representation (3.11) does not hold for arbitrary $\mathbf{p}$, but only for the interpolation points in parameter space.

In [DVW10, Maz14] constant (global) reduced order bases are used to compute the local reduced order models. The method of [AF11] includes an additional step, where different bases are made consistent, so parameter-dependent (local) reduced order bases can be used. [PMEL10] also includes a way of making the reduced order models compatible, but interpolation is not done on manifolds.

Interpolation on matrix manifolds preserves special matrix properties. In [DVW10] the manifold of invertible matrices is used for the square system matrix. In [AF11] also the manifold of symmetric positive definite matrices is used, which is important for the preservation of stability (see subsection 4.3.1).

The order of the resulting reduced order model at an arbitrary point in parameter space is the same as the order r of the local reduced order models and does not depend on the number of interpolation points N in parameter space.

We adopt the method from [AF11] for our special case ($-\mathbf{A}(\mathbf{p})$, $\mathbf{E}(\mathbf{p})$ symmetric positive definite, $\mathbf{W}(\mathbf{p}) = \mathbf{V}(\mathbf{p})$ one-sided projection) and describe the details in the next subsection. In chapter 4 we connect it with interpolation on sparse grids (see section 2.3) and describe specific choices, extensions and modifications.

Also in [GBP+14] the method from [AF11] is connected with interpolation on sparse grids. However, stability of the resulting reduced order models is not guaranteed. Furthermore, the model order reduction method that is used is not optimal and not included into the method.

A stability-preserving parametric model order reduction method which does not involve matrix manifolds is presented in [ECP+11, GPWL14]. However, non-negative weighting functions are required, so interpolation on sparse grids is not allowed.

A completely different parametric model order reduction method that involves interpolation, namely the interpolation of the transfer function, is suggested in [BB09]. The main problem of this approach is that when the system poles vary over the parameter space spurious poles occur. To deal with higher-dimensional parameter spaces interpolation is done on sparse grids.

In [BBBG11] sparse grids are mentioned, too. However, not in terms of interpolation, but merely for sampling at the grid points. At each point a local reduced order basis is computed and then all are united to build a big global reduced order basis.

3.4.1 Interpolation of Local Reduced System Matrices on Manifolds

In this section we describe the method of [AF11].

First, N points in parameter space $\mathbf{p}_1, \dots, \mathbf{p}_N$ and a reduction order r are selected. Second, for $j = 1, \dots, N$ a (non-parametric) model order reduction method is applied to the parametric full order model at the fixed parameter point $\mathbf{p}_j$ to construct a local reduced order basis of dimension r spanned by $\mathbf{V}_j$. Then the local reduced order bases are transformed (Step A) such that they fit together. This is the method's offline phase.

In the online phase of the method (Step B) we interpolate the system matrices of the local reduced order models on manifolds.

Step A: Congruence Transformations

The system matrices of the local reduced order models are not unique, as there exist infinitely many *equivalent* systems (see subsection 2.2.1). For a meaningful interpolation the entries of the system matrices, that is the coordinates, have to be expressed in the same reduced order basis. Since $\mathbf{V}(\cdot)$ and also its span vary over the parameter space this cannot be achieved exactly. Still, the $\mathbf{V}(\mathbf{p}_j)$ can be adapted to a reference projection matrix $\mathbf{V}_{\text{ref}}$ as follows.

So before we apply interpolation we transform each pre-computated reduced order model into a consistent set of generalized coordinates.

Multiplying a basis matrix $\mathbf{V}$ with an orthogonal matrix $\mathbf{Q}$ changes the basis, but not the system itself—the systems are *equivalent*. We choose

$$\mathbf{Q}_j = \underset{\mathbf{Q}\in\mathbb{R}^{r\times r}\text{: orth.}}{\operatorname{argmin}} \|\mathbf{V}_j \cdot \mathbf{Q} - \mathbf{V}_{\text{ref}}\|_F^2,$$

where $\mathbf{V}_{\text{ref}}$ is the representation of the reference basis all other systems are adapted to.

The solution of this *orthogonal Procrustes problem* is given by $\mathbf{Q}_j := \mathbf{U}_j \cdot \mathbf{Z}_j^T$, where $\mathbf{U}_j \cdot \mathbf{\Sigma}_j \cdot \mathbf{Z}_j^T = \mathbf{P}_j$ is a *singular value decomposition* of

$$\mathbf{P}_j := \mathbf{V}_j^T \cdot \mathbf{V}_{\text{ref}}. \tag{3.12}$$

The basis changes lead to transformed system matrices of the local reduced order models at the interpolation points in parameter space

$$\begin{aligned} \mathbf{E}_r^{\mathbf{p}_j} &\leftarrow \mathbf{Q}_j^T \cdot \mathbf{E}_r^{\mathbf{p}_j} \cdot \mathbf{Q}_j, \\ \mathbf{A}_r^{\mathbf{p}_j} &\leftarrow \mathbf{Q}_j^T \cdot \mathbf{A}_r^{\mathbf{p}_j} \cdot \mathbf{Q}_j, \\ \mathbf{B}_r^{\mathbf{p}_j} &\leftarrow \mathbf{Q}_j^T \cdot \mathbf{B}_r^{\mathbf{p}_j}, \quad \text{and} \\ \mathbf{C}_r^{\mathbf{p}_j} &\leftarrow \mathbf{C}_r^{\mathbf{p}_j} \cdot \mathbf{Q}_j. \end{aligned} \tag{3.13}$$

Step B: Interpolation on Matrix Manifolds

Instead of interpolating the matrix entries directly, we interpolate the associated linear operators on their appropriate matrix manifold.

(a) The manifold of symmetric positive definite matrices of size r for $-\mathbf{A}_r$ and $\mathbf{E}_r$.

(b) The manifold of $r \times m$ real matrices, $\mathbb{R}^{r\times m}$, for $\mathbf{B}_r$.

(c) The manifold of $q \times r$ real matrices, $\mathbb{R}^{q\times r}$, for $\mathbf{C}_r$.

The interpolation consists of three steps. First, the matrices are mapped from the manifold onto the tangent space of one fixed matrix $\mathbf{X}$ of the manifold (logarithm mapping). Next, the interpolation process takes place in this linear tangent space. Last, the resulting matrix is mapped back onto the manifold (exponential mapping). Table 3.1 shows which mapping corresponds to which manifold. exp and log denote the matrix exponential (see Definition 2.5) and logarithm (see Theorem 2.6).

Table 3.1: Exponential and logarithm mappings for matrix manifolds of interest.

manifold	$\mathbb{R}^{M\times N}$	symmetric positive definite matrices
$\mathrm{Exp}_{\mathbf{X}}(\mathbf{\Gamma})$	$\mathbf{X}+\mathbf{\Gamma}$	$\mathbf{X}^{1/2}\cdot\exp(\mathbf{\Gamma})\cdot\mathbf{X}^{1/2}$
$\mathrm{Log}_{\mathbf{X}}(\mathbf{Y})$	$\mathbf{Y}-\mathbf{X}$	$\log(\mathbf{X}^{-1/2}\cdot\mathbf{Y}\cdot\mathbf{X}^{-1/2})$

4 A Stability-preserving Parametric Model Order Reduction Method

The method we propose in this thesis (see section 4.1) is based on the method from [AF11] (see subsection 3.4.1). We make some specific choices, extensions and modifications, which we explain in detail in section 4.2, and connect it with interpolation on sparse grids (see section 2.3).

Features of the method are its

- stability guarantee,
- suitability for higher-dimensional parameter spaces,
- flexibility for any interpolation method,
- automatism, so no tuning is necessary,
- accurate reduced order models due to local reduced order bases,
- reusability of old local reduced order models when new interpolation points are added.

The only drawback of our method is its limitation to symmetric systems.

To run the method one simply needs to choose a reduction order and an interpolation method, i.e., a grid and basis functions. We suggest sparse grids with global polynomial basis functions.

4.1 Algorithm

The method is summarized in the following two algorithms, which form the two distinct phases of the method.

The first phase is the offline or reduction phase, where we perform model order reduction at discrete points in parameter space (Algorithm 4.1). The second phase is the online or interpolation phase, where we interpolate the local reduced order models in the parameter space (Algorithm 4.2).

Figure 4.1 depicts a flow chart of the method.

Algorithm 4.1 PMOR for symmetric LTI systems - Offline phase (reduction).

Compute reduced order models at parameter grid
(`orth` denotes the orthogonalization of the column vectors.)

1. Perform Iterative Rational Krylov Algorithm (see Algorithm 2.1) at center point of parameter space to compute reference reduced order basis $\mathbf{V}_{\text{ref}} = \mathbf{V}(\mathbf{p}_{\text{center}})$, $\mu_i^{\text{ref}} = \mu_i(\mathbf{p}_{\text{center}})$ and $\mathbf{b}_i^{\text{ref}} = \mathbf{b}_i(\mathbf{p}_{\text{center}})$ for $i = 1, \ldots, r$.
2. Choose sparse grid (see section 2.3) as parameter grid.
3. For every point $\mathbf{p}_j$ of parameter grid:
   a) Compute reduced order basis
   $\check{\mathbf{V}}_j = \texttt{orth}\left[\left(\mu_i^{\text{ref}} \cdot \mathbf{E}(\mathbf{p}_j) - \mathbf{A}(\mathbf{p}_j)\right)^{-1} \mathbf{B}(\mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}}\right]_{i=1}^{r}$ with $\operatorname{rank}(\check{\mathbf{V}}_j) = r$.
   b) Adjust reduced order basis:
      i. Compute $\mathbf{P} = \check{\mathbf{V}}_j^T \cdot \mathbf{V}_{\text{ref}} \in \mathbb{R}^{r \times r}$.
      ii. Compute singular value decomposition $\mathbf{P} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{Z}^T$ of $\mathbf{P}$.
      iii. Compute $\mathbf{Q} = \mathbf{U} \cdot \mathbf{Z}^T \in \mathbb{R}^{r \times r}$ (orthogonal).
      iv. Set $\mathbf{V}_j = \check{\mathbf{V}}_j \cdot \mathbf{Q}$.
   c) Project onto reduced order basis
   $\check{\mathbf{A}}_r^{\mathbf{p}_j} = \mathbf{V}_j^T \cdot \mathbf{A}(\mathbf{p}_j) \cdot \mathbf{V}_j, \quad \check{\mathbf{B}}_r^{\mathbf{p}_j} = \mathbf{V}_j^T \cdot \mathbf{B}(\mathbf{p}_j),$
   $\check{\mathbf{E}}_r^{\mathbf{p}_j} = \mathbf{V}_j^T \cdot \mathbf{E}(\mathbf{p}_j) \cdot \mathbf{V}_j, \quad \check{\mathbf{C}}_r^{\mathbf{p}_j} = \mathbf{C}(\mathbf{p}_j) \cdot \mathbf{V}_j.$
   d) Transform system to eliminate $\check{\mathbf{E}}_r^{\mathbf{p}_j}$ ($\rightsquigarrow \mathbf{I}_r$):
   Compute Cholesky *decomposition* $\check{\mathbf{E}}_r^{\mathbf{p}_j} = \mathbf{R}^T \cdot \mathbf{R}$ of $\check{\mathbf{E}}_r^{\mathbf{p}_j}$ (triangular).
   Set $\mathbf{A}_r^{\mathbf{p}_j} = \mathbf{R}^{-T} \cdot \check{\mathbf{A}}_r^{\mathbf{p}_j} \cdot \mathbf{R}^{-1}$, $\mathbf{B}_r^{\mathbf{p}_j} = \mathbf{R}^{-T} \cdot \check{\mathbf{B}}_r^{\mathbf{p}_j}$ and $\mathbf{C}_r^{\mathbf{p}_j} = \check{\mathbf{C}}_r^{\mathbf{p}_j} \cdot \mathbf{R}^{-1}$.

Algorithm 4.2 PMOR for symmetric LTI systems - Online phase (interpolation).

Interpolate system matrices of reduced order models on their appropriate manifold
(exp and log denote the matrix exponential and matrix logarithm, respectively.)

1. For every point $\mathbf{p}_j$ of parameter grid:
   Map $-\mathbf{A}_r^{\mathbf{p}_j}$ to tangent space of matrix manifold:
   $\mathfrak{A}_r^{\mathbf{p}_j} = \log(-\mathbf{A}_r^{\mathbf{p}_j})$
2. Interpolate in linear tangent space:
   Interpolate matrices $\mathfrak{A}_r^{\mathbf{p}_j}$, $\mathbf{B}_r^{\mathbf{p}_j}$ and $\mathbf{C}_r^{\mathbf{p}_j}$ entry-wise (see subsection 2.3.4) by interpolation on sparse grid (see section 2.3).
   a) Set up interpolants $\widetilde{\mathfrak{A}}_r$, $\widetilde{\mathbf{B}}_r$ and $\widetilde{\mathbf{C}}_r$ (see (2.79)).
   b) Evaluate interpolants at $\widehat{\mathbf{p}}$ to obtain $\widetilde{\mathfrak{A}}_r(\widehat{\mathbf{p}})$, $\widetilde{\mathbf{B}}_r(\widehat{\mathbf{p}})$ and $\widetilde{\mathbf{C}}_r(\widehat{\mathbf{p}})$.
3. Map $\widetilde{\mathfrak{A}}_r(\widehat{\mathbf{p}})$ back to matrix manifold:
   $\widetilde{\mathbf{A}}_r^{\widehat{\mathbf{p}}} = -\exp\left(\widetilde{\mathfrak{A}}_r(\widehat{\mathbf{p}})\right)$

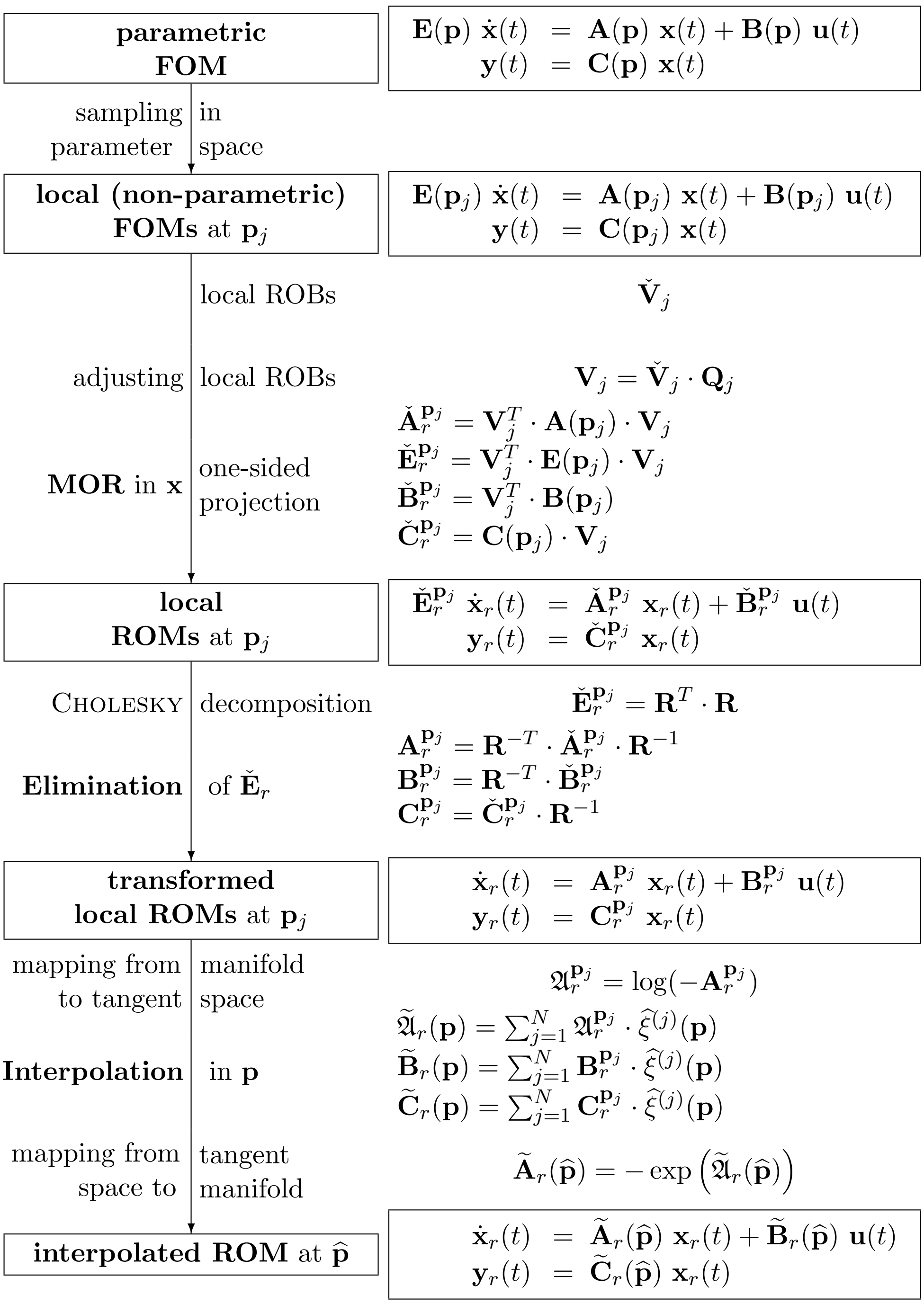


Figure 4.1: Flow chart of the method.

4.1.1 Performance Improvements

To efficiently implement the method, we would interlace some steps of the algorithm. Since this makes the succession of steps less clear, we just mention the modifications for implementation here.

From Algorithm 4.2 we would shift mapping to the tangent space (step 1) and setting up the interpolant (step 2a), i.e., the coefficients for the interpolation basis functions, to Algorithm 4.1 (offline phase). This just leaves the evaluation of the interpolant at point $\widehat{\mathbf{p}}$ (step 2b) , i.e., evaluation of the interpolation basis functions, and mapping back to the manifold (step 3) in Algorithm 4.2 (online phase).

A completely different idea to save some memory is to exploit the symmetry of the system matrices. This nearly halves the number of matrix entries that have to be interpolated.

4.2 Design Decisions

The parametric model order reduction method from [AF11] is a good method to base on since it possesses some important properties. The other methods applicable to linear time-invariant systems (see section 3.4) lack at least one of these properties.

1. The online phase is completely independent of the full order n. This is essential for any efficient method.
   In [AF08] basis matrices of size n are interpolated and multiplied with the system matrices of the full order model.

2. The reduction order r is independent of the number of interpolation points in parameter space N.

3. Some special matrix properties are preserved, e.g., symmetric positive definiteness, which is essential to preserve stability.
   In [PMEL10] the system matrices are interpolated directly and not on manifolds. In [GBP$^+$14] the system matrices are interpolated on the manifolds of regular and real matrices, but not symmetric positive definite.

4. There is no restriction on the interpolation method, e.g., concerning coefficients, basis functions or interpolation points.
   In [ECP$^+$11, GPWL14] only non-negative weighting functions (basis functions for which the coefficients are the values at the grid points) are allowed. When we interpolate on sparse grids, the weighting functions can become negative (see subsection 2.3.3).

5. The reduced order basis to compute the reduced order models need not be constant, but can vary over parameter space. So even reduced order models with a small reduction order can have good accuracy.
   In [BBBG11, BF14, DVW10, Maz14] constant (global) reduced order bases are used.

We do not use the method from [BB09] due to its problems with poles. The second and the fourth property are the reason why the PMOR method from [AF11] can efficiently be applied to higher-dimensional parameter spaces when connected with sparse grids. We make some choices, extensions and modifications to the original method and deal with them in the following.

Our method is tailored to symmetric systems (see subsection 4.2.1).

For the offline phase:

- We choose the reference reduced order basis $\mathbf{V}_{\mathrm{ref}}$ from $\mathbf{p}_{\mathrm{center}}$ (see subsection 4.2.2).
- We suggest a model order reduction method (see subsection 4.2.3) which produces guaranteed stable reduced order models by a projection matrix $\mathbf{V}$ with smooth parameter dependence and is nearly optimally.
- We switch the projection and the transformation step (see subsection 4.2.4).
- We transform the reduced order models to eliminate $\mathbf{E}_r$ (see subsection 4.2.5), which nearly halves interpolation effort in the online phase.

For the online phase:

- We use the tangent space at $\mathbf{X} = \mathbf{I}_r$ (not $\mathbf{A}_r(\mathbf{p}_{\mathrm{ref}})$) (see subsection 4.2.6) and at $\mathbf{X} = \mathbf{O}$ for $\mathbf{B}_r$ and $\mathbf{C}_r$.

4.2.1 System Structure

The method of [ECP+11, GPWL14] guarantees stability without any restriction on the system structure. However, only non-negative weighting functions can be used, so it is not compatible with interpolation on sparse grids.

Since we want to interpolate on sparse grids to handle higher-dimensional parameter spaces, we require a certain system structure, namely symmetry. Concretely, we require that $-\mathbf{A}(\mathbf{p})$ and $\mathbf{E}(\mathbf{p})$ are symmetric positive definite (Requirement 3.3).

4.2.2 Reference Basis

In [AF11] an index $\ell_0 \in \{1, \ldots, N\}$ is chosen and the reference reduced order basis is set to $\mathbf{V}_{\mathrm{ref}} = \mathbf{V}_{\ell_0}$. We chose ℓ_0 such that $\mathbf{p}_{\ell_0} = \mathbf{p}_{\mathrm{center}}$.

4.2.3 Model Order Reduction Method

Stability Preservation

Even if all local reduced order models are stable an interpolated reduced order model at an arbitrary parameter point can in general be unstable when no precautions are taken. Stabilizing an interpolated reduced order model is not preferable since it might destroy important properties such as a smooth parameter dependence and involves extra computational effort. Hence, we include stability preservation in our method.

To guarantee that the local reduced order models are stable, we assume that the matrices $-\mathbf{A}(\mathbf{p})$ and $\mathbf{E}(\mathbf{p})$ are symmetric positive definite for all $\mathbf{p} \in \mathcal{P}$ (Requirement 3.3). We use one-sided projection (see subsection 2.2.2), i.e., $\mathbf{W}(\mathbf{p}) = \mathbf{V}(\mathbf{p})$, which preserves the symmetric positive definiteness. Furthermore, we interpolate the system matrices $-\mathbf{A}_r(\mathbf{p})$ (and $\mathbf{E}_r(\mathbf{p})$) on appropriate matrix manifolds as in [AF11] (see subsection 3.4.1), which preserves their symmetric positive definiteness. Thus the interpolated reduced order model is stable (see Corollary 4.2 in subsection 4.3.1 for a detailed proof).

Preserve Smooth Parameter Dependence

We want to preserve possibly smooth parameter dependence during the reduction phase to have an efficient interpolation phase, especially when using global polynomials.

The congruence transformations of the reduced order bases (see section 3.4.1 or Algorithm 4.1 step 3b) make the reduced order models more consistent by changing the basis representations $\mathbf{V}_j$ of the local subspaces. However, we need to ensure that the local subspaces fit well together in the first place.

The methods from [AF11, PMEL10] allow for varying reduced order bases, but they do not suggest a specific method to generate them. In one of the examples the authors use a rational Krylov method called moment matching that bases on the interpolation of the transfer function in frequency space (see section 2.2.4). However, the interpolation points are not chosen optimally. We suggest a strategy for automatically choosing good, nearly optimal frequency interpolation points.

A small reduction error can be achieved by running the $\mathcal{H}_2$-optimal Iterative Rational Krylov Algorithm (Algorithm 2.1) at every interpolation point in parameter space $\mathbf{p}_j$. However, the computed $\mathcal{H}_2$-optimal interpolation points $\mu_i(\mathbf{p})$ in frequency space generally do not have a smooth parameter dependence, even if there is a smooth parameter dependence in the parametric full order model and even when using the same initial values for the μ_i.

We assume that the $\mathcal{H}_2$-optimal frequency interpolation points μ_i at some reference point $\mathbf{p}_{\text{ref}} \in \mathcal{P}$ are good, nearly $\mathcal{H}_2$-optimal frequency interpolation points for the whole parameter space $\mathcal{P}$. Therefore we propose to do non-optimal rational Krylov interpolation for all $\mathbf{p}_j$ with the $\mathcal{H}_2$-optimal frequency interpolation points and tangent directions of $\mathbf{p}_{\text{ref}}$ and choose $\mathbf{p}_{\text{ref}} = \mathbf{p}_{\text{center}}$. Note that the rational Krylov method corresponds to one IRKA-loop, i.e., the Iterative Rational Krylov Algorithm without iteration, and hence needs no extra implementation effort.

Since the proposed model order reduction method is interpolatory we have transfer function interpolation properties at every point $\mathbf{p}_j$ (see Theorem 4.13 in subsection 4.3.3 for details).

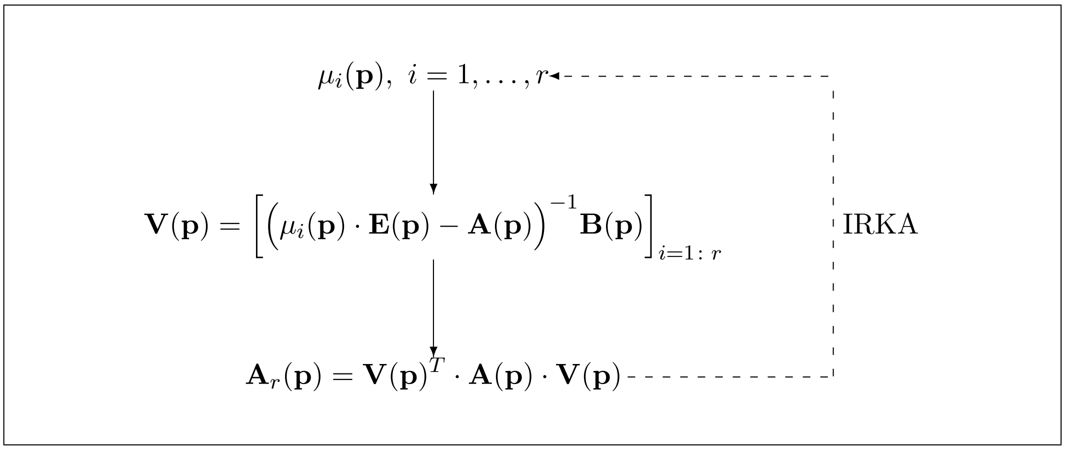


Figure 4.2: Parameter dependence of reduced order basis.

Unfortunately we could not prove the preservation of smooth parameter dependence so far. However, intuition suggests that our method preserves the smooth parameter dependence of the full order model. The orthogonalization might destroy the smoothness, but the PROCRUSTES step should repair this. The computational experiments in section 7.2 support this conjecture.

Let us note that in [BG15] a similar method for model order reduction at different points in parameter space is proposed. In contrast to our method, the transfer function is not only interpolated at some points in frequency space, but also approximated at some other points in frequency space. The frequency interpolation points are chosen as the $\mathcal{H}_2$-optimal frequency interpolation points of one particular parameter point and the frequency approximation points are chosen logarithmically spaced over the frequency range of interest. This method is computationally more expensive, but gives reduced order models of better accuracy than pure interpolation. However, stability of the reduced order models is not guaranteed.

4.2.4 Switch Projection and Reduced Order Basis Transformation

Instead of projecting

$$
\begin{aligned}
\mathbf{E}_r^{\mathbf{p}_j} &= \mathbf{V}_j^T \cdot \mathbf{E}(\mathbf{p}_j) \cdot \mathbf{V}_j, \\
\mathbf{A}_r^{\mathbf{p}_j} &= \mathbf{V}_j^T \cdot \mathbf{A}(\mathbf{p}_j) \cdot \mathbf{V}_j, \\
\mathbf{B}_r^{\mathbf{p}_j} &= \mathbf{V}_j^T \cdot \mathbf{B}(\mathbf{p}_j), \\
\mathbf{C}_r^{\mathbf{p}_j} &= \mathbf{C}(\mathbf{p}_j) \cdot \mathbf{V}_j,
\end{aligned}
$$

(cf. (3.11) with $\mathbf{W} = \mathbf{V}$ at $\mathbf{p}_j$) and then transforming the system matrices

$$\begin{aligned}
\mathbf{E}_r^{\mathbf{p}_j} &\leftarrow \mathbf{Q}_j^T \cdot \mathbf{E}_r^{\mathbf{p}_j} \cdot \mathbf{Q}_j, \\
\mathbf{A}_r^{\mathbf{p}_j} &\leftarrow \mathbf{Q}_j^T \cdot \mathbf{A}_r^{\mathbf{p}_j} \cdot \mathbf{Q}_j, \\
\mathbf{B}_r^{\mathbf{p}_j} &\leftarrow \mathbf{Q}_j^T \cdot \mathbf{B}_r^{\mathbf{p}_j}, \\
\mathbf{C}_r^{\mathbf{p}_j} &\leftarrow \mathbf{C}_r^{\mathbf{p}_j} \cdot \mathbf{Q}_j,
\end{aligned}$$

(cf. (3.13)) we include $\mathbf{Q}_j$ in $\mathbf{V}_j$

$$\mathbf{V}_j \leftarrow \mathbf{V}_j \cdot \mathbf{Q}_j$$

and then project

$$\begin{aligned}
\mathbf{E}_r^{\mathbf{p}_j} &= \mathbf{V}_j^T \cdot \mathbf{E}(\mathbf{p}_j) \cdot \mathbf{V}_j, \\
\mathbf{B}_r^{\mathbf{p}_j} &= \mathbf{V}_j^T \cdot \mathbf{B}(\mathbf{p}_j), \\
\mathbf{A}_r^{\mathbf{p}_j} &= \mathbf{V}_j^T \cdot \mathbf{A}(\mathbf{p}_j) \cdot \mathbf{V}_j, \\
\mathbf{C}_r^{\mathbf{p}_j} &= \mathbf{C}(\mathbf{p}_j) \cdot \mathbf{V}_j
\end{aligned}$$

(see steps 3b and 3c in Algorithm 4.1). This way the actual reduced order bases become visible.

4.2.5 Additional Reduced Order Basis Transformation to Simplify System

Additionally to step 3b in Algorithm 4.1 (corresponding to Step A in section 3.4.1) we perform another transformation of the reduced order basis (step 3d) before we interpolate (Algorithm 4.2, corresponding to Step B in section 3.4.1). Our goal is to achieve $\mathbf{E}_r = \mathbf{I}_r$ on the one hand and to preserve the symmetric negative definiteness of $\mathbf{A}_r$ on the other hand.

The simple transformation $\mathbf{A}_r \rightsquigarrow \mathbf{E}_r^{-1} \cdot \mathbf{A}_r$ and $\mathbf{E}_r \rightsquigarrow \mathbf{I}_r$ does preserve the symmetry only if $\mathbf{A}_r$ and $\mathbf{E}_r$ commute, which is generally not the case.

We propose to use the Cholesky factors of $\mathbf{E}_r = \mathbf{R}^T\mathbf{R}$ (Theorem A.15). Then we can transform $\mathbf{A}_r \rightsquigarrow \mathbf{R}^{-T} \cdot \mathbf{A}_r \cdot \mathbf{R}^{-1}$ and $\mathbf{E}_r \rightsquigarrow \mathbf{I}_r$. Of course, $\mathbf{B}_r$ and $\mathbf{C}_r$ need to be transformed as well: $\mathbf{B}_r \rightsquigarrow \mathbf{R}^{-T} \cdot \mathbf{B}_r$ and $\mathbf{C}_r \rightsquigarrow \mathbf{C}_r \cdot \mathbf{R}^{-1}$, so we achieve an *equivalent* system.

We could also use the matrix square root instead of the Cholesky factors to eliminate the matrix $\mathbf{E}_r$. However, this is computationally more expensive since the computation of the matrix square root for symmetric positive definite matrices involves the Cholesky decomposition anyway [Hig08, Algorithm 6.21].

Eliminating the full order model system matrix $\mathbf{E}$ before performing model order reduction is not advisable as it might destroy the sparsity of the matrix $\mathbf{A}$ and can lead to problems in model order reduction. However, for the interpolation of the reduced order models we experienced no significant changes in the error in our computational experiments (see subsection 7.2.2). If it would be a problem for other benchmarks, this step can simply be skipped.

The elimination of $\mathbf{E}_r$ seems to have no negative effects, but is has several advantages. Since the number of matrix entries is nearly halved, we need less memory to store the local reduced order models at the interpolation points in parameter space and we need less time to interpolate the local reduced order models. Additionally, the analysis of the interpolation error is much simpler.

4.2.6 Reference Point for Tangent Space

As the reference point for the manifold's tangent space we choose $\mathbf{X} = \mathbf{O}$ for arbitrary rectangular matrices and $\mathbf{X} = \mathbf{I}_r$ for symmetric positive definite matrices instead of $\mathbf{X} = -\mathbf{A}_r^{\mathbf{p}_{j_0}}$ for some index $j_0 \in \{1, \ldots, N\}$ as in [AF11]. That way, the mappings between a manifold and its tangent space are simplified to the identity and the common matrix exponential and matrix logarithm, respectively.

The theoretical background for this choice is given in [AFPA07]. Obviously, the mapping is cheaper to compute and we do not need to speculate or test which $\mathbf{p}_{j_0}$ would be best. Practical experience (cf. [Maz14]) give evidence that this simplification yields similar results and suffers from less numerical stability problems. Our experiments (see chapter 7) also indicate that this is an appropriate choice for the reference point.

4.3 Properties

4.3.1 Stability Preservation

A system with matrices $\mathbf{A}$ and $\mathbf{E}$ is stable if $\mathrm{Re}(\lambda) < 0$ for all $\lambda \in \Lambda(\mathbf{E}^{-1}\mathbf{A})$ (Definition 2.10). Due to Lemma B.7 it is sufficient that $\mathbf{A}$ is symmetric negative definite and $\mathbf{E}$ is symmetric positive definite.

Theorem 4.1. *The system matrix of the interpolated reduced order model $\widetilde{\mathbf{A}}_r(\widehat{\mathbf{p}})$ computed with Algorithm 4.1 and Algorithm 4.2 is symmetric negative definite for every $\widehat{\mathbf{p}} \in \mathcal{P}$.*

Proof. In the following we show that all square system matrices which appear in our method are symmetric negative definite and symmetric positive definite, respectively.

We demand $\mathbf{A}(\mathbf{p})$ to be symmetric negative definite and $\mathbf{E}(\mathbf{p})$ to be symmetric positive definite at every point $\mathbf{p}$ in the parameter space $\mathcal{P}$ (Requirement 3.3). Hence, the parametric full order model is stable in the whole parameter space.

We obtain the system matrices of the reduced order model by projection. Due to Theorem A.16 and Corollary A.17 the matrices $\check{\mathbf{A}}_r(\mathbf{p})$ and $\check{\mathbf{E}}_r(\mathbf{p})$ are symmetric negative definite are symmetric positive definite at every point $\mathbf{p} \in \mathcal{P}$, respectively. Hence, the parametric reduced order model is stable in the whole parameter space $\mathcal{P}$.

After the elimination of $\check{\mathbf{E}}_r(\mathbf{p})$, the matrices $\mathbf{A}_r(\mathbf{p})$ are symmetric negative definite at every point $\mathbf{p} \in \mathcal{P}$ due to Lemma B.6.

We obtain the system matrix $\widetilde{\mathbf{A}}_r(\widehat{\mathbf{p}})$ of the interpolated reduced order model by interpolation of $-\widetilde{\mathbf{A}}_r(\mathbf{p})$ on the matrix manifold of symmetric positive definite matrices. So the matrix $\widetilde{\mathbf{A}}_r(\widehat{\mathbf{p}})$ is symmetric negative definite for every point $\widehat{\mathbf{p}}$ in the parameter space $\mathcal{P}$. □

Corollary 4.2. *The interpolated reduced order model ($\widetilde{\mathbf{A}}_r(\widehat{\mathbf{p}})$, $\widetilde{\mathbf{B}}_r(\widehat{\mathbf{p}})$, $\widetilde{\mathbf{C}}_r(\widehat{\mathbf{p}})$) computed with Algorithm 4.1 and Algorithm 4.2 is stable for every $\widehat{\mathbf{p}} \in \mathcal{P}$.*

4.3.2 Preservation of Model Requirements for Well-defined Error Measures

In this section we prove that the requirements for the parametric full order model from subsection 3.1.1 (Requirement 3.1–Requirement 3.3) ensure that the interpolated reduced order model from our method fulfills Requirement 3.4 and Requirement 3.5. Recall that this implies that the transfer function of the parametric reduced order model is in $\mathcal{L}_\infty(\mathcal{P}, \mathcal{H}_\infty)$ and $\mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$ (see Lemma 3.6 and Lemma 3.7).

Table 4.1: Overview of preservation of requirements for well-defined error measures.

model/ matrices	requirement/ lemma	
full order model	Requirement 3.1	Requirement 3.2
$\mathbf{A}, \mathbf{E}, \mathbf{B}, \mathbf{C}$	$\mathbf{A}, \mathbf{A}^{-1}, \mathbf{E}, \mathbf{E}^{-1}$	$\mathbf{B} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{n\times m})$
	$\in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{n\times n})$	$\mathbf{C} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{q\times n})$
reduced order model	Lemma 4.4	Lemma 4.5
$\check{\mathbf{A}}_r, \check{\mathbf{E}}_r, \check{\mathbf{B}}_r, \check{\mathbf{C}}_r$	$\check{\mathbf{A}}_r, \check{\mathbf{A}}_r^{-1}, \check{\mathbf{E}}_r, \check{\mathbf{E}}_r^{-1}$	$\check{\mathbf{B}}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$
	$\in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$	$\check{\mathbf{C}}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$
transformed reduced order model	Lemma 4.6	Lemma 4.7
$\mathbf{A}_r, \mathbf{B}_r, \mathbf{C}_r$	$\mathbf{A}_r, \mathbf{A}_r^{-1}$	$\mathbf{B}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$
	$\in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$	$\mathbf{C}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$
in tangent space	Lemma 4.8	
$\mathfrak{A}_r$	$\mathfrak{A}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$	
interpolated in tangent space	Lemma 4.10	
$\widetilde{\mathfrak{A}}_r$	$\widetilde{\mathfrak{A}}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$	
interpolated reduced order model	Lemma 4.11	Lemma 4.12
$\widetilde{\mathbf{A}}_r, \widetilde{\mathbf{B}}_r, \widetilde{\mathbf{C}}_r$	$\widetilde{\mathbf{A}}, \widetilde{\mathbf{A}}^{-1}$	$\widetilde{\mathbf{B}}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$
	$\in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$	$\widetilde{\mathbf{C}}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$
	Requirement 3.4	Requirement 3.5

According to subsection 4.3.1, all models that appear in our method are stable. The transfer function of a stable system is in $\mathcal{H}_\infty$ [Ant05, Theorem 5.18] and in $\mathcal{H}_2$ [ABG10]. Thus, everything is fine for a fixed parameter point and we just have to deal with the norms over the parameter space and prove that they are finite.

Full Order Model

For the parametric full order model we state requirements in section 3.1. We have an equivalent characterization of Requirement 3.1 when $-\mathbf{A}$ and $\mathbf{E}$ are symmetric positive definite (Requirement 3.3).

Lemma 4.3. *When $-\mathbf{A}(\mathbf{p})$ and $\mathbf{E}(\mathbf{p})$ are symmetric positive definite at every point $\mathbf{p}$ in the parameter space $\mathcal{P}$, then $\mathbf{A}, \mathbf{A}^{-1}, \mathbf{E}, \mathbf{E}^{-1} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{n\times n})$ is equivalent to*

$$\begin{aligned}
-\infty &< \operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}} \lambda_{\min}(\mathbf{A}(\mathbf{p})), & \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \lambda_{\max}(\mathbf{A}(\mathbf{p})) &< 0,\\
0 &< \operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}} \lambda_{\min}(\mathbf{E}(\mathbf{p})), & \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \lambda_{\max}(\mathbf{E}(\mathbf{p})) &< \infty.
\end{aligned}$$

Proof.

$$\begin{aligned}
\|\mathbf{A}\|_{\mathcal{L}_\infty} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{A}(\mathbf{p})\|_2\\
&= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \left(-\lambda_{\min}(\mathbf{A}(\mathbf{p}))\right)\\
&= -\operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}} \lambda_{\min}(\mathbf{A}(\mathbf{p}))
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{A}^{-1}\|_{\mathcal{L}_\infty} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{A}(\mathbf{p})^{-1}\|_2\\
&= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \left(-\lambda_{\max}(\mathbf{A}(\mathbf{p}))^{-1}\right)\\
&= -\operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}} \left(\lambda_{\max}(\mathbf{A}(\mathbf{p}))^{-1}\right)\\
&= -\left(\operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \lambda_{\max}(\mathbf{A}(\mathbf{p}))\right)^{-1}
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{E}\|_{\mathcal{L}_\infty} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{E}(\mathbf{p})\|_2\\
&= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \lambda_{\max}(\mathbf{E}(\mathbf{p}))
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{E}^{-1}\|_{\mathcal{L}_\infty} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{E}(\mathbf{p})^{-1}\|_2\\
&= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \left(\lambda_{\min}(\mathbf{E}(\mathbf{p}))^{-1}\right)\\
&= \left(\operatorname*{ess\,inf}_{\mathbf{p}\in\mathcal{P}} \lambda_{\min}(\mathbf{E}(\mathbf{p}))\right)^{-1}
\end{aligned}$$

□

Reduced Order Model

Since we compute the reduced system matrices $\check{\mathbf{A}}_r$ and $\check{\mathbf{E}}_r$ by one-sided projection, we can show the following.

Lemma 4.4. *It holds* $\check{\mathbf{A}}_r, \check{\mathbf{A}}_r^{-1}, \check{\mathbf{E}}_r, \check{\mathbf{E}}_r^{-1} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$.

Proof. We have $\check{\mathbf{E}}_r(\mathbf{p}) = \mathbf{V}(\mathbf{p})^T \cdot \mathbf{E}(\mathbf{p}) \cdot \mathbf{V}(\mathbf{p})$ and $\check{\mathbf{A}}_r(\mathbf{p}) = \mathbf{V}(\mathbf{p})^T \cdot \mathbf{A}(\mathbf{p}) \cdot \mathbf{V}(\mathbf{p})$. Applying Corollary A.19 we get

$$\begin{aligned} \lambda_{\min}(\mathbf{A}(\mathbf{p})) &\le \lambda_{\min}(\check{\mathbf{A}}_r(\mathbf{p})), & \lambda_{\max}(\check{\mathbf{A}}_r(\mathbf{p})) &\le \lambda_{\max}(\mathbf{A}(\mathbf{p})), \\ \lambda_{\min}(\mathbf{E}(\mathbf{p})) &\le \lambda_{\min}(\check{\mathbf{E}}_r(\mathbf{p})), & \lambda_{\max}(\check{\mathbf{E}}_r(\mathbf{p})) &\le \lambda_{\max}(\mathbf{E}(\mathbf{p})). \end{aligned}$$

Hence, it holds

$$\begin{aligned} \|\check{\mathbf{A}}_r(\mathbf{p})\|_2 = -\lambda_{\min}(\check{\mathbf{A}}_r(\mathbf{p})) &\le -\lambda_{\min}(\mathbf{A}(\mathbf{p})) = \|\mathbf{A}(\mathbf{p})\|_2, \\ \|\check{\mathbf{A}}_r(\mathbf{p})^{-1}\|_2 = -\lambda_{\max}(\check{\mathbf{A}}_r(\mathbf{p}))^{-1} &\le -\lambda_{\max}(\mathbf{A}(\mathbf{p}))^{-1} = \|\mathbf{A}(\mathbf{p})^{-1}\|_2, \\ \|\check{\mathbf{E}}_r(\mathbf{p})^{-1}\|_2 = \lambda_{\min}(\check{\mathbf{E}}_r(\mathbf{p}))^{-1} &\le \lambda_{\min}(\mathbf{E}(\mathbf{p}))^{-1} = \|\mathbf{E}(\mathbf{p})^{-1}\|_2, \\ \|\check{\mathbf{E}}_r(\mathbf{p})\|_2 = \lambda_{\max}(\check{\mathbf{E}}_r(\mathbf{p})) &\le \lambda_{\max}(\mathbf{E}(\mathbf{p})) = \|\mathbf{E}(\mathbf{p})\|_2. \end{aligned}$$

Now we apply the essential supremum over the parameter space to bound the $\mathcal{L}_\infty$-norms, which are finite due to Requirement 3.1 and the equivalent characterization Lemma 4.3 together with Requirement 3.3. □

For the parametric input- and output-map of the reduced order model we can show the following.

Lemma 4.5. *It holds* $\check{\mathbf{B}}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$ *and* $\check{\mathbf{C}}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$.

Proof. We have

$$\begin{aligned} \|\check{\mathbf{B}}_r(\mathbf{p})\|_2 &= \|\mathbf{V}(\mathbf{p})^T\mathbf{B}(\mathbf{p})\|_2 \\ &\le \|\mathbf{V}(\mathbf{p})^T\|_2 \cdot \|\mathbf{B}(\mathbf{p})\|_2 \\ &= \|\mathbf{B}(\mathbf{p})\|_2 \end{aligned}$$

since $\mathbf{V}(\mathbf{p})$ has orthonormal columns. The inequality $\|\check{\mathbf{C}}_r(\mathbf{p})\|_2 \le \|\mathbf{C}(\mathbf{p})\|_2$ follows analogously. We apply the essential supremum over the parameter space to bound the $\mathcal{L}_\infty$-norms, which are finite due to Requirement 3.2. □

Transformed Reduced Order Model

Lemma 4.6. *It holds* $\mathbf{A}_r, \mathbf{A}_r^{-1} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$.

Proof. By definition of $\mathbf{R}$ it holds $\check{\mathbf{E}}_r(\mathbf{p}) = \mathbf{R}^T\mathbf{R}$. Hence, it holds $\check{\mathbf{E}}_r(\mathbf{p})^{-1} = \mathbf{R}^{-1}\mathbf{R}^{-T}$. Note that $\mathbf{R}^{-1}$ is an upper triangular matrix with positive diagonal entries, too.

We have

$$\begin{aligned}\|\mathbf{A}_r(\mathbf{p})\|_2 &= \|\mathbf{R}^{-T} \cdot \check{\mathbf{A}}_r(\mathbf{p}) \cdot \mathbf{R}^{-1}\|_2 \\ &\le \|\check{\mathbf{A}}_r(\mathbf{p})\|_2 \cdot \|\mathbf{R}^{-1}\|_2^2 \\ &= \|\check{\mathbf{A}}_r(\mathbf{p})\|_2 \cdot \|\check{\mathbf{E}}_r(\mathbf{p})^{-1}\|_2\end{aligned}$$

and

$$\begin{aligned}\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2 &= \|\mathbf{R}^{T} \cdot \check{\mathbf{A}}_r(\mathbf{p})^{-1} \cdot \mathbf{R}\|_2 \\ &\le \|\check{\mathbf{A}}_r(\mathbf{p})^{-1}\|_2 \cdot \|\mathbf{R}\|_2^2 \\ &= \|\check{\mathbf{A}}_r(\mathbf{p})^{-1}\|_2 \cdot \|\check{\mathbf{E}}_r(\mathbf{p})\|_2,\end{aligned}$$

due to Lemma B.8 for $\mathbf{M} = \check{\mathbf{E}}_r(\mathbf{p})^{-1}$ and $\mathbf{M} = \check{\mathbf{E}}_r(\mathbf{p})$, respectively.

We apply the essential supremum over the parameter space to bound the $\mathcal{L}_\infty$-norms, which are finite due to Lemma 4.4. □

Lemma 4.7. *It holds* $\mathbf{B}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$ *and* $\mathbf{C}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$.

Proof. By definition of $\mathbf{R}$ it holds $\check{\mathbf{E}}_r(\mathbf{p}) = \mathbf{R}^T\mathbf{R}$. Hence, it holds $\check{\mathbf{E}}_r(\mathbf{p})^{-1} = \mathbf{R}^{-1}\mathbf{R}^{-T}$. Note that $\mathbf{R}^{-1}$ is an upper triangular matrix with positive diagonal entries, too.

We have

$$\begin{aligned}\|\mathbf{B}_r(\mathbf{p})\|_2 &= \|\mathbf{R}^{-T} \cdot \check{\mathbf{B}}_r(\mathbf{p})\|_2 \\ &\le \|\mathbf{R}^{-T}\|_2 \cdot \|\check{\mathbf{B}}_r(\mathbf{p})\|_2 \\ &= \|\check{\mathbf{E}}_r(\mathbf{p})^{-1}\|_2^{1/2} \cdot \|\check{\mathbf{B}}_r(\mathbf{p})\|_2\end{aligned}$$

and

$$\begin{aligned}\|\mathbf{C}_r(\mathbf{p})\|_2 &= \|\check{\mathbf{C}}_r(\mathbf{p}) \cdot \mathbf{R}^{-1}\|_2 \\ &\le \|\check{\mathbf{C}}_r(\mathbf{p})\|_2 \cdot \|\mathbf{R}^{-1}\|_2 \\ &= \|\check{\mathbf{C}}_r(\mathbf{p})\|_2 \cdot \|\check{\mathbf{E}}_r(\mathbf{p})^{-1}\|_2^{1/2}\end{aligned}$$

due to Lemma B.8 for $\mathbf{M} = \check{\mathbf{E}}_r(\mathbf{p})^{-1}$.

We apply the essential supremum over the parameter space to bound the $\mathcal{L}_\infty$-norms, which are finite due to Lemma 4.4 and Lemma 4.5. □

Lemma 4.8. *It holds* $\mathfrak{A}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$.

Proof. We have $-\|\mathbf{A}_r(\mathbf{p})\|_2 = \lambda_{\min}(\mathbf{A}_r(\mathbf{p})) \le \lambda_{\max}(\mathbf{A}_r(\mathbf{p})) = -\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2^{-1}$, so

$$\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2^{-1} \le \|\mathbf{A}_r(\mathbf{p})\|_2. \tag{4.1}$$

$$\|\mathbf{A}_r(\mathbf{p})\|_2^{-1} \le \|\mathbf{A}_r(\mathbf{p})^{-1}\|_2, \tag{4.2}$$

To rewrite the norm of the matrix logarithm as the logarithm of the matrix norm we use Lemma B.9. So it holds

$$
\begin{aligned}
&\|\mathfrak{A}_r(\mathbf{p})\|_2 \\
&= \|\log(-\mathbf{A}_r(\mathbf{p}))\|_2 \\
&= \max\left\{|\log(\|-\mathbf{A}_r(\mathbf{p})\|_2)|, |\log(\|-\mathbf{A}_r(\mathbf{p})^{-1}\|_2)|\right\} \\
&= \max\left\{|\log(\|\mathbf{A}_r(\mathbf{p})\|_2)|, |\log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2)|\right\} \\
&= \max\left\{\log(\|\mathbf{A}_r(\mathbf{p})\|_2), -\log(\|\mathbf{A}_r(\mathbf{p})\|_2), \log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2), -\log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2)\right\} \\
&= \max\left\{\log(\|\mathbf{A}_r(\mathbf{p})\|_2), \log(\|\mathbf{A}_r(\mathbf{p})\|_2^{-1}), \log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2), \log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2^{-1})\right\} \\
&\overset{(4.1)}{\leq} \max\left\{\log(\|\mathbf{A}_r(\mathbf{p})\|_2), \log(\|\mathbf{A}_r(\mathbf{p})\|_2^{-1}), \log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2), \log(\|\mathbf{A}_r(\mathbf{p})\|_2)\right\} \\
&\overset{(4.2)}{\leq} \max\left\{\log(\|\mathbf{A}_r(\mathbf{p})\|_2), \log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2), \log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2), \log(\|\mathbf{A}_r(\mathbf{p})\|_2)\right\} \\
&= \max\left\{\log(\|\mathbf{A}_r(\mathbf{p})\|_2), \log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2)\right\}.
\end{aligned}
$$

We apply the essential supremum over the parameter space to bound the $\mathcal{L}_\infty$-norm. The norm

$$
\begin{aligned}
\|\mathfrak{A}_r\|_{\mathcal{L}_\infty} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathfrak{A}_r(\mathbf{p})\|_2 \\
&\leq \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \max\left\{\log(\|\mathbf{A}_r(\mathbf{p})\|_2), \log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2)\right\} \\
&= \max\left\{\operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \log(\|\mathbf{A}_r(\mathbf{p})\|_2), \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \log(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2)\right\} \\
&= \max\left\{\log(\operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{A}_r(\mathbf{p})\|_2), \log(\operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{A}_r(\mathbf{p})^{-1}\|_2)\right\} \\
&= \max\left\{\log(\|\mathbf{A}_r\|_{\mathcal{L}_\infty}), \log(\|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty})\right\}
\end{aligned}
$$

is finite due to Lemma 4.6. □

Interpolated Reduced Order Model

Lemma 4.9. *Let* $\mathbf{F} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{n\times m})$. *Then* $\widetilde{\mathbf{F}}^{(N)} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{n\times m})$ *for* $N \in \{N_\mathrm{S}, N_\mathrm{F}\}$.

Proof. We have

$$\begin{aligned}
\|\widetilde{\mathbf{F}}^{(N)}\|_{\mathcal{L}_\infty} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\widetilde{\mathbf{F}}^{(N)}(\mathbf{p})\|_2 \\
&\overset{(2.79)}{=} \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \left\| \sum_{j=1}^{N} \mathbf{F}(\mathbf{p}^{(j)}) \cdot \widehat{\xi}^{(j)}(\mathbf{p}) \right\|_2 \\
&\le \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \sum_{j=1}^{N} \|\mathbf{F}(\mathbf{p}^{(j)})\|_2 \cdot \|\widehat{\xi}^{(j)}(\mathbf{p})\|_2 \\
&\le \sum_{j=1}^{N} \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{F}(\mathbf{p})\|_2 \cdot \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\widehat{\xi}^{(j)}(\mathbf{p})\|_2 \\
&= \sum_{j=1}^{N} \|\mathbf{F}\|_{\mathcal{L}_\infty} \cdot \|\widehat{\xi}^{(j)}\|_{\mathcal{L}_\infty} \\
&\le N \cdot \|\mathbf{F}\|_{\mathcal{L}_\infty} \cdot \max_{j=1:\,N} \|\widehat{\xi}^{(j)}\|_{\mathcal{L}_\infty} \\
&< \infty
\end{aligned}$$

since $\max_{j=1:\,N} \|\widehat{\xi}^{(j)}\|_{\mathcal{L}_\infty} < \infty$ (see inequality (2.78) in subsection 2.3.3). □

Lemma 4.10. *It holds* $\widetilde{\mathfrak{A}}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$.

Proof. Apply Lemma 4.8 and Lemma 4.9 for $\mathbf{F} = \mathfrak{A}_r$. □

The next two lemmas show that our method produces reduced order models which fulfill Requirement 3.4 and Requirement 3.5.

Lemma 4.11. *It holds* $\widetilde{\mathbf{A}}_r, \widetilde{\mathbf{A}}_r^{-1} \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$.

Proof. We have

$$\begin{aligned}
\|\widetilde{\mathbf{A}}_r(\mathbf{p})^{\pm 1}\|_2 &= \|-\exp(\pm\widetilde{\mathfrak{A}}_r(\mathbf{p}))\|_2 \\
&= \|\exp(\pm\widetilde{\mathfrak{A}}_r(\mathbf{p}))\|_2 \\
&\le \exp(\|\pm\widetilde{\mathfrak{A}}_r(\mathbf{p})\|_2) \\
&= \exp(\|\widetilde{\mathfrak{A}}_r(\mathbf{p})\|_2)
\end{aligned}$$

applying Lemma A.20 for the inequality. Now we get

$$\begin{aligned}
\|\widetilde{\mathbf{A}}_r^{\pm 1}\|_{\mathcal{L}_\infty} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\widetilde{\mathbf{A}}_r^{\pm 1}(\mathbf{p})\|_2 \\
&\le \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \exp(\|\widetilde{\mathfrak{A}}_r(\mathbf{p})\|_2) \\
&= \exp(\operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\widetilde{\mathfrak{A}}_r(\mathbf{p})\|_2) \\
&= \exp(\|\widetilde{\mathfrak{A}}_r\|_{\mathcal{L}_\infty}),
\end{aligned}$$

which is finite due to Lemma 4.10. □

Lemma 4.12. *It holds* $\widetilde{\mathbf{B}}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$ *and* $\widetilde{\mathbf{C}}_r \in \mathcal{L}_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$.

Proof. Apply Lemma 4.5 and Lemma 4.9 with $\mathbf{F} = \mathbf{B}_r$ and $\mathbf{F} = \mathbf{C}_r$, respectively. □

4.3.3 Transfer Function Interpolation Property

Theorem 4.13. *The transfer functions of the full order model and the interpolated reduced order model,* $\mathbf{H}$ *and* $\widetilde{\mathbf{H}}_r$*, respectively, fulfill the one-sided tangential interpolation condition*

$$\mathbf{H}(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}} = \widetilde{\mathbf{H}}_r(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}}$$

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

Proof. We compute the local reduced order models with

$$\check{\mathbf{V}}_j = \texttt{orth}\left[\left(\mu_i^{\text{ref}} \cdot \mathbf{E}(\mathbf{p}_j) - \mathbf{A}(\mathbf{p}_j)\right)^{-1} \mathbf{B}(\mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}}\right]_{i=1:\,r}$$

so due to part 1 of Theorem 2.12 it holds

$$\mathbf{H}(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}} = \mathbf{H}_r(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}}$$

for $i = 1, \ldots, r$ and $j = 1, \ldots, N$. We interpolate the reduced system matrices in parameter space, so at the parameter interpolation points it holds

$$\mathbf{H}_r(s; \mathbf{p}_j) = \widetilde{\mathbf{H}}_r(s; \mathbf{p}_j)$$

for $s \in \overline{\mathbb{C}^+}$ and $j = 1, \ldots, N$. □

Since we use one-sided projection those properties are only right-tangential interpolation properties for the transfer function and not for its derivatives.

In the special case of $\mathbf{C}(\mathbf{p})^T = \mathbf{B}(\mathbf{p})$ we have bi-tangential interpolation properties for the transfer function and its first derivative with respect to the frequency.

Theorem 4.14. *If* $\mathbf{C}(\mathbf{p})^T = \mathbf{B}(\mathbf{p})$ *for all* $\mathbf{p} \in \mathcal{P}$*, then the transfer functions of the full order model and the interpolated reduced order model,* $\mathbf{H}$ *and* $\widetilde{\mathbf{H}}_r$*, respectively, fulfill the two-sided tangential* Hermite *interpolation conditions*

$$\begin{aligned} \mathbf{H}(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}} &= \widetilde{\mathbf{H}}_r(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}}, \\ \mathbf{c}_i^{\text{ref}T}\ \mathbf{H}(\mu_i^{\text{ref}}; \mathbf{p}_j) &= \mathbf{c}_i^{\text{ref}T}\ \widetilde{\mathbf{H}}_r(\mu_i^{\text{ref}}; \mathbf{p}_j), \qquad \textit{and} \\ \mathbf{c}_i^{\text{ref}T} \frac{\mathrm{d}}{\mathrm{d}s}\mathbf{H}(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}} &= \mathbf{c}_i^{\text{ref}T} \frac{\mathrm{d}}{\mathrm{d}s}\widetilde{\mathbf{H}}_r(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}} \end{aligned}$$

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

Proof. We compute the local reduced order models with

$$\begin{aligned}\check{\mathbf{V}}_j &= \texttt{orth}\left[\left(\mu_i^{\text{ref}} \cdot \mathbf{E}(\mathbf{p}_j) - \mathbf{A}(\mathbf{p}_j)\right)^{-1} \mathbf{B}(\mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}}\right]_{i=1:\,r} \\ &= \texttt{orth}\left[\left(\mu_i^{\text{ref}} \cdot \mathbf{E}(\mathbf{p}_j) - \mathbf{A}(\mathbf{p}_j)\right)^{-T} \mathbf{C}(\mathbf{p}_j)^T\ \mathbf{c}_i^{\text{ref}}\right]_{i=1:\,r}\end{aligned}$$

so due to Theorem 2.12 it holds

$$\begin{aligned}\mathbf{H}(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}} &= \mathbf{H}_r(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}}, \\ \mathbf{c}_i^{\text{ref}T}\ \mathbf{H}(\mu_i^{\text{ref}}; \mathbf{p}_j) &= \mathbf{c}_i^{\text{ref}T}\ \mathbf{H}_r(\mu_i^{\text{ref}}; \mathbf{p}_j), \qquad \text{and} \\ \mathbf{c}_i^{\text{ref}T} \frac{\mathrm{d}}{\mathrm{d}s}\mathbf{H}(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}} &= \mathbf{c}_i^{\text{ref}T} \frac{\mathrm{d}}{\mathrm{d}s}\mathbf{H}_r(\mu_i^{\text{ref}}; \mathbf{p}_j)\ \mathbf{b}_i^{\text{ref}}\end{aligned}$$

for $i = 1, \dots, r$ and $j = 1, \dots, N$. We interpolate the reduced system matrices in parameter space, so at the parameter interpolation points it holds

$$\begin{aligned}\mathbf{H}_r(s; \mathbf{p}_j) &= \widetilde{\mathbf{H}}_r(s; \mathbf{p}_j), \qquad \text{and} \\ \frac{\mathrm{d}}{\mathrm{d}s}\mathbf{H}_r(s; \mathbf{p}_j) &= \frac{\mathrm{d}}{\mathrm{d}s}\widetilde{\mathbf{H}}_r(s; \mathbf{p}_j)\end{aligned}$$

for $s \in \overline{\mathbb{C}^+}$ and $j = 1, \dots, N$. In fact, all derivatives of $\mathbf{H}_r$ and $\widetilde{\mathbf{H}}_r$ are equal. □

5 Interpolation Error Bounds

In this chapter we investigate the error of our proposed parametric model order reduction method. We focus on the analysis of the error which originates from interpolation of the reduced order models to relate it with the model order reduction error. First, we split the error into a model order reduction and an interpolation part. Then we derive bounds for the interpolation error.

Error Splitting

We consider the transfer functions of the original full order model, the reduced order model and the interpolated reduced order model, denoted by $\mathbf{H}$, $\mathbf{H}_r$ and $\widetilde{\mathbf{H}}_r$, respectively. We split the transfer function error into a model order reduction and an interpolation part using triangle inequalities (see Lemma B.10) as follows:

$$\|\mathbf{H} - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} \leq \underbrace{\|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)}}_{\text{model order reduction}} + \underbrace{\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)}}_{\text{interpolation}},$$

$$\|\mathbf{H} - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)} \leq \underbrace{\|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)}}_{\text{model order reduction}} + \underbrace{\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)}}_{\text{interpolation}}.$$

Interpolation Error Bounds

The interpolation algorithm which is used in our parametric model order reduction method acts on the matrices in tangent space. However, we want to quantify the interpolation error in terms of the transfer function, which is defined via the system matrices. Therefore, we derive upper bounds for the interpolation error in the transfer function which depend on the error in the tangent space matrices we interpolate.

We consider two transfer function norms: the $\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)$-norm (which contains the matrix spectral norm; see section 5.3) and the $\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)$-norm (which contains the matrix Frobenius norm; see section 5.4). Before we deal with the whole parameter space we bound the $\mathcal{H}_\infty$-norm and the $\mathcal{H}_2$-norm at a fixed parameter point (see section 5.1 and section 5.2).

The system matrices $\mathbf{B}_r(\mathbf{p})$ and $\mathbf{C}_r(\mathbf{p})$ are interpolated directly. Instead of $\mathbf{A}_r(\mathbf{p})$, we interpolate $\mathfrak{A}_r(\mathbf{p}) = \log\left(-\mathbf{A}_r(\mathbf{p})\right)$, where "log" denotes the matrix logarithm (see Figure 5.1). Hence, the interpolation error bound for the transfer function should depend on the matrix interpolation errors $\|\mathbf{B}_r(\mathbf{p}) - \widetilde{\mathbf{B}}_r(\mathbf{p})\|$, $\|\mathbf{C}_r(\mathbf{p}) - \widetilde{\mathbf{C}}_r(\mathbf{p})\|$ and $\|\mathfrak{A}_r(\mathbf{p}) - \widetilde{\mathfrak{A}}_r(\mathbf{p})\|$.

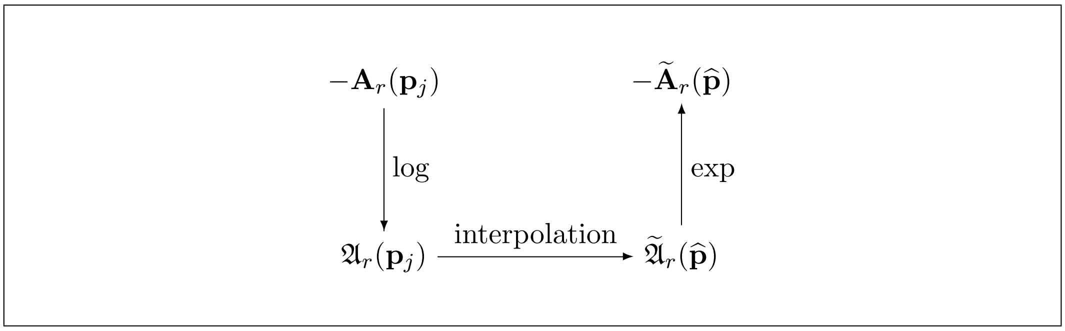


Figure 5.1: Interpolation scheme for system matrix $\mathbf{A}_r(\mathbf{p})$.

Before we start, we need some notation. In the following *we omit the index r* most of the time for the sake of simplicity. All considered quantities belong to the reduced system or the interpolated reduced system, marked with "$\sim$". The symbol Δ in front of a quantity x denotes the difference between the reduced and the interpolated reduced quantity $\Delta x = x - \widetilde{x}$. We also leave out the dependence on the parameter $\mathbf{p}$, as we consider everything at an arbitrary, but fixed parameter point $\mathbf{p}$.

The definitions of the $\mathcal{H}_\infty$-norm and the $\mathcal{H}_2$-norm involve the frequency variable s of the transfer function only on the imaginary axis, so $s \in \imath\mathbb{R}$ throughout.

The system matrix $\mathbf{A}$ is always required to be symmetric negative definite. Hence, the magnitude of its largest eigenvalue $|\lambda_{\max}(\mathbf{A})|$ equals $|\lambda|_{\min}(\mathbf{A})$, the magnitude of the eigenvalue with the smallest magnitude. To avoid confusion, we stick to the first notation. Consequently, $\|\mathbf{A}^{-1}\|_2 = |\lambda_{\max}(\mathbf{A})|^{-1}$.

With the definition $\mathfrak{A} = \log(-\mathbf{A})$ it is easy to verify that $\mathbf{A} = -\exp(\mathfrak{A})$ and $\mathbf{A}^{-1} = -\exp(-\mathfrak{A})$, where "exp" denotes the matrix exponential.

The interpolated system matrix $\widetilde{\mathbf{A}}$ is always assumed to be symmetric negative definite, too, as this is guaranteed by our algorithm (see Theorem 4.1), and we define $\widetilde{\mathfrak{A}} = \log(-\widetilde{\mathbf{A}})$.

Technical lemmas are deferred to the end of this chapter (see section 5.5) in order to focus on the core ideas in the following.

The first step of deriving our bound is to separate the three parts stemming from $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, respectively.

Lemma 5.1. *Let* $\mathbf{M} = \mathbf{M}(s) = (s\mathbf{I} - \mathbf{A})^{-1}$ *and* $\widetilde{\mathbf{M}} = \widetilde{\mathbf{M}}(s) = (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}$. *Then*

$$\|\mathbf{H}(s) - \widetilde{\mathbf{H}}(s)\| \leq \|\mathbf{C} \cdot \Delta\mathbf{M}(s) \cdot \widetilde{\mathbf{B}}\| + \|\mathbf{C} \cdot \mathbf{M}(s)\| \cdot \|\Delta\mathbf{B}\| + \|\widetilde{\mathbf{M}}(s) \cdot \widetilde{\mathbf{B}}\| \cdot \|\Delta\mathbf{C}\| \tag{5.1}$$

for any submultiplicative matrix norm.

Proof. We introduce some helpful zeros, use the triangle inequality of the matrix norm, submultiplicativity and yield

$$\|\mathbf{H} - \widetilde{\mathbf{H}}\| = \|\mathbf{CMB} - \widetilde{\mathbf{C}}\widetilde{\mathbf{M}}\widetilde{\mathbf{B}}\|$$

$$\begin{aligned}&\leq \|\mathbf{CMB}-\mathbf{CM}\widetilde{\mathbf{B}}\|+\|\mathbf{CM}\widetilde{\mathbf{B}}-\mathbf{C}\widetilde{\mathbf{M}}\widetilde{\mathbf{B}}\|+\|\mathbf{C}\widetilde{\mathbf{M}}\widetilde{\mathbf{B}}-\widetilde{\mathbf{C}}\widetilde{\mathbf{M}}\widetilde{\mathbf{B}}\|\\ &= \|\mathbf{C}\cdot\mathbf{M}\cdot(\mathbf{B}-\widetilde{\mathbf{B}})\|+\|\mathbf{C}\cdot(\mathbf{M}-\widetilde{\mathbf{M}})\cdot\widetilde{\mathbf{B}}\|+\|(\mathbf{C}-\widetilde{\mathbf{C}})\cdot\widetilde{\mathbf{M}}\cdot\widetilde{\mathbf{B}}\|\\ &\leq \|\mathbf{C}\cdot(\mathbf{M}-\widetilde{\mathbf{M}})\cdot\widetilde{\mathbf{B}}\|+\|\mathbf{C}\cdot\mathbf{M}\|\cdot\|\mathbf{B}-\widetilde{\mathbf{B}}\|+\|\widetilde{\mathbf{M}}\cdot\widetilde{\mathbf{B}}\|\cdot\|\mathbf{C}-\widetilde{\mathbf{C}}\|. \quad\square\end{aligned}$$

In the following we treat the $\mathcal{H}_\infty$-norm and the $\mathcal{H}_2$-norm case separately.

5.1 Bound for $\mathcal{H}_\infty$-norm at Fixed Point

In this section we derive an upper bound for the interpolation error in the $\mathcal{H}_\infty$-norm at a fixed point $\mathbf{p}$ in parameter space. Recall that

$$\|\mathbf{H}_r(\cdot\,;\mathbf{p})-\widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty}=\sup_{s\in\imath\mathbb{R}}\|\mathbf{H}_r(s;\mathbf{p})-\widetilde{\mathbf{H}}_r(s;\mathbf{p})\|_2.$$

In section 5.3 we will use this bound to derive an upper bound for the interpolation error in the $\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)$-norm. Recall that

$$\|\mathbf{H}_r-\widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)}=\operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}}\|\mathbf{H}_r(\cdot\,;\mathbf{p})-\widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty}.$$

5.1.1 Final Theorem

We prove an upper bound for the interpolation error of the transfer function in the $\mathcal{H}_\infty$-norm in dependence on the error of the interpolated matrices in the spectral norm.

Theorem 5.2. *Let $\mathbf{A}_r(\mathbf{p})$ and $\widetilde{\mathbf{A}}_r(\mathbf{p})$ be symmetric negative definite matrices. Define $\mathfrak{A}(\mathbf{p})=\log(-\mathbf{A}_r(\mathbf{p}))$ and $\widetilde{\mathfrak{A}}(\mathbf{p})=\log(-\widetilde{\mathbf{A}}_r(\mathbf{p}))$. Then*

$$\begin{aligned}&\|\mathbf{H}_r(\cdot\,;\mathbf{p})-\widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty}\\ &\leq\widetilde{\alpha}(\mathbf{p})\cdot(e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2}-1)+\widetilde{\beta}(\mathbf{p})\cdot\|\Delta\mathbf{B}_r(\mathbf{p})\|_2+\widetilde{\gamma}(\mathbf{p})\cdot\|\Delta\mathbf{C}_r(\mathbf{p})\|_2\end{aligned}$$

for

$$\begin{aligned}\widetilde{\alpha}(\mathbf{p})&:=\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2\cdot\|\widetilde{\mathbf{B}}_r(\mathbf{p})\|_2\cdot\|\mathbf{C}_r(\mathbf{p})\|_2\\ \widetilde{\beta}(\mathbf{p})&:=\|\mathbf{C}_r(\mathbf{p})\cdot\mathbf{A}_r(\mathbf{p})^{-1}\|_2\leq\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2\cdot\|\mathbf{C}_r(\mathbf{p})\|_2,\\ \widetilde{\gamma}(\mathbf{p})&:=\|\widetilde{\mathbf{A}}_r(\mathbf{p})^{-1}\cdot\widetilde{\mathbf{B}}_r(\mathbf{p})\|_2\leq\|\widetilde{\mathbf{A}}_r(\mathbf{p})^{-1}\|_2\cdot\|\widetilde{\mathbf{B}}_r(\mathbf{p})\|_2.\end{aligned}$$

Proof. We apply Lemma 5.1 for the spectral norm $\|\cdot\|_2$ and get

$$\begin{aligned}\|\mathbf{H}_r(\cdot\,;\mathbf{p})-\widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty}&=\sup_{s\in\imath\mathbb{R}}\|\mathbf{H}_r(s;\mathbf{p})-\widetilde{\mathbf{H}}_r(s;\mathbf{p})\|_2\\ &\leq\sup_{s\in\imath\mathbb{R}}\Big\{\|\mathbf{C}_r(\mathbf{p})\cdot\Delta\mathbf{M}(s)\cdot\widetilde{\mathbf{B}}_r(\mathbf{p})\|_2\\ &\qquad+\|\mathbf{C}_r(\mathbf{p})\cdot\mathbf{M}(s)\|_2\cdot\|\Delta\mathbf{B}_r(\mathbf{p})\|_2\end{aligned}$$

$$
\begin{aligned}
&\qquad + \|\widetilde{\mathbf{M}}(s)\cdot\widetilde{\mathbf{B}}_r(\mathbf{p})\|_2\cdot\|\Delta\mathbf{C}_r(\mathbf{p})\|_2\Big\} \\
&\le \sup_{s\in\imath\mathbb{R}}\|\mathbf{C}_r(\mathbf{p})\cdot\Delta\mathbf{M}(s)\cdot\widetilde{\mathbf{B}}_r(\mathbf{p})\|_2 \\
&\quad + \sup_{s\in\imath\mathbb{R}}\Big\{\|\mathbf{C}_r(\mathbf{p})\cdot\mathbf{M}(s)\|_2\cdot\|\Delta\mathbf{B}_r(\mathbf{p})\|_2\Big\} \\
&\quad + \sup_{s\in\imath\mathbb{R}}\Big\{\|\widetilde{\mathbf{M}}(s)\cdot\widetilde{\mathbf{B}}_r(\mathbf{p})\|_2\cdot\|\Delta\mathbf{C}_r(\mathbf{p})\|_2\Big\} \\
&= \underbrace{\sup_{s\in\imath\mathbb{R}}\|\mathbf{C}_r(\mathbf{p})\cdot\Delta\mathbf{M}(s)\cdot\widetilde{\mathbf{B}}_r(\mathbf{p})\|_2}_{=:f_{\mathbf{M}}(\mathbf{p})} \\
&\quad + \underbrace{\sup_{s\in\imath\mathbb{R}}\|\mathbf{C}_r(\mathbf{p})\cdot\mathbf{M}(s)\|_2\cdot\|\Delta\mathbf{B}_r(\mathbf{p})\|_2}_{=:f_{\mathbf{B}}(\mathbf{p})} \\
&\quad + \underbrace{\sup_{s\in\imath\mathbb{R}}\|\widetilde{\mathbf{M}}(s)\cdot\widetilde{\mathbf{B}}_r(\mathbf{p})\|_2\cdot\|\Delta\mathbf{C}_r(\mathbf{p})\|_2}_{=:f_{\mathbf{C}}(\mathbf{p})}.
\end{aligned}
$$

Finally, we use Proposition 5.7 to bound $f_{\mathbf{M}}(\mathbf{p})$ and Proposition 5.8 to bound $f_{\mathbf{B}}(\mathbf{p})$ and $f_{\mathbf{C}}(\mathbf{p})$. □

In subsection 7.3.1 we perform some numerical experiments to see how the bound from Theorem 5.2 behaves in practice.

Since the bound still contains the unknown quantities $\widetilde{\mathbf{A}}_r(\mathbf{p})$ and $\widetilde{\mathbf{B}}_r(\mathbf{p})$, we replace them in the following corollary.

Corollary 5.3. *Let $\mathbf{A}_r(\mathbf{p})$ and $\widetilde{\mathbf{A}}_r(\mathbf{p})$ be symmetric negative definite matrices. Define $\mathfrak{A}(\mathbf{p}) = \log(-\mathbf{A}_r(\mathbf{p}))$ and $\widetilde{\mathfrak{A}}(\mathbf{p}) = \log(-\widetilde{\mathbf{A}}_r(\mathbf{p}))$. Then*

$$
\begin{aligned}
&\|\mathbf{H}_r(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty} \\
&\le \|\mathbf{A}_r(\mathbf{p})^{-1}\|_2\cdot\Big\{\alpha(\mathbf{p})\cdot(e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2}-1) + \beta(\mathbf{p})\cdot\|\Delta\mathbf{B}_r(\mathbf{p})\|_2 + \gamma(\mathbf{p})\cdot\|\Delta\mathbf{C}_r(\mathbf{p})\|_2\Big\} \\
&\approx \|\mathbf{A}_r(\mathbf{p})^{-1}\|_2\cdot\Big\{\alpha(\mathbf{p})\cdot\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2 + \beta(\mathbf{p})\cdot\|\Delta\mathbf{B}_r(\mathbf{p})\|_2 + \gamma(\mathbf{p})\cdot\|\Delta\mathbf{C}_r(\mathbf{p})\|_2\Big\} \\
&\qquad \text{for small } \|\Delta\mathfrak{A}_r(\mathbf{p})\|_2,
\end{aligned}
$$

for

$$
\begin{aligned}
\alpha(\mathbf{p}) &:= (\|\mathbf{B}_r(\mathbf{p})\|_2 + \|\Delta\mathbf{B}_r(\mathbf{p})\|_2)\cdot\|\mathbf{C}_r(\mathbf{p})\|_2, \\
\beta(\mathbf{p}) &:= \|\mathbf{C}_r(\mathbf{p})\|_2, \\
\gamma(\mathbf{p}) &:= e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2}\cdot(\|\mathbf{B}_r(\mathbf{p})\|_2 + \|\Delta\mathbf{B}_r(\mathbf{p})\|_2).
\end{aligned}
$$

Proof. The claim follows from Theorem 5.2 and the inequalities

$$
\|\widetilde{\mathbf{B}}_r(\mathbf{p})\|_2 \le \|\mathbf{B}_r(\mathbf{p})\|_2 + \|\Delta\mathbf{B}_r(\mathbf{p})\|_2
$$

and

$$\|\widetilde{\mathbf{A}}_r(\mathbf{p})^{-1}\|_2 \leq \|\mathbf{A}_r(\mathbf{p})^{-1}\|_2 \cdot e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2}.$$

The second inequality stems from Lemma 5.16. □

First, we deal with the first term of the right hand side of bound (5.1) in Lemma 5.1, which belongs to $\mathbf{M}$, and hence to $\mathbf{A}$. Later, we deal with the other two terms belonging to $\mathbf{B}$ and $\mathbf{C}$, respectively.

5.1.2 The ΔM-term

We analyze $\|\mathbf{C} \cdot \Delta\mathbf{M}(s) \cdot \widetilde{\mathbf{B}}\|$ and use three different approaches to derive an upper bound.

The first is a quite naive approach. It involves a locality restriction and leads to a rather bad bound. For the second approach we assume that $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ commute. This leads to an exact expression and results in a better bound. In the third approach we derive the same bound without the strong assumption that $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ commute.

Motivated by the common denominator approach for scalar fractions, we obtain the following generalization for matrices:

$$\mathbf{X}^{-1} - \mathbf{Y}^{-1} = \mathbf{X}^{-1} \cdot (\mathbf{Y} - \mathbf{X}) \cdot \mathbf{Y}^{-1}. \tag{5.2}$$

In the following, this trick is used several times.

Approach (I): Perturbation of the Inverse

In this first approach we use a theorem from perturbation theory of the matrix inverse.

Proposition 5.4. *Let* $\|(s\mathbf{I} - \mathbf{A})^{-1} \cdot \Delta\mathbf{A}\|_2 \leq 1/2$. *Then*

$$\begin{aligned} &\sup_{s \in \imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_2 \\ &\leq 2 \cdot \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2^2 \cdot \|\widetilde{\mathbf{B}}\|_2 \cdot \|\mathbf{C}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1). \end{aligned}$$

Proof. We have

$$\begin{aligned} &\sup_{s \in \imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_2 \\ &\leq \sup_{s \in \imath\mathbb{R}} \Big\{\|\mathbf{C}\|_2 \cdot \|(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}\|_2\Big\} \\ &= \|\mathbf{C}\|_2 \cdot \sup_{s \in \imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}\|_2 \\ &\leq \|\mathbf{C}\|_2 \cdot 2 \cdot \sup_{s \in \imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1}\|_2^2 \cdot \|\Delta\mathbf{A}\|_2 \cdot \|\widetilde{\mathbf{B}}\|_2 \end{aligned}$$

using Corollary A.22 with $\mathbf{M} = s\mathbf{I} - \mathbf{A}$ and $\mathbf{F} = \Delta\mathbf{A}$ and $\|(s\mathbf{I} - \mathbf{A})^{-1} \cdot \Delta\mathbf{A}\|_2 \leq 1/2$ for the last inequality.

By equation (5.9) from Lemma 5.17 with $\mathbf{Z} = \mathbf{I}$ we yield

$$\begin{aligned}\sup_{s\in\imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1}\|_2^2 \cdot \|\Delta\mathbf{A}\|_2 &= \|\mathbf{A}^{-1}\|_2^2 \cdot \|\Delta\mathbf{A}\|_2 \\ &\leq \|\mathbf{A}^{-1}\|_2^2 \cdot \|\mathbf{A}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1)\end{aligned}$$

with the inequality (5.5) from Lemma 5.15. □

The prerequisite $\|(s\mathbf{I} - \mathbf{A})^{-1} \cdot \Delta\mathbf{A}\|_2 \leq 1/2$ is a strong locality restriction, which we want to eliminate. Furthermore, the factor $\|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2^2 = \kappa_2(\mathbf{A}) \cdot \|\mathbf{A}^{-1}\|_2$ is rather big.

Approach (II): Assumption: $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ commute

This approach is based on the assumption that $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ commute, which leads to a significantly better bound. This assumption implies that $\mathfrak{A}$ and $\widetilde{\mathfrak{A}}$ commute.

In the following lemma we manipulate $\mathbf{C} \cdot \Delta\mathbf{M}(s) \cdot \widetilde{\mathbf{B}}$ to introduce $\Delta\mathfrak{A}$.

Lemma 5.5. *If* $\mathbf{A}$ *and* $\widetilde{\mathbf{A}}$ *commute, then*

$$\mathbf{C} \cdot \left\{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\right\} \cdot \widetilde{\mathbf{B}} = \mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1} \cdot (e^{\Delta\mathfrak{A}} - \mathbf{I}) \cdot (s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \cdot \widetilde{\mathbf{B}}.$$

Proof. We have

$$\begin{aligned}&(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \\ &\overset{(5.2)}{=} (s\mathbf{I} - \mathbf{A})^{-1} \cdot \left\{(s\mathbf{I} - \widetilde{\mathbf{A}}) - (s\mathbf{I} - \mathbf{A})\right\} \cdot (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \\ &= (s\mathbf{I} - \mathbf{A})^{-1} \cdot (\mathbf{A} - \widetilde{\mathbf{A}}) \cdot (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \\ &= (s\mathbf{I} - \mathbf{A})^{-1} \cdot (\mathbf{A} \cdot \widetilde{\mathbf{A}}^{-1} - \mathbf{I}) \cdot \widetilde{\mathbf{A}} \cdot (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \\ &= (s\mathbf{I} - \mathbf{A})^{-1} \cdot (\mathbf{A} \cdot \widetilde{\mathbf{A}}^{-1} - \mathbf{I}) \cdot (s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \\ &= (s\mathbf{I} - \mathbf{A})^{-1} \cdot \left\{\left(-\exp(\mathfrak{A})\right) \cdot \left(-\exp(-\widetilde{\mathfrak{A}})\right) - \mathbf{I}\right\} \cdot (s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \\ &= (s\mathbf{I} - \mathbf{A})^{-1} \cdot \left\{\exp(\mathfrak{A}) \cdot \exp(-\widetilde{\mathfrak{A}}) - \mathbf{I}\right\} \cdot (s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \\ &= (s\mathbf{I} - \mathbf{A})^{-1} \cdot \left\{\exp(\mathfrak{A} - \widetilde{\mathfrak{A}}) - \mathbf{I}\right\} \cdot (s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1}\end{aligned}$$

using the requirement that $\mathfrak{A}$ and $\widetilde{\mathfrak{A}}$ commute only in the last equality. □

Using submultiplicativity, we can eliminate the need for taking the supremum over $\imath\mathbb{R}$.

Proposition 5.6. *If* $\mathbf{A}$ *and* $\widetilde{\mathbf{A}}$ *commute, then*

$$\begin{aligned}\sup_{s\in\imath\mathbb{R}} \|\mathbf{C} \cdot \left\{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\right\} \cdot \widetilde{\mathbf{B}}\|_2 &\leq \|\widetilde{\mathbf{B}}\|_2 \cdot \|\mathbf{C}\mathbf{A}^{-1}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1) \\ &\leq \|\mathbf{A}^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}\|_2 \cdot \|\mathbf{C}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1).\end{aligned}$$

Proof. First, we apply Lemma 5.5 and yield

$$\begin{aligned}
&\sup_{s\in\imath\mathbb{R}} \|\mathbf{C}\cdot\{(s\mathbf{I}-\mathbf{A})^{-1}-(s\mathbf{I}-\widetilde{\mathbf{A}})^{-1}\}\cdot\widetilde{\mathbf{B}}\|_2\\
&=\sup_{s\in\imath\mathbb{R}} \|\mathbf{C}\cdot(s\mathbf{I}-\mathbf{A})^{-1}\cdot(e^{\Delta\mathfrak{A}}-\mathbf{I})\cdot(s\widetilde{\mathbf{A}}^{-1}-\mathbf{I})^{-1}\cdot\widetilde{\mathbf{B}}\|_2\\
&\le\sup_{s\in\imath\mathbb{R}}\Big\{\|\mathbf{C}\cdot(s\mathbf{I}-\mathbf{A})^{-1}\|_2\cdot\|e^{\Delta\mathfrak{A}}-\mathbf{I}\|_2\cdot\|(s\widetilde{\mathbf{A}}^{-1}-\mathbf{I})^{-1}\cdot\widetilde{\mathbf{B}}\|_2\Big\}\\
&\le\sup_{s\in\imath\mathbb{R}}\|(s\mathbf{I}-\mathbf{A})^{-1}\cdot\mathbf{C}^T\|_2\cdot\|e^{\Delta\mathfrak{A}}-\mathbf{I}\|_2\cdot\sup_{s\in\imath\mathbb{R}}\|(s\widetilde{\mathbf{A}}^{-1}-\mathbf{I})^{-1}\cdot\widetilde{\mathbf{B}}\|_2.
\end{aligned}$$

Using equation (5.9) and (5.8) from Lemma 5.17 with $\mathbf{Z}=\mathbf{C}^T$ and with $\mathbf{Z}=\widetilde{\mathbf{B}}$, respectively, and making use of the symmetry of $\mathbf{A}$ we obtain

$$\begin{aligned}
\sup_{s\in\imath\mathbb{R}} \|\mathbf{C}\cdot\{(s\mathbf{I}-\mathbf{A})^{-1}-(s\mathbf{I}-\widetilde{\mathbf{A}})^{-1}\}\cdot\widetilde{\mathbf{B}}\|_2 &\le \|\mathbf{A}^{-1}\mathbf{C}^T\|_2\cdot\|e^{\Delta\mathfrak{A}}-\mathbf{I}\|_2\cdot\|\widetilde{\mathbf{B}}\|_2\\
&=\|\mathbf{C}\mathbf{A}^{-1}\|_2\cdot\|e^{\Delta\mathfrak{A}}-\mathbf{I}\|_2\cdot\|\widetilde{\mathbf{B}}\|_2\\
&\le\|\mathbf{C}\mathbf{A}^{-1}\|_2\cdot(e^{\|\Delta\mathfrak{A}\|_2}-1)\cdot\|\widetilde{\mathbf{B}}\|_2
\end{aligned}$$

with Corollary A.24. □

Compared to Approach (I) with perturbation of the inverse (see Proposition 5.4), we essentially gain a factor of $\kappa_2(\mathbf{A})=\|\mathbf{A}\|_2\cdot\|\mathbf{A}^{-1}\|_2$. Instead of the locality condition $\|(s\mathbf{I}-\mathbf{A})^{-1}\cdot\Delta\mathbf{A}\|\le 1/2$ we need the commutativity of $\mathbf{A}$ and $\widetilde{\mathbf{A}}$.

The assumption that $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ commute is rather strong and our interpolation framework cannot guarantee it. In the next part we show that it is possible to derive essentially the same bound as in Proposition 5.6 without this assumption.

Approach (III): No Assumption

The common denominator trick yields

$$\begin{aligned}
(s\mathbf{I}-\mathbf{A})^{-1}-(s\mathbf{I}-\widetilde{\mathbf{A}})^{-1}&=(s\mathbf{I}-\mathbf{A})^{-1}\cdot\Big\{(s\mathbf{I}-\widetilde{\mathbf{A}})-(s\mathbf{I}-\mathbf{A})\Big\}\cdot(s\mathbf{I}-\widetilde{\mathbf{A}})^{-1}\\
&=(s\mathbf{I}-\mathbf{A})^{-1}\cdot(\mathbf{A}-\widetilde{\mathbf{A}})\cdot(s\mathbf{I}-\widetilde{\mathbf{A}})^{-1}. \qquad (5.3)
\end{aligned}$$

If we would apply submultiplicativity for bounding $\|(s\mathbf{I}-\mathbf{A})^{-1}-(s\mathbf{I}-\widetilde{\mathbf{A}})^{-1}\|$, particularly for $s=0$, then we would obtain

$$\begin{aligned}
\|\mathbf{A}^{-1}-\widetilde{\mathbf{A}}^{-1}\| &= \|(s\mathbf{I}-\mathbf{A})^{-1}-(s\mathbf{I}-\widetilde{\mathbf{A}})^{-1}\|\Big|_{s=0}\\
&=\|(s\mathbf{I}-\mathbf{A})^{-1}\cdot(\mathbf{A}-\widetilde{\mathbf{A}})\cdot(s\mathbf{I}-\widetilde{\mathbf{A}})^{-1}\|\Big|_{s=0}\\
&\le\|\mathbf{A}^{-1}\|\cdot\|\mathbf{A}-\widetilde{\mathbf{A}}\|\cdot\|\widetilde{\mathbf{A}}^{-1}\|.
\end{aligned}$$

This looks unfavorable and indeed, for the $\mathcal{H}_\infty$-norm case (see Proposition B.16) we would lose a factor of $\|\mathbf{A}\|_2\cdot\|\widetilde{\mathbf{A}}^{-1}\|_2\approx\kappa_2(\mathbf{A})$ compared to Proposition 5.6 as with the perturbation of the inverse approach. Hence, we aim to achieve a formula expressed in terms of $\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}$ on the right hand side, which finally leads us to the desired bound.

This motivates the following remark: If $\mathbf{A}$ is not symmetric negative definite (and hence we cannot interpolate in the tangent space of a matrix manifold), then the interpolation of $\mathbf{A}^{-1}$ might be preferable to the interpolation of $\mathbf{A}$. However, the occurrence of instable interpolated reduced order models makes this study quite impossible.

To introduce $\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}$ on the right hand side, we apply the common denominator trick to equation (5.3) a second time for $\mathbf{X} = \mathbf{A}^{-1}$ and $\mathbf{Y} = \widetilde{\mathbf{A}}^{-1}$. This yields

$$\begin{aligned}(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} &= (s\mathbf{I} - \mathbf{A})^{-1} \cdot (\mathbf{A} - \widetilde{\mathbf{A}}) \cdot (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \\ &= (s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{A} \cdot (\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}) \cdot \widetilde{\mathbf{A}} \cdot (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \\ &= (s\mathbf{A}^{-1} - \mathbf{I})^{-1} \cdot (\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}) \cdot (s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1}. \end{aligned} \tag{5.4}$$

For the $\mathcal{H}_\infty$-norm case we can show:

Proposition 5.7. *It holds*

$$\sup_{s \in \imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_2 \le \|\mathbf{A}^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}\|_2 \cdot \|\mathbf{C}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1).$$

Proof. Plugging in equation (5.4), we get

$$\begin{aligned}&\sup_{s \in \imath\mathbb{R}} \|\mathbf{C} \cdot \left\{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\right\} \cdot \widetilde{\mathbf{B}}\|_2 \\ &\le \sup_{s \in \imath\mathbb{R}} \|\mathbf{C} \cdot (s\mathbf{A}^{-1} - \mathbf{I})^{-1} \cdot (\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}) \cdot (s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \cdot \widetilde{\mathbf{B}}\|_2 \\ &\le \sup_{s \in \imath\mathbb{R}} \left(\|\mathbf{C} \cdot (s\mathbf{A}^{-1} - \mathbf{I})^{-1}\|_2 \cdot \|\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}\|_2 \cdot \|(s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \cdot \widetilde{\mathbf{B}}\|_2\right) \\ &\le \sup_{s \in \imath\mathbb{R}} \|(s\mathbf{A}^{-1} - \mathbf{I})^{-1} \cdot \mathbf{C}^T\|_2 \cdot \|\mathbf{A}^{-1} - \widetilde{\mathbf{A}}^{-1}\|_2 \cdot \sup_{s \in \imath\mathbb{R}} \|(s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \cdot \widetilde{\mathbf{B}}\|_2.\end{aligned}$$

Using equation (5.8) from Lemma 5.17 with $\mathbf{Z} = \mathbf{C}^T$ and with $\mathbf{Z} = \widetilde{\mathbf{B}}$, respectively, we obtain

$$\begin{aligned}\sup_{s \in \imath\mathbb{R}} \left\|\mathbf{C} \cdot \left\{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\right\} \cdot \widetilde{\mathbf{B}}\right\|_2 &\le \|\mathbf{C}^T\|_2 \cdot \|\mathbf{A}^{-1} - \widetilde{\mathbf{A}}^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}\|_2 \\ &= \|\mathbf{C}\|_2 \cdot \|\mathbf{A}^{-1} - \widetilde{\mathbf{A}}^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}\|_2 \\ &\le \|\mathbf{C}\|_2 \cdot \|\mathbf{A}^{-1}\|_2 \cdot (e^{\|\mathfrak{A} - \widetilde{\mathfrak{A}}\|_2} - 1) \cdot \|\widetilde{\mathbf{B}}\|_2\end{aligned}$$

with inequality (5.6) from Lemma 5.15. □

Comparing Proposition 5.6 and Proposition 5.7, we see they give essentially the same result. But in Proposition 5.7 we do not need the strong assumption that $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ commute.

5.1.3 The ΔB- and the ΔC-term

To finish the proof of Theorem 5.2 it remains to bound $\sup_{s\in\imath\mathbb{R}}\|\mathbf{C}\cdot(s\mathbf{I}-\mathbf{A})^{-1}\|_2$ and $\sup_{s\in\imath\mathbb{R}}\|(s\mathbf{I}-\widetilde{\mathbf{A}})^{-1}\cdot\widetilde{\mathbf{B}}\|_2$. We show the following.

Proposition 5.8. *It holds*

$$\sup_{s\in\imath\mathbb{R}}\|\mathbf{C}\cdot(s\mathbf{I}-\mathbf{A})^{-1}\|_2 = \|\mathbf{C}\cdot\mathbf{A}^{-1}\|_2 \quad \text{and}$$
$$\sup_{s\in\imath\mathbb{R}}\|(s\mathbf{I}-\widetilde{\mathbf{A}})^{-1}\cdot\widetilde{\mathbf{B}}\|_2 = \|\widetilde{\mathbf{A}}^{-1}\cdot\widetilde{\mathbf{B}}\|_2.$$

Proof. Apply equation (5.9) from Lemma 5.17 with $\mathbf{Z}=\mathbf{C}^T$ and with $\mathbf{Z}=\widetilde{\mathbf{B}}$, respectively. □

5.2 Bound for $\mathcal{H}_2$-norm at Fixed Point

In this section we derive an upper bound for the interpolation error in the $\mathcal{H}_2$-norm at a fixed point $\mathbf{p}$ in parameter space. Recall that

$$\|\mathbf{H}_r(\cdot\,;\mathbf{p})-\widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2} = \left(\frac{1}{2\pi}\int_{\imath\mathbb{R}}\|\mathbf{H}_r(s;\mathbf{p})-\widetilde{\mathbf{H}}_r(s;\mathbf{p})\|_F^2\ \mathrm{d}s\right)^{1/2}.$$

In section 5.4 we will use this bound to derive an upper bound for the interpolation error in the $\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)$-norm. Recall that

$$\|\mathbf{H}_r-\widetilde{\mathbf{H}}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)} = \left(\int_{\mathcal{P}}\|\mathbf{H}_r(\cdot\,;\mathbf{p})-\widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2}^2\ \mathrm{d}\mathbf{p}\right)^{1/2}.$$

5.2.1 Final Theorem

We prove an upper bound for the interpolation error of the transfer function in the $\mathcal{H}_2$-norm in dependence on the error of the interpolated matrices in the Frobenius norm and the spectral norm.

The analysis is very similar to the $\mathcal{H}_\infty$-norm case (see section 5.1). However, the treatment of the integral is more involved. At some points, the analysis of the Frobenius norm requires to switch to the spectral norm via matrix norm inequalities (2.5).

Theorem 5.9 (cf. Theorem 5.2)**.** *Let $\mathbf{A}_r(\mathbf{p})$ and $\widetilde{\mathbf{A}}_r(\mathbf{p})$ be symmetric negative definite matrices. Define $\mathfrak{A}_r(\mathbf{p})=\log(-\mathbf{A}_r(\mathbf{p}))$ and $\widetilde{\mathfrak{A}}_r(\mathbf{p})=\log(-\widetilde{\mathbf{A}}_r(\mathbf{p}))$. Then*

$$\begin{aligned}&\|\mathbf{H}_r(\cdot\,;\mathbf{p})-\widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2}\\ &\le \widetilde{\alpha}(\mathbf{p})\cdot(e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2}-1)+\widetilde{\beta}(\mathbf{p})\cdot\|\Delta\mathbf{B}_r(\mathbf{p})\|_F+\widetilde{\gamma}(\mathbf{p})\cdot\|\Delta\mathbf{C}_r(\mathbf{p})\|_F\end{aligned}$$

for

$$\widetilde{\alpha}(\mathbf{p}) := \frac{\sqrt{r}}{\sqrt{2}}\cdot\sqrt{\|\mathbf{A}_r(\mathbf{p})\|_2}\cdot\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2\cdot\|\widetilde{\mathbf{B}}_r(\mathbf{p})\|_F\cdot\|\mathbf{C}_r(\mathbf{p})\|_F,$$

$$\widetilde{\beta}(\mathbf{p}) := \frac{1}{\sqrt{2}} \cdot \|\mathbf{C}_r(\mathbf{p}) \cdot (-\mathbf{A}_r(\mathbf{p}))^{-1/2}\|_F \leq \frac{1}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2} \cdot \|\mathbf{C}_r(\mathbf{p})\|_F,$$
$$\widetilde{\gamma}(\mathbf{p}) := \frac{1}{\sqrt{2}} \cdot \|(-\widetilde{\mathbf{A}}_r(\mathbf{p}))^{-1/2} \cdot \widetilde{\mathbf{B}}_r(\mathbf{p})\|_F \leq \frac{1}{\sqrt{2}} \cdot \sqrt{\|\widetilde{\mathbf{A}}_r(\mathbf{p})^{-1}\|_2} \cdot \|\widetilde{\mathbf{B}}_r(\mathbf{p})\|_F.$$

Proof. We apply Lemma 5.1 for the FROBENIUS norm and yield

$$\begin{aligned}
\|\mathbf{H}_r(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2} &= \left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\mathbf{H}_r(s;\mathbf{p}) - \widetilde{\mathbf{H}}_r(s;\mathbf{p})\|_F^2 \,\mathrm{d}s\right)^{1/2} \\
&\leq \left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \Big\{\|\mathbf{C}_r(\mathbf{p}) \cdot \Delta\mathbf{M}(s) \cdot \widetilde{\mathbf{B}}_r(\mathbf{p})\|_F \right. \\
&\qquad + \|\mathbf{C}_r(\mathbf{p}) \cdot \mathbf{M}(s)\|_F \cdot \|\Delta\mathbf{B}_r(\mathbf{p})\|_F \\
&\qquad \left. + \|\widetilde{\mathbf{M}}(s) \cdot \widetilde{\mathbf{B}}_r(\mathbf{p})\|_F \cdot \|\Delta\mathbf{C}_r(\mathbf{p})\|_F\Big\}^2 \,\mathrm{d}s\right)^{1/2} \\
&\leq \left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\mathbf{C}_r(\mathbf{p}) \cdot \Delta\mathbf{M}(s) \cdot \widetilde{\mathbf{B}}_r(\mathbf{p})\|_F^2 \,\mathrm{d}s\right)^{1/2} \\
&\quad + \left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\mathbf{C}_r(\mathbf{p}) \cdot \mathbf{M}(s)\|_F^2 \cdot \|\Delta\mathbf{B}_r(\mathbf{p})\|_F^2 \,\mathrm{d}s\right)^{1/2} \\
&\quad + \left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\widetilde{\mathbf{M}}(s) \cdot \widetilde{\mathbf{B}}_r(\mathbf{p})\|_F^2 \cdot \|\Delta\mathbf{C}_r(\mathbf{p})\|_F^2 \,\mathrm{d}s\right)^{1/2} \\
&= \underbrace{\left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\mathbf{C}_r(\mathbf{p}) \cdot \Delta\mathbf{M}(s) \cdot \widetilde{\mathbf{B}}_r(\mathbf{p})\|_F^2 \,\mathrm{d}s\right)^{1/2}}_{=: f_{\mathbf{M}}(\mathbf{p})} \\
&\quad + \underbrace{\left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\mathbf{C}_r(\mathbf{p}) \cdot \mathbf{M}(s)\|_F^2 \,\mathrm{d}s\right)^{1/2} \cdot \|\Delta\mathbf{B}_r(\mathbf{p})\|_F}_{=: f_{\mathbf{B}}(\mathbf{p})} \\
&\quad + \underbrace{\left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\widetilde{\mathbf{M}}(s) \cdot \widetilde{\mathbf{B}}_r(\mathbf{p})\|_F^2 \,\mathrm{d}s\right)^{1/2} \cdot \|\Delta\mathbf{C}_r(\mathbf{p})\|_F}_{=: f_{\mathbf{C}}(\mathbf{p})}.
\end{aligned}$$

Then we use Proposition 5.11 to bound $f_{\mathbf{M}}(\mathbf{p})$ and Proposition 5.12 to bound $f_{\mathbf{B}}(\mathbf{p})$ and $f_{\mathbf{C}}(\mathbf{p})$. This gives the desired result with $\widetilde{\alpha}(\mathbf{p})$, $\widetilde{\beta}(\mathbf{p})$ and $\widetilde{\gamma}(\mathbf{p})$.

The upper bounds for $\widetilde{\beta}(\mathbf{p})$ and $\widetilde{\gamma}(\mathbf{p})$ follow from Lemma 5.18. □

Corollary 5.10 (cf. Corollary 5.3)**.** *Let* $\mathbf{A}_r(\mathbf{p})$ *and* $\widetilde{\mathbf{A}}_r(\mathbf{p})$ *be symmetric negative definite matrices. Define* $\mathfrak{A}_r(\mathbf{p}) = \log(-\mathbf{A}_r(\mathbf{p}))$ *and* $\widetilde{\mathfrak{A}}_r(\mathbf{p}) = \log(-\widetilde{\mathbf{A}}_r(\mathbf{p}))$*. Then*

$$\begin{aligned}
\|\mathbf{H}_r(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2} \leq \frac{\sqrt{\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2}}{\sqrt{2}} \cdot \Big\{&\alpha(\mathbf{p}) \cdot (e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2} - 1) \\
&+ \beta(\mathbf{p}) \cdot \|\Delta\mathbf{B}_r(\mathbf{p})\|_F \\
&+ \gamma(\mathbf{p}) \cdot \|\Delta\mathbf{C}_r(\mathbf{p})\|_F\Big\}
\end{aligned}$$

$$\approx \frac{\sqrt{\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2}}{\sqrt{2}} \cdot \Big\{\alpha(\mathbf{p}) \cdot \|\Delta\mathfrak{A}_r(\mathbf{p})\|_2 + \beta(\mathbf{p}) \cdot \|\Delta\mathbf{B}_r(\mathbf{p})\|_F + \gamma(\mathbf{p}) \cdot \|\Delta\mathbf{C}_r(\mathbf{p})\|_F\Big\}$$
$$\text{for small } \|\Delta\mathfrak{A}_r(\mathbf{p})\|_2,$$

for

$$\begin{aligned}
\alpha(\mathbf{p}) &:= \sqrt{r} \cdot \sqrt{\kappa_2(\mathbf{A}_r(\mathbf{p}))} \cdot \Big(\|\mathbf{B}_r(\mathbf{p})\|_F + \|\Delta\mathbf{B}_r(\mathbf{p})\|_F\Big) \cdot \|\mathbf{C}_r(\mathbf{p})\|_F,\\
\beta(\mathbf{p}) &:= \|\mathbf{C}_r(\mathbf{p})\|_F,\\
\gamma(\mathbf{p}) &:= \sqrt{e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2}} \cdot \Big(\|\mathbf{B}_r(\mathbf{p})\|_F + \|\Delta\mathbf{B}_r(\mathbf{p})\|_F\Big).
\end{aligned}$$

Proof. The claim follows from Theorem 5.9 and the inequalities

$$\|\widetilde{\mathbf{B}}_r(\mathbf{p})\|_F \le \|\mathbf{B}_r(\mathbf{p})\|_F + \|\Delta\mathbf{B}_r(\mathbf{p})\|_F$$

and

$$\|\widetilde{\mathbf{A}}_r(\mathbf{p})^{-1}\|_2 \le \|\mathbf{A}_r(\mathbf{p})^{-1}\|_2 \cdot e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2}$$

using Lemma 5.16. □

5.2.2 The ΔM-term

To prove Theorem 5.9 we first bound $\int_{\imath\mathbb{R}} \|\mathbf{C} \cdot \Delta\mathbf{M}(s) \cdot \widetilde{\mathbf{B}}\|_F^2 \, \mathrm{d}s$.

Proposition 5.11 (cf. Proposition 5.7)**.** *It holds*

$$\begin{aligned}
&\left(\frac{1}{2\pi}\int_{\imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_F^2 \, \mathrm{d}s\right)^{1/2}\\
&\le \frac{1}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2} \cdot \|\widetilde{\mathbf{B}}\|_F \cdot \|\mathbf{C}\|_F \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1).
\end{aligned}$$

Proof. First, we bound the integrand

$$\begin{aligned}
&\|\mathbf{C} \cdot \Big\{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\Big\} \cdot \widetilde{\mathbf{B}}\|_F\\
&\overset{(5.4)}{=} \|\mathbf{C} \cdot (s\mathbf{A}^{-1} - \mathbf{I})^{-1} \cdot (\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}) \cdot (s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \cdot \widetilde{\mathbf{B}}\|_F\\
&\le \|\mathbf{C} \cdot (s\mathbf{A}^{-1} - \mathbf{I})^{-1} \cdot (\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1})\|_F \cdot \|(s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \cdot \widetilde{\mathbf{B}}\|_F\\
&\le \|\mathbf{C} \cdot (s\mathbf{A}^{-1} - \mathbf{I})^{-1}\|_F \cdot \|\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}\|_2 \cdot \|(s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \cdot \widetilde{\mathbf{B}}\|_F\\
&\le \|\mathbf{C}\|_F \cdot \max_{i=1:\, r} \left|\frac{1}{s/\lambda_i - 1}\right| \cdot \|\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}\|_2 \cdot \max_{i=1:\, r} \left|\frac{1}{s/\widetilde{\lambda}_i - 1}\right| \cdot \|\widetilde{\mathbf{B}}\|_F\\
&= \left|\frac{1}{s/\lambda_{\min} - 1}\right| \cdot \left|\frac{1}{s/\widetilde{\lambda}_{\min} - 1}\right| \cdot \|\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}\|_2 \cdot \|\mathbf{C}\|_F \cdot \|\widetilde{\mathbf{B}}\|_F
\end{aligned}$$

using Lemma 5.18 for the second inequality and Lemma 5.19 for $f(\mathbf{M}) = (s \cdot \mathbf{M}^{-1} - \mathbf{I})^{-1}$ with $\mathbf{M} = \mathbf{A}$, $\mathbf{N}^T = \mathbf{C}$ and $\mathbf{M} = \widetilde{\mathbf{A}}$, $\mathbf{N} = \widetilde{\mathbf{B}}$ respectively, for the third inequality. The last equality holds since $s \in \imath\mathbb{R}$.

Next, we bound the whole integral

$$\begin{aligned}
&\int_{\imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_F^2 \,\mathrm{d}s \\
&\le \int_{\imath\mathbb{R}} \left|\frac{1}{s/\lambda_{\min} - 1}\right|^2 \cdot \left|\frac{1}{s/\widetilde{\lambda}_{\min} - 1}\right|^2 \cdot \|\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}\|_2^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \,\mathrm{d}s \\
&= \int_{\imath\mathbb{R}} \left|\frac{1}{s/\lambda_{\min} - 1}\right|^2 \cdot \left|\frac{1}{s/\widetilde{\lambda}_{\min} - 1}\right|^2 \,\mathrm{d}s \cdot \|\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}\|_2^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \\
&= \pi \cdot \frac{|\lambda_{\min}| \cdot |\widetilde{\lambda}_{\min}|}{|\lambda_{\min}| + |\widetilde{\lambda}_{\min}|} \cdot \|\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}\|_2^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \\
&\le \pi \cdot \frac{|\lambda_{\min}| \cdot |\widetilde{\lambda}_{\min}|}{|\lambda_{\min}| + |\widetilde{\lambda}_{\min}|} \cdot \|\mathbf{A}^{-1}\|_2^2 \cdot (e^{\|\mathfrak{A} - \widetilde{\mathfrak{A}}\|_2} - 1)^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \\
&\le \pi \cdot |\lambda_{\min}| \cdot \|\mathbf{A}^{-1}\|_2^2 \cdot (e^{\|\mathfrak{A} - \widetilde{\mathfrak{A}}\|_2} - 1)^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \\
&= \pi \cdot \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2^2 \cdot (e^{\|\mathfrak{A} - \widetilde{\mathfrak{A}}\|_2} - 1)^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2
\end{aligned}$$

using equation (5.12) in Lemma 5.20 for the first equation and inequality (5.6) in Lemma 5.15 for the second inequality. □

The bound from Proposition 5.11 is a factor $\sqrt{\kappa_2(\mathbf{A})}$ worse than the bound we would achieve with the assumption that $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ commute (see Proposition B.18). However, it is by a factor of $\sqrt{\kappa_2(\mathbf{A})}$ better than the approach via $\mathbf{A} - \widetilde{\mathbf{A}}$ (see Proposition B.19). The latter is as bad as the perturbation of the matrix inverse approach (see Proposition B.17).

5.2.3 The △B- and the △C-term

To finish the proof of Theorem 5.9 it remains to bound $\int_{\imath\mathbb{R}} \|\mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1}\|_F^2 \,\mathrm{d}s$ and $\int_{\imath\mathbb{R}} \|(s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \cdot \widetilde{\mathbf{B}}\|_F^2 \,\mathrm{d}s$. We show the following.

Proposition 5.12 (cf. Proposition 5.8). *It holds*

$$\begin{aligned}
\left(\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1}\|_F^2 \,\mathrm{d}s\right)^{1/2} &= \frac{1}{\sqrt{2}} \cdot \|\mathbf{C} \cdot (-\mathbf{A})^{-1/2}\|_F \quad \textit{and} \\
\left(\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|(s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \cdot \widetilde{\mathbf{B}}\|_F^2 \,\mathrm{d}s\right)^{1/2} &= \frac{1}{\sqrt{2}} \cdot \|(-\widetilde{\mathbf{A}})^{-1/2} \cdot \widetilde{\mathbf{B}}\|_F.
\end{aligned}$$

Proof. We have

$$\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1}\|_F^2 \,\mathrm{d}s = \frac{1}{2\pi} \int_{\imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{C}^T\|_F^2 \,\mathrm{d}s$$

$$= \frac{1}{2\pi} \cdot \pi \cdot \|(-\mathbf{A})^{-1/2} \cdot \mathbf{C}^T\|_F^2$$
$$= \frac{1}{2} \cdot \|\mathbf{C} \cdot (-\mathbf{A})^{-1/2}\|_F^2$$

evaluating the integral by applying equation (5.14) from Lemma 5.21 with $\mathbf{Z} = \mathbf{C}^T$.

Applying equation (5.14) from Lemma 5.21 with $\mathbf{Z} = \widetilde{\mathbf{B}}$ we similarly yield the second result. □

5.3 Bound for $\mathcal{L}_\infty(\mathcal{P}, \mathcal{H}_\infty)$-norm

In section 5.1 we derived an upper bound for the interpolation error at a fixed point in parameter space. Now we bound the interpolation error over the whole parameter space.

From the final theorem for the interpolation error at a fixed parameter point in the $\mathcal{H}_\infty$-norm (see subsection 5.1.1) we deduce an upper bound for the interpolation error over the parameter space in the $\mathcal{L}_\infty(\mathcal{P}, \mathcal{H}_\infty)$-norm. Note that the $\mathcal{L}_\infty(\mathcal{P}, \mathcal{H}_\infty)$-norm of the transfer functions is well-defined (see section 3.3).

Theorem 5.13. *It holds*

$$\begin{aligned}&\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} \\ &\le \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \cdot \Big\{\alpha \cdot (e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} - 1) + \beta \cdot \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty} + \gamma \cdot \|\Delta\mathbf{C}_r\|_{\mathcal{L}_\infty}\Big\}\end{aligned}$$

for

$$\begin{aligned}\alpha &:= \big(\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}\big) \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty},\\ \beta &:= \|\mathbf{C}_r\|_{\mathcal{L}_\infty},\\ \gamma &:= e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} \cdot \big(\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}\big).\end{aligned}$$

Proof. Applying Corollary 5.3 we get

$$\begin{aligned}&\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} \\ &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{H}_r(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty} \\ &\le \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \Big(\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2 \cdot \Big\{\big(\|\mathbf{B}_r(\mathbf{p})\|_2 + \|\Delta\mathbf{B}_r(\mathbf{p})\|_2\big) \cdot \|\mathbf{C}_r(\mathbf{p})\|_2 \cdot (e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2} - 1) \\ &\qquad + \|\mathbf{C}_r(\mathbf{p})\|_2 \cdot \|\Delta\mathbf{B}_r(\mathbf{p})\|_2 \\ &\qquad + e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2} \cdot \big(\|\mathbf{B}_r(\mathbf{p})\|_2 + \|\Delta\mathbf{B}_r(\mathbf{p})\|_2\big) \cdot \|\Delta\mathbf{C}_r(\mathbf{p})\|_2\Big\}\Big) \\ &\le \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \cdot \Big\{\big(\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}\big) \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot (e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} - 1) \\ &\qquad + \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty} \\ &\qquad + e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} \cdot \big(\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}\big) \cdot \|\Delta\mathbf{C}_r\|_{\mathcal{L}_\infty}\Big\}.\end{aligned}$$

□

5.4 Bound for $\mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$-norm

In section 5.2 we derived an upper bound for the interpolation error at a fixed point in parameter space. Now we bound the interpolation error over the whole parameter space.

From the final theorem for the interpolation error at a fixed parameter point in the $\mathcal{H}_2$-norm (see subsection 5.2.1) we deduce an upper bound for the interpolation error over the parameter space in the $\mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$-norm. Note that the $\mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$-norm of the transfer functions is well-defined (see section 3.3).

Theorem 5.14. *It holds*

$$\begin{aligned}\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)} &\leq \frac{r \cdot \sqrt{mq}}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}} \cdot \left(\int_{\mathcal{P}} 1 \, \mathrm{d}\mathbf{p}\right)^{1/2} \\ &\quad \cdot \left\{\alpha \cdot (e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} - 1) + \beta \cdot \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty} + \gamma \cdot \|\Delta\mathbf{C}_r\|_{\mathcal{L}_\infty}\right\}\end{aligned}$$

for

$$\begin{aligned}\alpha &:= \sqrt{r} \cdot \sqrt{\|\mathbf{A}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}} \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}) \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty}, \\ \beta &:= \|\mathbf{C}_r\|_{\mathcal{L}_\infty}, \\ \gamma &:= \sqrt{e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}}} \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}).\end{aligned}$$

Proof. Applying Corollary 5.10 we get for an arbitrary $\mathbf{p} \in \mathcal{P}$

$$\begin{aligned}&\|\mathbf{H}_r(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2} \\ &= \frac{1}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2} \\ &\quad \cdot \{\sqrt{r} \cdot \sqrt{\kappa_2(\mathbf{A}_r(\mathbf{p}))} \cdot (\|\mathbf{B}_r(\mathbf{p})\|_F + \|\Delta\mathbf{B}_r(\mathbf{p})\|_F) \cdot \|\mathbf{C}_r(\mathbf{p})\|_F \cdot (e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2} - 1) \\ &\qquad + \|\mathbf{C}_r(\mathbf{p})\|_F \cdot \|\Delta\mathbf{B}_r(\mathbf{p})\|_F \\ &\qquad + \sqrt{e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2}} \cdot (\|\mathbf{B}_r(\mathbf{p})\|_F + \|\Delta\mathbf{B}_r(\mathbf{p})\|_F) \cdot \|\Delta\mathbf{C}_r(\mathbf{p})\|_F\} \\ &\leq \frac{1}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2} \\ &\quad \cdot \{\sqrt{r} \cdot \sqrt{\kappa_2(\mathbf{A}_r(\mathbf{p}))} \cdot \sqrt{rm} \cdot (\|\mathbf{B}_r(\mathbf{p})\|_2 + \|\Delta\mathbf{B}_r(\mathbf{p})\|_2) \cdot \sqrt{rq} \cdot \|\mathbf{C}_r(\mathbf{p})\|_2 \\ &\qquad \cdot (e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2} - 1) \\ &\qquad + \sqrt{rq} \cdot \|\mathbf{C}_r(\mathbf{p})\|_2 \cdot \sqrt{rm} \cdot \|\Delta\mathbf{B}_r(\mathbf{p})\|_2 \\ &\qquad + \sqrt{e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2}} \cdot \sqrt{rm} \cdot (\|\mathbf{B}_r(\mathbf{p})\|_2 + \|\Delta\mathbf{B}_r(\mathbf{p})\|_2) \cdot \sqrt{rq} \cdot \|\Delta\mathbf{C}_r(\mathbf{p})\|_2\} \\ &= \frac{r \cdot \sqrt{mq}}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}_r(\mathbf{p})^{-1}\|_2} \\ &\quad \cdot \{\sqrt{r} \cdot \sqrt{\|\mathbf{A}_r(\mathbf{p})\|_2 \cdot \|\mathbf{A}_r(\mathbf{p})^{-1}\|_2} \cdot (\|\mathbf{B}_r(\mathbf{p})\|_2 + \|\Delta\mathbf{B}_r(\mathbf{p})\|_2) \cdot \|\mathbf{C}_r(\mathbf{p})\|_2 \\ &\qquad \cdot (e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2} - 1)\end{aligned}$$

$$
\begin{aligned}
&\quad + \|\mathbf{C}_r(\mathbf{p})\|_2 \cdot \|\Delta\mathbf{B}_r(\mathbf{p})\|_2 \\
&\quad + \sqrt{e^{\|\Delta\mathfrak{A}_r(\mathbf{p})\|_2}} \cdot (\|\mathbf{B}_r(\mathbf{p})\|_2 + \|\Delta\mathbf{B}_r(\mathbf{p})\|_2) \cdot \|\Delta\mathbf{C}_r(\mathbf{p})\|_2\} \\
&\le \frac{r \cdot \sqrt{mq}}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}} \\
&\cdot \{\sqrt{r} \cdot \sqrt{\|\mathbf{A}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}} \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}) \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \\
&\quad \cdot (e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} - 1) \\
&\quad + \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty} + \sqrt{e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}}} \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}) \cdot \|\Delta\mathbf{C}_r\|_{\mathcal{L}_\infty}\} \\
&= \frac{r \cdot \sqrt{mq}}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}} \cdot \{\alpha \cdot (e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} - 1) + \beta \cdot \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty} + \gamma \cdot \|\Delta\mathbf{C}_r\|_{\mathcal{L}_\infty}\}
\end{aligned}
$$

for

$$
\begin{aligned}
&\alpha := \sqrt{r} \cdot \sqrt{\|\mathbf{A}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}} \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}) \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty}, \\
&\beta := \|\mathbf{C}_r\|_{\mathcal{L}_\infty}, \\
&\gamma := \sqrt{e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}}} \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}).
\end{aligned}
$$

This yields

$$
\begin{aligned}
&\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)} \\
&= \left(\int_{\mathcal{P}} \|\mathbf{H}_r(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2}^2 \,\mathrm{d}\mathbf{p}\right)^{1/2} \\
&\le \Bigg(\int_{\mathcal{P}} \left\{\frac{r \cdot \sqrt{mq}}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}}\right\}^2 \\
&\qquad \cdot \{\alpha \cdot (e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} - 1) + \beta \cdot \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty} + \gamma \cdot \|\Delta\mathbf{C}_r\|_{\mathcal{L}_\infty}\}^2 \,\mathrm{d}\mathbf{p}\Bigg)^{1/2} \\
&= \frac{r \cdot \sqrt{mq}}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}} \\
&\quad \cdot \{\alpha \cdot (e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} - 1) + \beta \cdot \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty} + \gamma \cdot \|\Delta\mathbf{C}_r\|_{\mathcal{L}_\infty}\} \cdot \left(\int_{\mathcal{P}} 1 \,\mathrm{d}\mathbf{p}\right)^{1/2}. \qquad \square
\end{aligned}
$$

5.5 Technical Lemmas

We state some lemmas, which are needed before in this chapter.

The matrices $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ are always required to be symmetric negative definite and we define $\mathfrak{A} = \log(-\mathbf{A})$ and $\widetilde{\mathfrak{A}} = \log(-\widetilde{\mathbf{A}})$ as before.

The transfer function and hence the transfer function error is expressed in $\mathbf{A}$ and not in $\mathfrak{A}$. Thus we need an upper bound involving $\Delta\mathfrak{A}$, which depends on $\mathbf{A}$.

Lemma 5.15. *It holds*

$$
\|\mathbf{A} - \widetilde{\mathbf{A}}\|_2 \le \|\mathbf{A}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1), \tag{5.5}
$$

$$\|\mathbf{A}^{-1} - \widetilde{\mathbf{A}}^{-1}\|_2 \leq \|\mathbf{A}^{-1}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1), \tag{5.6}$$

$$\|\mathbf{A} - \widetilde{\mathbf{A}}\|_F \leq \sqrt{r} \cdot \|\mathbf{A}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1), \tag{5.7}$$

Proof. We begin with proving the first two inequalities (for the spectral norm) simultaneously. It holds

$$\begin{aligned} \mathbf{A}^{\pm 1} - \widetilde{\mathbf{A}}^{\pm 1} &= -\exp(\pm\mathfrak{A}) + \exp(\pm\widetilde{\mathfrak{A}}) \\ &= \exp\big(\pm(\mathfrak{A} - (\mathfrak{A} - \widetilde{\mathfrak{A}}))\big) - \exp(\pm\mathfrak{A}). \end{aligned}$$

Since $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ are symmetric, so are $\mathfrak{A}$ and $\widetilde{\mathfrak{A}}$. Now we apply Theorem A.23 with the symmetric and hence normal matrices $\mathbf{M} = \pm\mathfrak{A}$ and $\mathbf{F} = \mp(\mathfrak{A} - \widetilde{\mathfrak{A}})$ and obtain

$$\begin{aligned} \|\mathbf{A}^{\pm 1} - \widetilde{\mathbf{A}}^{\pm 1}\|_2 &= \|\exp(\mathbf{M} + \mathbf{F}) - \exp(\mathbf{M})\|_2 \\ &\leq \|\exp(\mathbf{M})\|_2 \cdot (e^{\|\mathbf{F}\|_2} - 1) \\ &= \|\exp(\pm\mathfrak{A})\|_2 \cdot (e^{\|\mathfrak{A} - \widetilde{\mathfrak{A}}\|_2} - 1) \\ &= \|-\mathbf{A}^{\pm 1}\|_2 \cdot (e^{\|\mathfrak{A} - \widetilde{\mathfrak{A}}\|_2} - 1). \end{aligned}$$

The inequality for the FROBENIUS norm follows directly from inequality (5.5) for the spectral norm and the matrix norm inequalities (2.5) for the matrix dimension r of the reduced system matrices.

□

Now, an upper bound for $\|\widetilde{\mathbf{A}}^{-1}\|_2$ can easily be deduced.

Lemma 5.16. *It holds*

$$\|\widetilde{\mathbf{A}}^{-1}\|_2 \leq \|\mathbf{A}^{-1}\|_2 \cdot e^{\|\Delta\mathfrak{A}\|_2}.$$

Proof. The proof follows directly from the triangle inequality and inequality (5.6) from Lemma 5.15.

$$\begin{aligned} \|\widetilde{\mathbf{A}}^{-1}\|_2 &\leq \|\mathbf{A}^{-1}\|_2 + \|\mathbf{A}^{-1} - \widetilde{\mathbf{A}}^{-1}\|_2 \\ &\leq \|\mathbf{A}^{-1}\|_2 + \|\mathbf{A}^{-1}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1) \\ &= \|\mathbf{A}^{-1}\|_2 \cdot e^{\|\Delta\mathfrak{A}\|_2}. \end{aligned}$$ □

For the $\mathcal{H}_\infty$-norm case we need the following lemma.

Lemma 5.17. *Let $\mathbf{Z} \in \mathbb{R}^{r \times m}$ be an arbitrary real matrix. Then*

$$\sup_{s \in \imath\mathbb{R}} \|(s\mathbf{A}^{-1} - \mathbf{I})^{-1} \cdot \mathbf{Z}\|_2 = \|\mathbf{Z}\|_2, \tag{5.8}$$

$$\sup_{s \in \imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{Z}\|_2 = \|\mathbf{A}^{-1} \cdot \mathbf{Z}\|_2. \tag{5.9}$$

Lemma 5.17 states that in both cases the supremum is attained at $s = 0$.

Proof of Lemma 5.17. First, we prove equation (5.8).

The matrix $\mathbf{A}$ is symmetric (and negative definite) and hence orthogonally diagonalizable. Let $\lambda_1, \dots, \lambda_r$ be its eigenvalues. We have

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \operatorname{diag}(\lambda_1, \dots, \lambda_r)$$

with $\lambda_i \in \mathbb{R}^-$. Then for $s \in \imath\mathbb{R}$

$$\mathbf{Q}^T (s\mathbf{A}^{-1} - \mathbf{I})^{-1} \mathbf{Q} = \operatorname{diag}\left((s/\lambda_1 - 1)^{-1}, \dots, (s/\lambda_r - 1)^{-1}\right).$$

It holds

$$\begin{aligned}
\sup_{s\in\imath\mathbb{R}} \|(s\mathbf{A}^{-1} - \mathbf{I})^{-1}\|_2 &= \sup_{s\in\imath\mathbb{R}} \|\mathbf{Q}^T (s\mathbf{A}^{-1} - \mathbf{I})^{-1} \mathbf{Q}\|_2 \\
&= \sup_{s\in\imath\mathbb{R}} \left\|\operatorname{diag}\left((s/\lambda_1 - 1)^{-1}, \dots, (s/\lambda_r - 1)^{-1}\right)\right\|_2 \\
&= \sup_{s\in\imath\mathbb{R}} \max_{i=1:\,r} \left|(s/\lambda_i - 1)^{-1}\right| \\
&= 1,
\end{aligned}$$

so with submultiplicativity

$$\begin{aligned}
\sup_{s\in\imath\mathbb{R}} \|(s\mathbf{A}^{-1} - \mathbf{I})^{-1} \cdot \mathbf{Z}\|_2 &\le \sup_{s\in\imath\mathbb{R}} \|(s\mathbf{A}^{-1} - \mathbf{I})^{-1}\|_2 \cdot \|\mathbf{Z}\|_2 \\
&= \|\mathbf{Z}\|_2.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\sup_{s\in\imath\mathbb{R}} \|(s\mathbf{A}^{-1} - \mathbf{I})^{-1} \cdot \mathbf{Z}\|_2 &\ge \|(0 \cdot \mathbf{A}^{-1} - \mathbf{I})^{-1} \cdot \mathbf{Z}\|_2 \\
&= \|\mathbf{Z}\|_2.
\end{aligned}$$

This proves equation (5.8).

Equation (5.9) follows from equation (5.8) with $\mathbf{A}^{-1} \cdot \mathbf{Z}$ instead of $\mathbf{Z}$. □

The next lemma yields an improvement to submultiplicativity for the Frobenius norm.

Lemma 5.18. *Let* $\mathbf{M} \in \mathbb{C}^{r\times r}$ *be unitarily diagonalizable and* $\mathbf{N} \in \mathbb{C}^{r\times k}$*. Then*

$$\max\left\{\|\mathbf{M} \cdot \mathbf{N}\|_F, \|\mathbf{N}^T \cdot \mathbf{M}\|_F\right\} \le \|\mathbf{M}\|_2 \cdot \|\mathbf{N}\|_F.$$

Proof. We only prove the inequality for $\|\mathbf{M} \cdot \mathbf{N}\|_F$. The inequality for $\|\mathbf{N}^T \cdot \mathbf{M}\|_F$ follows because $\|\mathbf{N}^T \cdot \mathbf{M}\|_F = \|\mathbf{M}^T \cdot \mathbf{N}\|_F \le \|\mathbf{M}^T\|_2 \cdot \|\mathbf{N}\|_F = \|\mathbf{M}\|_2 \cdot \|\mathbf{N}\|_F$.

First, we show the claim for the diagonal matrix $\mathbf{D} = \operatorname{diag}(d_1, \dots, d_r)$ instead of $\mathbf{M}$. Let $\mathbf{N} = [n_{ij}]$. It holds

$$\|\mathbf{D} \cdot \mathbf{N}\|_F^2 = \sum_{i,j} |d_i \cdot n_{i,j}|^2$$

$$\begin{aligned} &\leq \max_{i=1:\,r} |d_i|^2 \cdot \sum_{i,j} |n_{i,j}|^2 \\ &= \|\mathbf{D}\|_2^2 \cdot \|\mathbf{N}\|_F^2 \end{aligned}$$

We yield

$$\|\mathbf{D} \cdot \mathbf{N}\|_F \leq \|\mathbf{D}\|_2 \cdot \|\mathbf{N}\|_F. \tag{5.10}$$

Now we prove the inequality for an arbitrary unitarily diagonalizable matrix $\mathbf{M} = \mathbf{UDU}^H$. We have

$$\begin{aligned} \|\mathbf{MN}\|_F &= \|\mathbf{U}^H\mathbf{MUU}^H\mathbf{N}\|_F \\ &= \|\mathbf{DU}^H\mathbf{N}\|_F \\ &\leq \|\mathbf{D}\|_2 \cdot \|\mathbf{U}^H\mathbf{N}\|_F \\ &= \|\mathbf{UDU}^H\|_2 \cdot \|\mathbf{U}^H\mathbf{N}\|_F \\ &= \|\mathbf{M}\|_2 \cdot \|\mathbf{N}\|_F \end{aligned}$$

using inequality (5.10). □

Lemma 5.19. *Let $\mathbf{M} \in \mathbb{C}^{r\times r}$ be unitarily diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_r$ and let f be a function of a matrix (see Definition 2.3. Let $\mathbf{N} \in \mathbb{C}^{r\times k}$. Then*

$$\max\left\{\|f(\mathbf{M}) \cdot \mathbf{N}\|_F, \|\mathbf{N}^T \cdot f(\mathbf{M})\|_F\right\} \leq \max_{i=1:\,r} |f(\lambda_i)| \cdot \|\mathbf{N}\|_F.$$

The same inequality holds for the spectral norm. There it follows immediately from submultiplicativity. For the Frobenius norm, however, we would get a bound that is worse than the bound claimed by Lemma 5.19 by a factor up to $\sqrt{r}$.

Proof of Lemma 5.19. We only prove the inequality for $\|f(\mathbf{M})\cdot\mathbf{N}\|_F$. The inequality for $\|\mathbf{N}^T \cdot f(\mathbf{M})\|_F$ can be derived analogously.

Let $\mathbf{M} = \mathbf{UDU}^H$. Then we have

$$\begin{aligned} \|f(\mathbf{M}) \cdot \mathbf{N}\|_F &= \|\mathbf{U} \cdot \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_r)) \cdot \mathbf{U}^H \cdot \mathbf{N}\|_F \\ &\leq \|\mathbf{U} \cdot \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_r)) \cdot \mathbf{U}^H\|_2 \cdot \|\mathbf{N}\|_F \\ &= \|\operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_r))\|_2 \cdot \|\mathbf{N}\|_F \\ &= \max_{i=1:\,r} |f(\lambda_i)| \cdot \|\widetilde{\mathbf{N}}\|_F \end{aligned}$$

where the inequality is due to Lemma 5.18. The claim follows from this and $\|\widetilde{\mathbf{N}}\|_F = \|\mathbf{U}^H \cdot \mathbf{N}\|_F = \|\mathbf{N}\|_F$. □

For the $\mathcal{H}_2$-norm case we need the following lemmas.

Lemma 5.20. *Let* $\lambda, \widetilde{\lambda} \in \mathbb{R}$. *Then*

$$\int_{\imath\mathbb{R}} \left| \frac{1}{s/\lambda - 1} \right|^2 \, \mathrm{d}s = \pi \cdot |\lambda|, \tag{5.11}$$

$$\int_{\imath\mathbb{R}} \left| \frac{1}{s/\lambda - 1} \right|^2 \cdot \left| \frac{1}{s/\widetilde{\lambda} - 1} \right|^2 \, \mathrm{d}s = \pi \cdot \frac{|\lambda| \cdot |\widetilde{\lambda}|}{|\lambda| + |\widetilde{\lambda}|}. \tag{5.12}$$

Proof. Firstly, we prove equation (5.11) by

$$\begin{aligned}
\int_{\imath\mathbb{R}} \left| \frac{1}{s/\lambda - 1} \right|^2 \, \mathrm{d}s &= \int_{\mathbb{R}} \frac{1}{(\omega/|\lambda|)^2 + 1} \, \mathrm{d}\omega \\
&= \left[|\lambda| \cdot \arctan\left(\frac{\omega}{|\lambda|} \right) \right]_{-\infty}^{+\infty} \\
&= |\lambda| \cdot \left\{ \frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right\} \\
&= \pi \cdot |\lambda|.
\end{aligned}$$

Now we prove equation (5.12) for the case $|\lambda| = |\widetilde{\lambda}|$, i.e.,

$$\begin{aligned}
\int_{\imath\mathbb{R}} \left| \frac{1}{s/\lambda - 1} \right|^4 \, \mathrm{d}s &= \int_{\mathbb{R}} \left(\frac{1}{(\omega/|\lambda|)^2 + 1} \right)^2 \, \mathrm{d}\omega \\
&\overset{(A.1)}{=} \left[|\lambda| \cdot \frac{1}{2} \cdot \left(\frac{\omega/|\lambda|}{(\omega/|\lambda|)^2 + 1} + \arctan\left(\frac{\omega}{|\lambda|} \right) \right) \right]_{-\infty}^{+\infty} \\
&= \frac{|\lambda|}{2} \cdot \left(\left\{ 0 + \frac{\pi}{2} \right\} - \left\{ 0 + \left(-\frac{\pi}{2} \right) \right\} \right) \\
&= \pi \cdot \frac{|\lambda|}{2} \\
&= \pi \cdot \frac{|\lambda| \cdot |\lambda|}{|\lambda| + |\lambda|}.
\end{aligned}$$

Finally, we prove equation (5.12) for $|\lambda| \neq |\widetilde{\lambda}|$ by

$$\begin{aligned}
&\int_{\imath\mathbb{R}} \left| \frac{1}{s/\lambda - 1} \right|^2 \cdot \left| \frac{1}{s/\widetilde{\lambda} - 1} \right|^2 \, \mathrm{d}s \\
&= \int_{\mathbb{R}} \frac{1}{(\omega/|\lambda|)^2 + 1} \cdot \frac{1}{(\omega/|\widetilde{\lambda}|)^2 + 1} \, \mathrm{d}\omega \\
&= \frac{1}{|\widetilde{\lambda}|^2 - |\lambda|^2} \int_{\mathbb{R}} \left(\frac{|\widetilde{\lambda}|^2}{(\omega/|\lambda|)^2 + 1} - \frac{|\lambda|^2}{(\omega/|\widetilde{\lambda}|)^2 + 1} \right) \, \mathrm{d}\omega \\
&= \frac{1}{|\widetilde{\lambda}|^2 - |\lambda|^2} \cdot \left[|\widetilde{\lambda}|^2 \cdot |\lambda| \cdot \arctan\left(\frac{\omega}{|\lambda|} \right) - |\lambda|^2 \cdot |\widetilde{\lambda}| \cdot \arctan\left(\frac{\omega}{|\widetilde{\lambda}|} \right) \right]_{-\infty}^{+\infty} \\
&= \frac{|\lambda| \cdot |\widetilde{\lambda}|}{|\widetilde{\lambda}|^2 - |\lambda|^2} \cdot \left[|\widetilde{\lambda}| \cdot \arctan\left(\frac{\omega}{|\lambda|} \right) - |\lambda| \cdot \arctan\left(\frac{\omega}{|\widetilde{\lambda}|} \right) \right]_{-\infty}^{+\infty}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{|\lambda| \cdot |\widetilde{\lambda}|}{|\widetilde{\lambda}|^2 - |\lambda|^2} \cdot \left(\left\{ |\widetilde{\lambda}| \cdot \frac{\pi}{2} - |\lambda| \cdot \frac{\pi}{2} \right\} - \left\{ |\widetilde{\lambda}| \cdot \left(-\frac{\pi}{2} \right) - |\lambda| \cdot \left(-\frac{\pi}{2} \right) \right\} \right) \\
&= \pi \cdot \frac{|\lambda| \cdot |\widetilde{\lambda}|}{|\widetilde{\lambda}|^2 - |\lambda|^2} \cdot \left(|\widetilde{\lambda}| - |\lambda| \right) \\
&= \pi \cdot \frac{|\lambda| \cdot |\widetilde{\lambda}|}{\left(|\widetilde{\lambda}| + |\lambda| \right) \cdot \left(|\widetilde{\lambda}| - |\lambda| \right)} \cdot \left(|\widetilde{\lambda}| - |\lambda| \right) \\
&= \pi \cdot \frac{|\lambda| \cdot |\widetilde{\lambda}|}{|\widetilde{\lambda}| + |\lambda|}. \qquad \square
\end{aligned}
$$

Lemma 5.21 (cf. Lemma 5.17)**.** *Let* $\mathbf{Z} \in \mathbb{R}^{r \times m}$ *be an arbitrary real matrix. Then*

$$
\int_{\imath\mathbb{R}} \|(s\mathbf{A}^{-1} - \mathbf{I})^{-1} \cdot \mathbf{Z}\|_F^2 \, \mathrm{d}s = \pi \cdot \|(-\mathbf{A})^{1/2} \cdot \mathbf{Z}\|_F^2. \tag{5.13}
$$

$$
\int_{\imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{Z}\|_F^2 \, \mathrm{d}s = \pi \cdot \|(-\mathbf{A})^{-1/2} \cdot \mathbf{Z}\|_F^2. \tag{5.14}
$$

Proof. We start proving equation (5.13).

First, we show the claim for the diagonal matrix $\mathbf{D} = \mathrm{diag}(\lambda_1, \dots, \lambda_r)$ instead of $\mathbf{A}$. Recall that $\lambda_1, \dots, \lambda_r < 0$. Let $\mathbf{Z} = [z_{ij}]_{i,j=1}^{r,m}$. We have

$$
\begin{aligned}
\int_{\imath\mathbb{R}} \|\left(s\mathbf{D}^{-1} - \mathbf{I}\right)^{-1} \cdot \mathbf{Z}\|_F^2 \, \mathrm{d}s &= \int_{\imath\mathbb{R}} \left\| \mathrm{diag}\left(\frac{1}{s/\lambda_1 - 1}, \dots, \frac{1}{s/\lambda_r - 1} \right) \cdot \mathbf{Z} \right\|_F^2 \mathrm{d}s \\
&= \int_{\imath\mathbb{R}} \left\{ \sum_{i=1}^{r} \sum_{j=1}^{m} \left(\left| \frac{1}{s/\lambda_i - 1} \right| \cdot z_{ij} \right)^2 \right\} \mathrm{d}s \\
&= \sum_{i=1}^{r} \sum_{j=1}^{m} \left\{ z_{ij}^2 \cdot \int_{\imath\mathbb{R}} \left| \frac{1}{s/\lambda_i - 1} \right|^2 \mathrm{d}s \right\} \\
&= \sum_{i=1}^{r} \sum_{j=1}^{m} \left\{ z_{ij}^2 \cdot \pi \cdot |\lambda_i| \right\} \\
&= \pi \cdot \sum_{i=1}^{r} \sum_{j=1}^{m} z_{ij}^2 \cdot |\lambda_i| \qquad (5.15) \\
&= \pi \cdot \left\| \mathrm{diag}\left(\sqrt{|\lambda_1|} \dots, \sqrt{|\lambda_r|} \right) \cdot \mathbf{Z} \right\|_F^2 \\
&= \pi \cdot \left\| \mathrm{diag}\left(\sqrt{-\lambda_1}, \dots, \sqrt{-\lambda_r} \right) \cdot \mathbf{Z} \right\|_F^2 \\
&= \pi \cdot \|(-\mathbf{D})^{1/2} \cdot \mathbf{Z}\|_F^2 \qquad (5.16)
\end{aligned}
$$

using equation (5.11) from Lemma 5.20 for evaluating the integral.

Now we show the claim for an arbitrary symmetric negative definite matrix $\mathbf{A}$. The matrix $\mathbf{A}$ is symmetric (and negative definite) and hence orthogonally diagonalizable

$$
\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \mathrm{diag}(\lambda_1, \dots, \lambda_r) =: \mathbf{D}
$$

with eigenvalues $\lambda_i \in \mathbb{R}^-$. Then for $s \in \imath\mathbb{R}$

$$\mathbf{Q}^T \left(s\mathbf{A}^{-1} - \mathbf{I}\right)^{-1} \mathbf{Q} = (s\mathbf{D}^{-1} - \mathbf{I})^{-1}$$

and

$$\mathbf{Q}^T (-\mathbf{A})^{1/2} \mathbf{Q} = (-\mathbf{D})^{1/2}.$$

With $\mathbf{Q}^T\mathbf{Z}$ instead of $\mathbf{Z}$ it follows

$$\begin{aligned}
\int_{\imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{Z}\|_F^2 \, \mathrm{d}s &= \int_{\imath\mathbb{R}} \left\|\mathbf{Q} \cdot (s\mathbf{I} - \mathbf{D})^{-1} \cdot \mathbf{Q}^T \cdot \mathbf{Z}\right\|_F^2 \, \mathrm{d}s \\
&= \int_{\imath\mathbb{R}} \left\|(s\mathbf{I} - \mathbf{D})^{-1} \cdot (\mathbf{Q}^T \cdot \mathbf{Z})\right\|_F^2 \, \mathrm{d}s \\
&\overset{(5.16)}{=} \pi \cdot \left\|(-\mathbf{D})^{1/2} \cdot \mathbf{Q}^T \cdot \mathbf{Z}\right\|_F^2 \\
&= \pi \cdot \left\|\mathbf{Q} \cdot (-\mathbf{D})^{1/2} \cdot (\mathbf{Q}^T \cdot \mathbf{Z})\right\|_F^2 \\
&= \pi \cdot \|(-\mathbf{A})^{1/2} \cdot \mathbf{Z}\|_F^2.
\end{aligned}$$

Equation (5.14) follows from equation (5.13) with $\mathbf{A}^{-1} \cdot \mathbf{Z}$ instead of $\mathbf{Z}$. □

6 Error Convergence Rates for Interpolation on Sparse Grids

We develop two convergence rate statements for the interpolation error (in the $\mathcal{L}_\infty(\mathcal{P}, \mathcal{H}_\infty)$-norm) in terms of the number of sparse grid interpolation points in parameter space, one for piece-wise linear (see section 6.1) and one for global polynomial (see section 6.2) basis functions.

Both sections are structured as follows.

In the first subsection we repeat the result from sparse grid interpolation. In the second subsection we generalize this for the entry-wise interpolation of matrices. In the third subsection we apply this to the system matrices and plug it into the bound from chapter 5.

To relate a matrix-valued function to its scalar-valued entry functions in the third subsections, we need the following result for a constant matrix and its entries.

Lemma 6.1. *Let* $\mathbf{M} \in \mathbb{R}^{n\times m}$*. For the matrix norm* $\|\cdot\| = \|\cdot\|_F$ *or* $\|\cdot\| = \|\cdot\|_2$ *it holds*

$$\max_{i,j} |M_{ij}|^2 \leq \|\mathbf{M}\|^2 \leq \sum_{i=1}^{n}\sum_{j=1}^{m} |M_{ij}|^2.$$

Proof. For the Frobenius norm

$$\|\mathbf{M}\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} |M_{ij}|^2 \geq \max_{i,j} |M_{ij}|^2$$

and for the spectral norm

$$\max_{i,j} |M_{ij}|^2 \leq \|\mathbf{M}\|_2^2 \leq \|\mathbf{M}\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} |M_{ij}|^2,$$

where the left inequality is from [GVL89, (2.3.8)]. □

6.1 Interpolation with Piece-wise Linear Basis Functions

We consider functions from the Sobolev space with bounded mixed second derivatives (see subsection 2.1.3) on $\mathcal{P}$ and interpolate on sparse grids with piece-wise linear basis functions (see subsection 2.3.1).

6.1.1 Interpolation of Scalar-valued Functions

From Lemma 2.20 we have for $f \in \mathcal{W}_p^2(\mathcal{P}, \mathbb{R})$ with $p \in \{2, \infty\}$ and $q \in \{2, \infty\}$

$$\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q} \leq c \cdot N^{-2} \cdot (\log N)^{3\cdot(D-1)} \cdot |||f|||_{\mathbf{2},q},$$

where N is the number of grid points and the constant c depends on the dimension D.

6.1.2 Entry-wise Interpolation of Matrix-valued Functions

We deduce from the scalar case. Therefore we relate the matrix-valued function $\mathbf{F}$ to its scalar-valued entry functions f_{ij} to obtain three results similar to the constant matrix case in Lemma 6.1. The first lemma is for the seminorms $|\cdot|_{\boldsymbol{\alpha},p}$. The second lemma is for the norms $\|\cdot\|_{\mathcal{W}_p^R}$. The third lemma is for $|||\cdot|||_{\boldsymbol{\alpha},p}$.

Lemma 6.2. *Let $D^{\boldsymbol{\alpha}} f_{ij} \in \mathcal{L}_p(\mathcal{P}, \mathbb{R})$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$ and $p \in \{2, \infty\}$. Let $\mathbf{F} = [f_{ij}]_{i,j=1}^{n,m}$. Then $D^{\boldsymbol{\alpha}}\mathbf{F} \in \mathcal{L}_p(\mathcal{P}, \mathbb{R}^{n\times m})$ and it holds*

$$\max_{i,j} |f_{ij}|^2_{\boldsymbol{\alpha},p} \leq |\mathbf{F}|^2_{\boldsymbol{\alpha},p} \leq \sum_{i=1}^{n}\sum_{j=1}^{m} |f_{ij}|^2_{\boldsymbol{\alpha},p}.$$

Proof. Firstly, we consider $p = \infty$. Applying the first inequality of Lemma 6.1, we have

$$\begin{aligned}
|\mathbf{F}|^2_{\boldsymbol{\alpha},\infty} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|D^{\boldsymbol{\alpha}}\mathbf{F}(\mathbf{p})\|_2^2 \\
&\geq \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \max_{i,j} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \\
&= \max_{i,j} \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \\
&= \max_{i,j} |f_{ij}|^2_{\boldsymbol{\alpha},\infty}
\end{aligned}$$

and applying the second inequality of Lemma 6.1, we have

$$\begin{aligned}
|\mathbf{F}|^2_{\boldsymbol{\alpha},\infty} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|D^{\boldsymbol{\alpha}}\mathbf{F}(\mathbf{p})\|_2^2 \\
&\leq \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \sum_{i=1}^{n}\sum_{j=1}^{m} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \\
&\leq \sum_{i=1}^{n}\sum_{j=1}^{m} \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} |f_{ij}|^2_{\boldsymbol{\alpha},\infty}.
\end{aligned}$$

Now we consider $p = 2$. Applying the first inequality of Lemma 6.1, we have

$$|\mathbf{F}|^2_{\boldsymbol{\alpha},2} = \int_{\mathcal{P}} \|D^{\boldsymbol{\alpha}}\mathbf{F}(\mathbf{p})\|_F^2 \, \mathrm{d}\mathbf{p}$$

$$
\begin{aligned}
&\geq \int_{\mathcal{P}} \max_{i,j} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \ \mathrm{d}\mathbf{p} \\
&\geq \max_{i,j} \int_{\mathcal{P}} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \ \mathrm{d}\mathbf{p} \\
&= \max_{i,j} |f_{ij}|^2_{\boldsymbol{\alpha},2}
\end{aligned}
$$

and applying the second inequality of Lemma 6.1, we have

$$
\begin{aligned}
|\mathbf{F}|^2_{\boldsymbol{\alpha},2} &= \int_{\mathcal{P}} \|D^{\boldsymbol{\alpha}} \mathbf{F}(\mathbf{p})\|_F^2 \ \mathrm{d}\mathbf{p} \\
&\leq \int_{\mathcal{P}} \sum_{i=1}^{n} \sum_{j=1}^{m} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \ \mathrm{d}\mathbf{p} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} \int_{\mathcal{P}} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \ \mathrm{d}\mathbf{p} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} |f_{ij}|^2_{\boldsymbol{\alpha},2}. \qquad \square
\end{aligned}
$$

Lemma 6.3. *Let* $f_{ij} \in \mathcal{W}_p^R(\mathcal{P}, \mathbb{R})$ *for* $i = 1, \ldots, n$ *and* $j = 1, \ldots, m$ *with* $R \geq 0$ *and* $p \in \{2, \infty\}$*. Let* $\mathbf{F} = [f_{ij}]_{i,j=1}^{n,m}$*. Then* $\mathbf{F} \in \mathcal{W}_p^R(\mathcal{P}, \mathbb{R}^{n \times m})$ *and it holds*

$$
\max_{i,j} \|f_{ij}\|^2_{\mathcal{W}_p^R} \leq \|\mathbf{F}\|^2_{\mathcal{W}_p^R} \leq \sum_{i=1}^{n} \sum_{j=1}^{m} \|f_{ij}\|^2_{\mathcal{W}_p^R}.
$$

Proof. Firstly, we consider $p = \infty$. Applying the first inequality of Lemma 6.2, we have

$$
\begin{aligned}
\|\mathbf{F}\|^2_{\mathcal{W}_\infty^R} &= \max_{|\boldsymbol{\alpha}|_\infty \leq R} |\mathbf{F}|^2_{\boldsymbol{\alpha},\infty} \\
&\geq \max_{|\boldsymbol{\alpha}|_\infty \leq R} \max_{i,j} |f_{ij}|^2_{\boldsymbol{\alpha},\infty} \\
&= \max_{i,j} \max_{|\boldsymbol{\alpha}|_\infty \leq R} |f_{ij}|^2_{\boldsymbol{\alpha},\infty} \\
&= \max_{i,j} \|f_{ij}\|^2_{\mathcal{W}_\infty^R}
\end{aligned}
$$

and applying the second inequality of Lemma 6.2, we have

$$
\begin{aligned}
\|\mathbf{F}\|^2_{\mathcal{W}_\infty^R} &= \max_{|\boldsymbol{\alpha}|_\infty \leq R} |\mathbf{F}|^2_{\boldsymbol{\alpha},\infty} \\
&\leq \max_{|\boldsymbol{\alpha}|_\infty \leq R} \sum_{i=1}^{n} \sum_{j=1}^{m} |f_{ij}|^2_{\boldsymbol{\alpha},\infty} \\
&\leq \sum_{i=1}^{n} \sum_{j=1}^{m} \max_{|\boldsymbol{\alpha}|_\infty \leq R} |f_{ij}|^2_{\boldsymbol{\alpha},\infty}
\end{aligned}
$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} \|f_{ij}\|_{\mathcal{W}_\infty^R}^2.$$

Now we consider $p = 2$. Applying the first inequality of Lemma 6.2, we have

$$\begin{aligned}\|\mathbf{F}\|_{\mathcal{W}_2^R}^2 &= \sum_{|\boldsymbol{\alpha}|_\infty \le R} |\mathbf{F}|_{\boldsymbol{\alpha},2}^2 \\ &\ge \sum_{|\boldsymbol{\alpha}|_\infty \le R} \max_{i,j} |f_{ij}|_{\boldsymbol{\alpha},2}^2 \\ &\ge \max_{i,j} \sum_{|\boldsymbol{\alpha}|_\infty \le R} |f_{ij}|_{\boldsymbol{\alpha},2}^2 \\ &= \max_{i,j} \|f_{ij}\|_{\mathcal{W}_2^R}^2\end{aligned}$$

and applying the second inequality of Lemma 6.2, we have

$$\begin{aligned}\|\mathbf{F}\|_{\mathcal{W}_2^R}^2 &= \sum_{|\boldsymbol{\alpha}|_\infty \le R} |\mathbf{F}|_{\boldsymbol{\alpha},2}^2 \\ &\le \sum_{|\boldsymbol{\alpha}|_\infty \le R} \sum_{i=1}^{n} \sum_{j=1}^{m} |f_{ij}|_{\boldsymbol{\alpha},2}^2 \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{|\boldsymbol{\alpha}|_\infty \le R} |f_{ij}|_{\boldsymbol{\alpha},2}^2 \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} \|f_{ij}\|_{\mathcal{W}_2^R}^2.\end{aligned}$$

□

Lemma 6.4. *Let* $f_{ij} \in \mathcal{W}_p^R(\mathcal{P}, \mathbb{R})$ *and* $|||f_{ij}|||_{\boldsymbol{\alpha},p} < \infty$ *for* $i = 1, \ldots, n$ *and* $j = 1, \ldots, m$ *with* $R \ge 0$ *and* $p \in \{2, \infty\}$ *and* $|\boldsymbol{\alpha}|_\infty \le R$. *Let* $\mathbf{F} = [f_{ij}]_{i,j=1}^{n,m}$. *Then* $\mathbf{F} \in \mathcal{W}_p^R(\mathcal{P}, \mathbb{R}^{n\times m})$ *and* $|||\mathbf{F}|||_{\boldsymbol{\alpha},p} < \infty$ *and it holds*

$$\max_{i,j} |||f_{ij}|||_{\boldsymbol{\alpha},p}^2 \le |||\mathbf{F}|||_{\boldsymbol{\alpha},p}^2 \le \sum_{i=1}^{n} \sum_{j=1}^{m} |||f_{ij}|||_{\boldsymbol{\alpha},p}^2.$$

Proof. Applying the first inequality of Lemma 6.2, we have

$$\begin{aligned}|||\mathbf{F}|||_{\boldsymbol{\alpha},p}^2 &= \max_{\substack{\boldsymbol{\ell} \in \{-1,0,*\}^D \\ |\boldsymbol{\ell}^{(*)}| > 0}} \{|\mathbf{F}_{\boldsymbol{\ell},*}|_{\boldsymbol{\alpha},p}\} \\ &\ge \max_{\substack{\boldsymbol{\ell} \in \{-1,0,*\}^D \\ |\boldsymbol{\ell}^{(*)}| > 0}} \left\{\max_{i,j} |(f_{ij})_{\boldsymbol{\ell},*}|_{\boldsymbol{\alpha},p}\right\} \\ &= \max_{i,j} \max_{\substack{\boldsymbol{\ell} \in \{-1,0,*\}^D \\ |\boldsymbol{\ell}^{(*)}| > 0}} \{|(f_{ij})_{\boldsymbol{\ell},*}|_{\boldsymbol{\alpha},p}\} \\ &= \max_{i,j} |||f_{ij}|||_{\boldsymbol{\alpha},p}^2\end{aligned}$$

and applying the second inequality of Lemma 6.2, we have

$$\begin{aligned}
|||\mathbf{F}|||^2_{\boldsymbol{\alpha},p} &= \max_{\substack{\boldsymbol{\ell}\in\{-1,0,*\}^D\\ |\boldsymbol{\ell}^{(*)}|>0}} \{|\mathbf{F}_{\boldsymbol{\ell},*}|_{\boldsymbol{\alpha},p}\} \\
&\le \max_{\substack{\boldsymbol{\ell}\in\{-1,0,*\}^D\\ |\boldsymbol{\ell}^{(*)}|>0}} \left\{\sum_{i=1}^{n}\sum_{j=1}^{m} |(f_{ij})_{\boldsymbol{\ell},*}|^2_{\boldsymbol{\alpha},p}\right\} \\
&\le \sum_{i=1}^{n}\sum_{j=1}^{m} \max_{\substack{\boldsymbol{\ell}\in\{-1,0,*\}^D\\ |\boldsymbol{\ell}^{(*)}|>0}} \left\{|(f_{ij})_{\boldsymbol{\ell},*}|^2_{\boldsymbol{\alpha},p}\right\} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} |||f_{ij}|||^2_{\boldsymbol{\alpha},p}. \qquad\square
\end{aligned}$$

Lemma 6.5 (cf. Lemma 6.10)**.** *Let* $\mathbf{F}\in\mathcal{W}^2_p(\mathcal{P},\mathbb{R}^{n\times m})$ *and* $|||\mathbf{F}|||_{\mathbf{2},p}<\infty$ *with* $p\in\{2,\infty\}$*. Then it holds*

$$\|\mathbf{F}-\widetilde{\mathbf{F}}^{(N)}\|_{\mathcal{L}_p} \le c\cdot N^{-2}\cdot(\log N)^{3\cdot(D-1)}\cdot\sqrt{nm}\cdot|||\mathbf{F}|||_{\mathbf{2},p},$$

where N *is the number of grid points. The constant* c *depends on the dimension* D*.*

Proof. Since $\mathbf{F}\in\mathcal{W}^2_p([0,1]^D,\mathbb{R}^{n\times m})$ and $|||\mathbf{F}|||_{\mathbf{2},p}<\infty$ we have $f_{ij}\in\mathcal{W}^2_p([0,1]^D,\mathbb{R})$ and $|||f_{ij}|||_{\mathbf{2},p}<\infty$ for $i=1,\dots,n$ and $j=1,\dots,m$.

Applying the second inequality of Lemma 6.3 and then Lemma 2.20 with $p=q$ for scalar-valued functions we get

$$\begin{aligned}
\|\mathbf{F}-\widetilde{\mathbf{F}}^{(N)}\|^2_{\mathcal{L}_p} &= \|\mathbf{F}-\widetilde{\mathbf{F}}^{(N)}\|^2_{\mathcal{W}^0_p} \\
&\le \sum_{i=1}^{n}\sum_{j=1}^{m} \|f_{ij}-\widetilde{f}^{(N)}_{ij}\|^2_{\mathcal{W}^0_p} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \|f_{ij}-\widetilde{f}^{(N)}_{ij}\|^2_{\mathcal{L}_p} \\
&\le \sum_{i=1}^{n}\sum_{j=1}^{m} \left(c\cdot N^{-2}\cdot(\log N)^{3\cdot(d-1)}\right)^2\cdot|||f_{ij}|||^2_{\mathbf{2},p} \\
&= \left(c\cdot N^{-2}\cdot(\log N)^{3\cdot(d-1)}\right)^2\cdot\sum_{i=1}^{n}\sum_{j=1}^{m}|||f_{ij}|||^2_{\mathbf{2},p} \\
&\le \left(c\cdot N^{-2}\cdot(\log N)^{3\cdot(d-1)}\right)^2\cdot n\cdot m\cdot|||\mathbf{F}|||^2_{\mathbf{2},p}
\end{aligned}$$

with finally applying the first inequality of Lemma 6.4 $n\cdot m$ times. □

6.1.3 Transfer Function

We require the following of the matrices we interpolate.

Requirement 6.6 (smoothness)**.** *Let* $\mathfrak{A}_r \in \mathcal{W}^2_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$ *and* $|||\mathfrak{A}_r|||_{\mathbf{2},\infty} < \infty$, $\mathbf{B}_r \in \mathcal{W}^2_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$ *and* $|||\mathbf{B}_r|||_{\mathbf{2},\infty} < \infty$, *and* $\mathbf{C}_r \in \mathcal{W}^2_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$ *and* $|||\mathbf{C}_r|||_{\mathbf{2},\infty} < \infty$.

We get the following convergence rate for the interpolation error of those matrices.

Lemma 6.7 (cf. Lemma 6.12)**.** *Under Requirement 6.6 it holds*

$$\begin{aligned}
\|\mathfrak{A}_r - \widetilde{\mathfrak{A}}_r^{(N)}\|_{\mathcal{L}_\infty} &\le c\cdot N^{-2}\cdot(\log N)^{3\cdot(D-1)}\cdot r\cdot\|\mathfrak{A}_r\|_{\mathcal{W}^2_\infty},\\
\|\mathbf{B}_r - \widetilde{\mathbf{B}}_r^{(N)}\|_{\mathcal{L}_\infty} &\le c\cdot N^{-2}\cdot(\log N)^{3\cdot(D-1)}\cdot\sqrt{rm}\cdot\|\mathbf{B}_r\|_{\mathcal{W}^2_\infty},\\
\|\mathbf{C}_r - \widetilde{\mathbf{C}}_r^{(N)}\|_{\mathcal{L}_\infty} &\le c\cdot N^{-2}\cdot(\log N)^{3\cdot(D-1)}\cdot\sqrt{rq}\cdot\|\mathbf{C}_r\|_{\mathcal{W}^2_\infty},
\end{aligned}$$

where N *is the number of grid points. The constant* c *depends on the dimension* D.

Proof. Apply Lemma 6.5 for $p = \infty$. □

Now we prove a convergence rate for the interpolation error of the transfer function.

Theorem 6.8 (cf. Theorem 6.14)**.** *Under Requirement 6.6 it holds*

$$\begin{aligned}
\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r^{(N)}\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} \le c\cdot N^{-2}\cdot(\log N)^{3\cdot(D-1)}\cdot\|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}&\\
\cdot\Big\{4r\cdot\|\mathbf{B}_r\|_{\mathcal{L}_\infty}\cdot\|\mathbf{C}_r\|_{\mathcal{L}_\infty}\cdot\|\mathfrak{A}\|_{\mathcal{W}^2_\infty}&\\
+\sqrt{rm}\cdot\|\mathbf{C}_r\|_{\mathcal{L}_\infty}\cdot\|\mathbf{B}_r\|_{\mathcal{W}^2_\infty}&\\
+2e\cdot\sqrt{rq}\cdot\|\mathbf{B}_r\|_{\mathcal{L}_\infty}\cdot\|\mathbf{C}_r\|_{\mathcal{W}^2_\infty}\Big\}&
\end{aligned}$$

when the number of grid points N *is sufficiently large. The constant* c *depends on the dimension* D.

Proof. For sufficiently large N we have

$$\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty} \le 1, \tag{6.1}$$

so $e^{\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}} - 1 \le 2\cdot\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}$ and $e^{\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}} \le e$, and

$$\|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty} \le \|\mathbf{B}_r\|_{\mathcal{L}_\infty}. \tag{6.2}$$

With Theorem 5.13 we get

$$\begin{aligned}
&\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r^{(N)}\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)}\\
&\le \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}\cdot\Big\{(\|\mathbf{B}_r\|_{\mathcal{L}_\infty}+\|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty})\cdot\|\mathbf{C}_r\|_{\mathcal{L}_\infty}\cdot(e^{\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}}-1)\\
&\qquad+\|\mathbf{C}_r\|_{\mathcal{L}_\infty}\cdot\|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty}\\
&\qquad+e^{\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}}\cdot(\|\mathbf{B}_r\|_{\mathcal{L}_\infty}+\|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty})\cdot\|\Delta\mathbf{C}_r^{(N)}\|_{\mathcal{L}_\infty}\Big\}\\
&\overset{(6.3)}{\le}\|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty}\cdot\Big\{2\cdot(\|\mathbf{B}_r\|_{\mathcal{L}_\infty}+\|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty})\cdot\|\mathbf{C}_r\|_{\mathcal{L}_\infty}\cdot\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}
\end{aligned}$$

$$
\begin{aligned}
&\qquad + \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta \mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty} \\
&\qquad + e \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta \mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty}) \cdot \|\Delta \mathbf{C}_r^{(N)}\|_{\mathcal{L}_\infty} \Big\} \\
\overset{(6.4)}{\leq}\ & \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \cdot \Big\{ 2 \cdot 2 \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta \mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty} \\
&\qquad + \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta \mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty} \\
&\qquad + 2 \cdot e \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta \mathbf{C}_r^{(N)}\|_{\mathcal{L}_\infty} \Big\} \\
\leq\ & \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \cdot c \cdot N^{-2} \cdot (\log N)^{3 \cdot (D-1)} \cdot \Big\{ 4 \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot r \cdot \|\mathfrak{A}_r\|_{\mathcal{W}^2_\infty} \\
&\qquad + \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \sqrt{rm} \cdot \|\mathbf{B}_r\|_{\mathcal{W}^2_\infty} \\
&\qquad + 2e \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \sqrt{rq} \cdot \|\mathbf{C}_r\|_{\mathcal{W}^2_\infty} \Big\}
\end{aligned}
$$

applying Lemma 6.7. □

6.2 Interpolation with Global Polynomial Basis Functions

We consider functions which are continuously differentiable up to some order R (see subsection 2.1.4) on $\mathcal{P} = [-1, 1]^D$ (as in subsection 2.3.2) and interpolate on Chebyshev sparse grids with global polynomial basis functions (see subsection 2.3.2).

6.2.1 Interpolation of Scalar-valued Functions

From Lemma 2.21 we have for $f \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R})$

$$\|f - \widetilde{f}^{(N)}\|_{\mathcal{C}^0_\infty} \leq c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot \|f\|_{\mathcal{C}^R_\infty},$$

where N is the number of grid points and the constant c depends on the dimension D of the parameter space and the smoothness R.

6.2.2 Entry-wise Interpolation of Matrix-valued Functions

We want to deduce from the scalar case. Therefore we relate the matrix-valued function $\mathbf{F}$ to its scalar-valued entry functions f_{ij} to obtain a result similar to the constant matrix case in Lemma 6.1.

Lemma 6.9 (cf. Lemma 6.3)**.** *Let $f_{ij} \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R})$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$ with $R \geq 0$. Let $\mathbf{F} = [f_{ij}]_{i,j=1}^{n,m}$. Then $\mathbf{F} \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R}^{n \times m})$ and it holds*

$$\max_{i,j} \|f_{ij}\|^2_{\mathcal{C}^R_\infty} \leq \|\mathbf{F}\|^2_{\mathcal{C}^R_\infty} \leq \sum_{i=1}^n \sum_{j=1}^m \|f_{ij}\|^2_{\mathcal{C}^R_\infty}.$$

Proof. Applying the first inequality of Lemma 6.1, we have

$$\|\mathbf{F}\|^2_{\mathcal{C}^R_\infty} = \max_{|\boldsymbol{\alpha}|_\infty \leq R} \sup_{\mathbf{p} \in \mathcal{P}} \|D^{\boldsymbol{\alpha}} \mathbf{F}(\mathbf{p})\|_2^2$$

$$
\begin{aligned}
&\geq \max_{|\boldsymbol{\alpha}|_\infty \leq R} \sup_{\mathbf{p}\in\mathcal{P}} \max_{i,j} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \\
&= \max_{i,j} \max_{|\boldsymbol{\alpha}|_\infty \leq R} \sup_{\mathbf{p}\in\mathcal{P}} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \\
&= \max_{i,j} \|f_{ij}\|^2_{\mathcal{C}^R_\infty}
\end{aligned}
$$

and applying the second inequality of Lemma 6.1, we have

$$
\begin{aligned}
\|\mathbf{F}\|^2_{\mathcal{C}^R_\infty} &= \max_{|\boldsymbol{\alpha}|_\infty \leq R} \sup_{\mathbf{p}\in\mathcal{P}} \|D^{\boldsymbol{\alpha}} \mathbf{F}(\mathbf{p})\|_2^2 \\
&\leq \max_{|\boldsymbol{\alpha}|_\infty \leq R} \sup_{\mathbf{p}\in\mathcal{P}} \sum_{i=1}^{n}\sum_{j=1}^{m} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \\
&\leq \sum_{i=1}^{n}\sum_{j=1}^{m} \max_{|\boldsymbol{\alpha}|_\infty \leq R} \sup_{\mathbf{p}\in\mathcal{P}} |D^{\boldsymbol{\alpha}} f_{ij}(\mathbf{p})|^2 \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \|f_{ij}\|^2_{\mathcal{C}^R_\infty}. \qquad \square
\end{aligned}
$$

Lemma 6.10 (cf. Lemma 6.5)**.** *Let* $\mathbf{F} \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R}^{n\times m})$*. Then it holds*

$$
\|\mathbf{F} - \widetilde{\mathbf{F}}^{(N)}\|_{\mathcal{C}^0_\infty} \leq c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot \sqrt{nm} \cdot \|\mathbf{F}\|_{\mathcal{C}^R_\infty}
$$

where N *is the number of grid points. The constant* c *depends on the dimension* D *of the parameter space and the smoothness* R*.*

Proof. Since $\mathbf{F} \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R}^{n\times m})$ we have $f_{ij} \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R})$ for $i = 1, \dots, n$, $j = 1, \dots, m$.

Applying the second inequality of Lemma 6.9 for $R = 0$ and then Lemma 2.21 for scalar-valued functions we get

$$
\begin{aligned}
\|\mathbf{F} - \widetilde{\mathbf{F}}^{(N)}\|^2_{\mathcal{C}^0_\infty} &\leq \sum_{i=1}^{n}\sum_{j=1}^{m} \|f_{ij} - \widetilde{f}^{(N)}_{ij}\|^2_{\mathcal{C}^0_\infty} \\
&\leq \sum_{i=1}^{n}\sum_{j=1}^{m} \left(c \cdot N^{-R} \cdot (\log N)^{(R+2)(d-1)+1}\right)^2 \cdot \|f_{ij}\|^2_{\mathcal{C}^R_\infty} \\
&= \left(c \cdot N^{-R} \cdot (\log N)^{(R+2)(d-1)+1}\right)^2 \cdot \sum_{i=1}^{n}\sum_{j=1}^{m} \|f_{ij}\|^2_{\mathcal{C}^R_\infty} \\
&\leq \left(c \cdot N^{-R} \cdot (\log N)^{(R+2)(d-1)+1}\right)^2 \cdot n \cdot m \cdot \|\mathbf{F}\|^2_{\mathcal{C}^R_\infty}
\end{aligned}
$$

with finally applying the first inequality of Lemma 6.9 $n \cdot m$ times. □

6.2.3 Transfer Function

We require the following of the matrices we interpolate.

Requirement 6.11 (smoothness)**.** *Let* $\mathfrak{A}_r \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$*,* $\mathbf{B}_r \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$ *and* $\mathbf{C}_r \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$*.*

We get the following convergence rate for the interpolation error of those matrices.

Lemma 6.12 (cf. Lemma 6.7)**.** *Under Requirement 6.11 it holds*

$$\begin{aligned}
\|\mathfrak{A}_r - \widetilde{\mathfrak{A}}_r^{(N)}\|_{\mathcal{C}^0_\infty} &\le c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot r \cdot \|\mathfrak{A}_r\|_{\mathcal{C}^R_\infty}, \\
\|\mathbf{B}_r - \widetilde{\mathbf{B}}_r^{(N)}\|_{\mathcal{C}^0_\infty} &\le c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot \sqrt{rm} \cdot \|\mathbf{B}_r\|_{\mathcal{C}^R_\infty}, \\
\|\mathbf{C}_r - \widetilde{\mathbf{C}}_r^{(N)}\|_{\mathcal{C}^0_\infty} &\le c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot \sqrt{rq} \cdot \|\mathbf{C}_r\|_{\mathcal{C}^R_\infty},
\end{aligned}$$

where N *is the number of grid points. The constant* c *depends on the dimension* D *of the parameter space and the smoothness* R*.*

Proof. Apply Lemma 6.10. □

For continuous functions the $\mathcal{C}^0_\infty$-norm and the $\mathcal{L}_\infty$-norm coincide.

Corollary 6.13. *Under Requirement 6.11 it holds*

$$\begin{aligned}
\|\mathfrak{A}_r - \widetilde{\mathfrak{A}}_r^{(N)}\|_{\mathcal{L}_\infty} &\le c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot r \cdot \|\mathfrak{A}_r\|_{\mathcal{C}^R_\infty}, \\
\|\mathbf{B}_r - \widetilde{\mathbf{B}}_r^{(N)}\|_{\mathcal{L}_\infty} &\le c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot \sqrt{rm} \cdot \|\mathbf{B}_r\|_{\mathcal{C}^R_\infty}, \\
\|\mathbf{C}_r - \widetilde{\mathbf{C}}_r^{(N)}\|_{\mathcal{L}_\infty} &\le c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot \sqrt{rq} \cdot \|\mathbf{C}_r\|_{\mathcal{C}^R_\infty},
\end{aligned}$$

where N *is the number of grid points. The constant* c *depends on the dimension* D *of the parameter space and the smoothness* R*.*

Now we prove a convergence rate for the interpolation error of the transfer function.

Theorem 6.14 (cf. Theorem 6.8)**.** *Under Requirement 6.11 it holds*

$$\begin{aligned}
\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r^{(N)}\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} \le c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \\
\cdot \Big\{ 4r \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\mathfrak{A}_r\|_{\mathcal{C}^R_\infty} \\
+ \sqrt{rm} \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{B}_r\|_{\mathcal{C}^R_\infty} \\
+ 2e \cdot \sqrt{rq} \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{C}_r\|_{\mathcal{C}^R_\infty} \Big\}
\end{aligned}$$

when the number of grid points N *is sufficiently large. The constant* c *depends on the dimension* D *of the parameter space and the smoothness* R*.*

Proof. For sufficiently large N we have

$$\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty} \le 1, \tag{6.3}$$

so $e^{\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}} - 1 \le 2 \cdot \|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}$ and $e^{\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}} \le e$, and

$$\|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty} \le \|\mathbf{B}_r\|_{\mathcal{L}_\infty}. \tag{6.4}$$

With Theorem 5.13 we get

$$
\begin{aligned}
&\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r^{(N)}\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} \\
&\leq \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \cdot \Big\{ (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty}) \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot (e^{\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}} - 1) \\
&\qquad + \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty} \\
&\qquad + e^{\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty}} \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty}) \cdot \|\Delta\mathbf{C}_r^{(N)}\|_{\mathcal{L}_\infty} \Big\} \\
&\overset{(6.3)}{\leq} \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \cdot \Big\{ 2 \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty}) \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty} \\
&\qquad + \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty} \\
&\qquad + e \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty}) \cdot \|\Delta\mathbf{C}_r^{(N)}\|_{\mathcal{L}_\infty} \Big\} \\
&\overset{(6.4)}{\leq} \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \cdot \Big\{ 2 \cdot 2 \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty} \\
&\qquad + \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty} \\
&\qquad + 2 \cdot e \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta\mathbf{C}_r^{(N)}\|_{\mathcal{L}_\infty} \Big\} \\
&\leq \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \cdot c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot \Big\{ 4 \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot r \cdot \|\mathfrak{A}_r\|_{\mathcal{C}_\infty^R} \\
&\qquad + \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \sqrt{rm} \cdot \|\mathbf{B}_r\|_{\mathcal{C}_\infty^R} \\
&\qquad + 2e \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \sqrt{rq} \|\mathbf{C}_r\|_{\mathcal{C}_\infty^R} \Big\} .
\end{aligned}
$$

applying Corollary 6.13. □

7 Computational Experiments

In this section we perform computational experiments to analyze our method and theoretical results.

We consider error per number of interpolation points. These are the two quantities that are important to us, since the number of interpolation points directly influences the memory consumption and computation time of the local reduced order models and the interpolation. In subsection 7.2.6 we also plot the error over the parameter space, which is one-dimensional there. We do not give times, except in subsection 7.2.6, since the implementation is not time-optimal and also involves e.g., expensive error computations.

7.1 Settings

7.1.1 Benchmarks

There are two major benchmark collections for model order reduction. The first one is the *Oberwolfach Model Reduction Benchmark Collection* [Obe]. The second one is part of the *Model Order Reduction Wiki* [MC]. Besides benchmarks the wiki contains pages explaining model order reduction methods and lists model order reduction software.

We choose two benchmarks which have symmetric system matrices and depend on more than one parameter.

Thermal Model

The benchmark is from [Obe] and available at `http://portal.uni-freiburg.de/imteksimulation/downloads/benchmark/ThermalModel(38865)` under the name "Boundary Condition Independent Thermal Model (38865)".

The mathematical model is a heat transfer equation with convection (Robin) boundary conditions. The film coefficients h_{top}, h_{side}, h_{bottom} are the parameters that describe the heat exchange at the interfaces.

The parameter dependence is affine in

$$\mathbf{A}(\mathbf{p}) = \mathbf{A} - p_1 \cdot \mathbf{A}_1 - p_2 \cdot \mathbf{A}_2 - p_3 \cdot \mathbf{A}_3$$

with $p_1 = h_{\text{top}}, p_2 = h_{\text{side}}, p_3 = h_{\text{bottom}} \in \mathcal{P} = [10^0, 10^4] \times [10^0, 10^4] \times [10^3, 10^4]$, so $D = 3$. The matrices $\mathbf{E}$, $\mathbf{B}$ and $\mathbf{C}$ are not parameter-dependent. The state dimension of the full order model is $n = 4257$. The input dimension is $m = 1$ and the output dimension is $q = 7$. In the default setting we make the model SISO by summing up the columns in the matrix $\mathbf{C}$.

This is the default benchmark. When we use this benchmark with less than 3 parameters, we set the remaining parameters to 10^4. The default dimension of the parameter space is 2.

Silicon Nitride Membrane

The benchmark is from [MC] and available at `http://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Silicon_nitride_membrane`.

The mathematical model is a heat transfer equation with a homogeneous heat generation rate, Dirichlet boundary condition at the bottom and convection (Robin) boundary condition at the top.

The parameter dependence is affine in

$$\mathbf{A}(\mathbf{p}) = \mathbf{A} + p_1 \cdot \mathbf{A}_1 + p_2 \cdot \mathbf{A}_2 \qquad \text{and}$$
$$\mathbf{E}(\mathbf{p}) = \mathbf{E} + p_3 \cdot p_4 \cdot \mathbf{E}_1$$

with

$$\begin{aligned} p_1 &= \kappa && \in [2,5] \text{ thermal conductivity,} \\ p_2 &= h && \in [10,12] \text{ heat transfer coefficient,} \\ p_3 &= c_p && \in [400,750] \text{ specific heat capacity} \qquad \text{and} \\ p_4 &= \rho && \in [3000,3200] \text{ density.} \end{aligned}$$

So we have $D = 4$. The state dimension of the full order model is $n = 60{,}020$. The input dimension is $m = 1$ and the output dimension is $q = 2$.

7.1.2 Model Order Reduction

We choose $r = 14$ as the default reduction order. At the center of the parameter space we run the Iterative Rational Krylov Algorithm (see Algorithm 2.1) with a symmetric initialization strategy (see Algorithm 7.1). We perform one-sided projection such that the square system matrices stay symmetric.

Algorithm 7.1 Symmetric initialization of Iterative Rational Krylov Algorithm.

Compute r largest magnitude eigenvalues λ_i and associated right eigenvectors $\mathbf{x}_i$ for which $\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{E}\mathbf{x}_i$. Then:

a) Set $\mathbf{V} = [\mathbf{x}_1, \ldots, \mathbf{x}_r]$ and $\mathbf{W} = \mathbf{V}$.

b) Set $\mathbf{A}_r = \mathbf{W}^T\mathbf{A}\mathbf{V}$, $\mathbf{E}_r = \mathbf{W}^T\mathbf{E}\mathbf{V}$, $\mathbf{B}_r = \mathbf{W}^T\mathbf{B}$ and $\mathbf{C}_r = \mathbf{C}\mathbf{V}$.

c) Compute eigenvalues λ_i and associated right and left eigenvectors $\mathbf{x}_i$ and $\mathbf{y}_i^H$ for which $\mathbf{A}_r\mathbf{x}_i = \lambda_i\mathbf{E}_r\mathbf{x}_i$, $\mathbf{y}_i^H\mathbf{A}_r = \lambda_i\mathbf{y}_i^H\mathbf{E}_r$ and $\mathbf{y}_i^H\mathbf{E}_r\mathbf{x}_j = \delta_{ij}$.

d) Set $\mu_i = -\lambda_i$, $\mathbf{b}_i^T = \mathbf{y}_i^H\mathbf{B}_r$ and $\mathbf{c}_i = \mathbf{C}_r\mathbf{x}_i$ for $i = 1, \ldots, r$.

e) Set $\mu_i \leftarrow |\mathrm{Re}(\mu_i)| + \imath\,\mathrm{Im}(\mu_i)$.

7.1.3 Discrete Error Measures

The error measures from section 3.3 are continuous. To compute an approximation of them, we discretize the frequency space and the parameter space.

For the frequency space we choose the finite interval $[10^{-3}, 10^{6}]$ and measure at 100 logarithmically spaced points. We call this grid $\mathcal{R}_{\text{test}}$.

So we can approximate the $\mathcal{H}_\infty$-norm by

$$\begin{aligned}\|\mathbf{H}(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty} &= \sup_{s\in\imath\mathbb{R}} \|\mathbf{H}(s;\mathbf{p})\|_2 \\ &\approx \max_{\omega\in\mathcal{R}_{\text{test}}} \|\mathbf{H}(\imath\omega;\mathbf{p})\|_2.\end{aligned}$$

We discretize the parameter space $\mathcal{P}$ with a grid $\mathcal{P}_{\text{test}}$, which is chosen as a fine full grid with base 3.

We compute an approximation of the $\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)$-norm and not the $\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)$-norm, as it is easier to compute. For the overall error we have the absolute norm

$$\begin{aligned}\|\mathbf{H}-\widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{H}(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty} \\ &\approx \max_{\mathbf{p}\in\mathcal{P}_{\text{test}}} \|\mathbf{H}(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty}\end{aligned}$$

and the relative norm

$$\begin{aligned}\|\mathbf{H}-\widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\text{rel}\mathcal{H}_\infty)} &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{H}(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\text{rel}\mathcal{H}_\infty} \\ &= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \frac{\|\mathbf{H}(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty}}{\|\mathbf{H}(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty}} \\ &\approx \max_{\mathbf{p}\in\mathcal{P}_{\text{test}}} \frac{\|\mathbf{H}(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty}}{\|\mathbf{H}(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty}}.\end{aligned}$$

The reduction errors $\|\mathbf{H}-\mathbf{H}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)}$ and $\|\mathbf{H}-\mathbf{H}_r\|_{\mathcal{L}_\infty(\mathcal{P},\text{rel}\mathcal{H}_\infty)}$ and the interpolation errors $\|\mathbf{H}_r-\widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)}$ and $\|\mathbf{H}_r-\widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\text{rel}\mathcal{H}_\infty)}$ are approximated analogously.

7.1.4 Sparse Grid Interpolation Toolbox

Klimke developed a sparse grid interpolation toolbox for MATLAB® [KW05, Kli07], which is available from `http://www.ians.uni-stuttgart.de/spinterp`. The toolbox is well documented and has several nice features we make use of. Besides “Piecewise linear basis functions” [Kli07, Section 1.4]) it includes also “Polynomial basis functions” [Kli07, Section 1.5]. Furthermore, it can deal with “Multiple output variables” [Kli07, Section 2.2] at once, what we need to interpolate all matrix entries simultaneously.

We use version v5.1.1 in all computations.

7.1.5 Types of Sparse Grids

We consider the following sparse grid types from the Sparse Grid Interpolation Toolbox (see subsection 7.1.4).

The sparse grids with boundary and piece-wise linear basis functions from subsection 2.3.1 are implemented as the sparse grid type “Maximum”. For the sparse grid with boundary type called “Clenshaw–Curtis”, the resolution on the boundary is not equal, but smaller than in the interior (see Figure 7.1 and Figure 7.2 for two-dimensional grids up to $n = 5$). This results in a smaller number of grid points (see Table 7.1).

The name of the sparse grid type “Maximum” is due to the $\mathcal{L}_\infty$-norm optimality of the grid. The name of the sparse grid type “Clenshaw–Curtis” stems from a grid that originally was made of Chebyshev-distributed and not equidistant nodes.

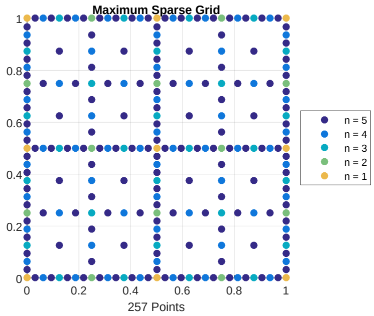


Figure 7.1: Sparse grid with boundary “Maximum” for $D = 2$.

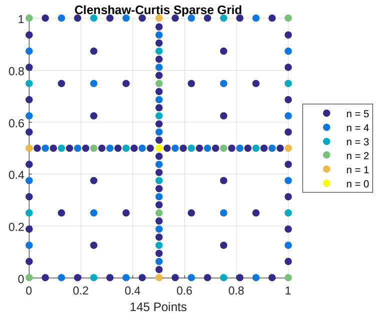


Figure 7.2: Sparse grid with boundary “Clenshaw–Curtis” for $D = 2$.

Table 7.1: Number of grid points for sparse grids with different boundary resolution.

[Kli07] $D = 2$	interior	boundary	overall
“Maximum”	1/5/17/49/129	8/16/32/ 64/128	9/21/49/113/257
“Clenshaw–Curtis”	1/5/13/33/ 81	4/ 8/16/ 32/ 64	5/13/29/ 65/145

The sparse grids with boundary and global polynomial basis functions from subsection 2.3.2 are implemented as the sparse grid type “Chebyshev”, where the number of interpolation points coincides with the sparse grid type “Clenshaw–Curtis”—just the location of the points is different (see Figure 7.5 and Figure 7.6).

We do not consider energy-norm optimal sparse grids since we do not interpolate solutions of partial differential equations and so this norm has no theoretical background for us. Furthermore, the parameter space dimensions of our benchmarks are moderate, so there is no need for the more complicated grid type so far.

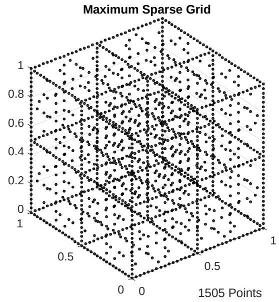


Figure 7.3: Sparse grid with boundary "Maximum" for $D = 3$.

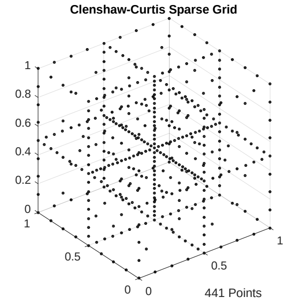


Figure 7.4: Sparse grid with boundary "Clenshaw–Curtis" for $D = 3$.

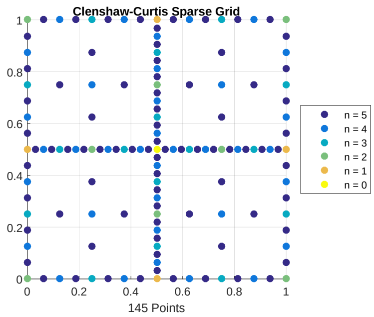


Figure 7.5: Sparse grid with boundary "Clenshaw–Curtis" for $D = 2$.

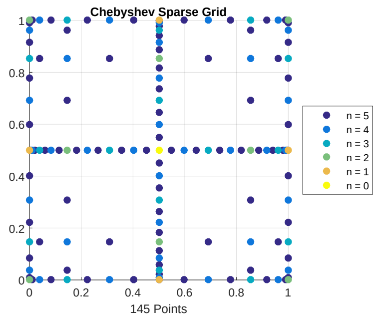


Figure 7.6: Sparse grid with boundary "Chebyshev" for $D = 2$.

7.2 Tests

7.2.1 Full Grids and Different Types of Sparse Grids

We compare interpolation on full grids and different types of sparse grids (see subsection 7.1.5) in terms of error per number of interpolation points (in parameter space).

The left plot in Figure 7.7 depicts the overall error (between the interpolated reduced order models and the parametric full order model). The black horizontal line represents the reduction error (between the directly reduced order models and the parametric full order model). As expected, the overall error levels off when it reaches the reduction error (for the Chebyshev sparse grid).

The right plot in Figure 7.7 depicts the interpolation error (between the interpolated reduced order models and the directly reduced order models). The curves are the same as for the overall error until the reduction error is reached (for the

Chebyshev sparse grid). So from now on we will only plot the overall error.

Interpolation on any sparse grid needs less interpolation points than interpolation on a full grid to achieve the same error. The Clenshaw–Curtis sparse grid is like the usual Maximum sparse grid, but with fewer boundary points. Interpolation on the Clenshaw–Curtis sparse grid is slightly more efficient for this benchmark. Interpolation on sparse grids with global polynomial basis functions (Chebyshev) is more efficient than with piece-wise linear basis functions (Maximum, Clenshaw–Curtis) if the parameter dependence is smooth.

To sum up, in this two-dimensional benchmark the number of interpolation points can be decreased from about 16,000 (full grid) down to about 700 (sparse grid) or even 150 (sparse grid with global polynomial basis functions).

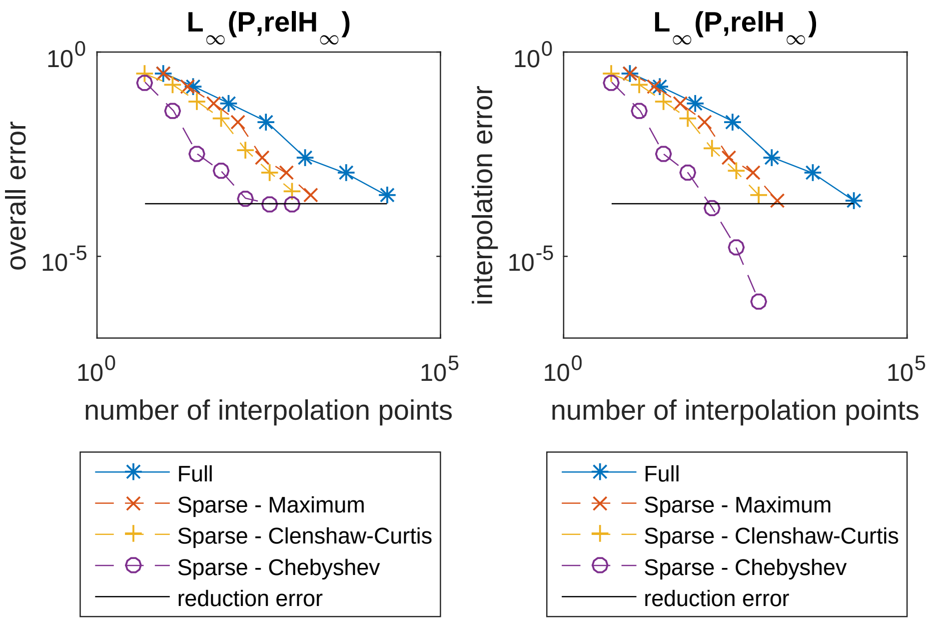
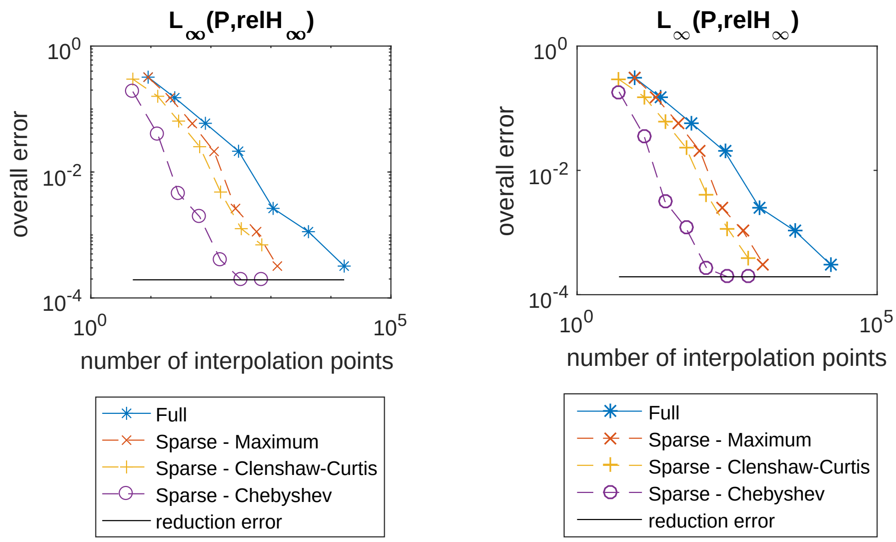


Figure 7.7: The overall (left) and interpolation (right) error for interpolation on full grids and different types of sparse grids.

7.2.2 Influence of Additional Reduced Order Basis Transformation

We analyze the impact of step 3d in Algorithm 4.1.

When we do not eliminate $\mathbf{E}_r$ by a basis transformation to $\mathbf{I}_r$ (i.e., skip step 3d in Algorithm 4.1), we need to interpolate it the same way as $-\mathbf{A}_r$ (in Algorithm 4.2). That is, the number of matrix entries that are interpolated approximately doubles and the matrix exponential is applied twice. However, the results are very similar—at least for this benchmark (see Figure 7.8).



(a) Keep and interpolate $\mathbf{E}_r$.

(b) Eliminate $\mathbf{E}_r$.

Figure 7.8: The overall error for without (left) and with (right) elimination of $\mathbf{E}_r$.

7.2.3 Reduction Order

We analyze the dependence on the reduction order, i.e., the size of the reduced order models. When we increase the reduction order r, the reduction error decreases (see Figure 7.9).

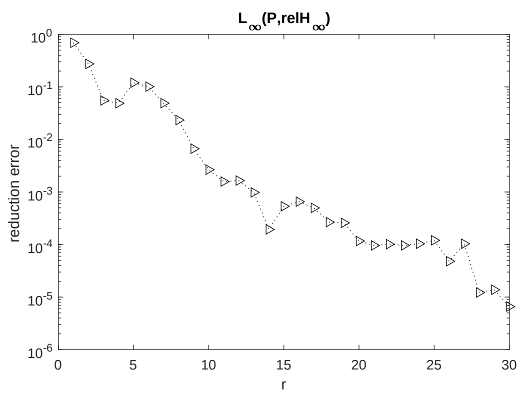
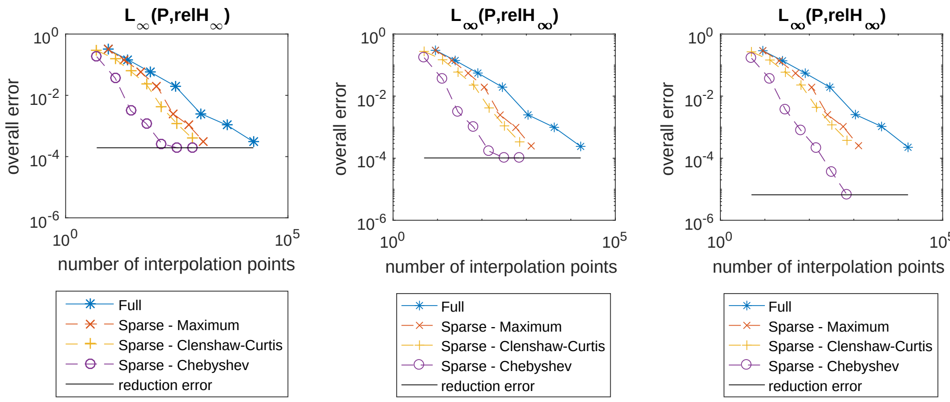


Figure 7.9: Reduction error for different reduction order r.

In Figure 7.10 the curves are the same for all three values of the reduction order until the reduction error is reached (for the CHEBYSHEV sparse grid). That is, the reduction order just influences the smallest reachable error.

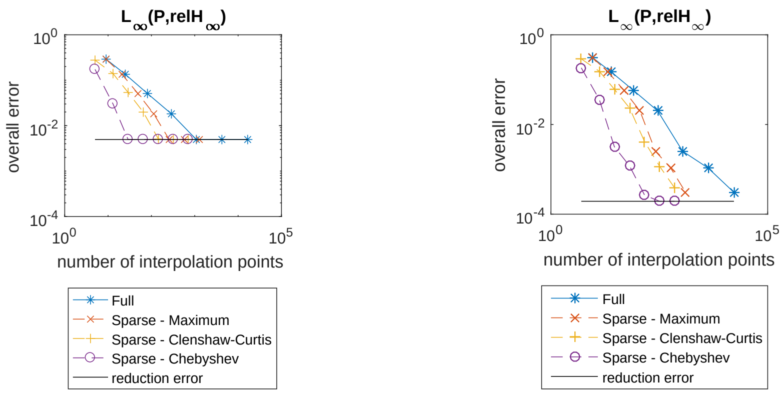


(a) $r = 14$. (b) $r = 22$. (c) $r = 30$.

Figure 7.10: The overall error for different reduction orders.

7.2.4 Multi-input, Multi-output and Single-input, Single-output Systems

We compare the results for multi-input, multi-output and single-input, single-output systems. The reduction errors differ, but the curves of the overall error are very similar (see Figure 7.11) for all the different grid types (until the reduction error is reached).



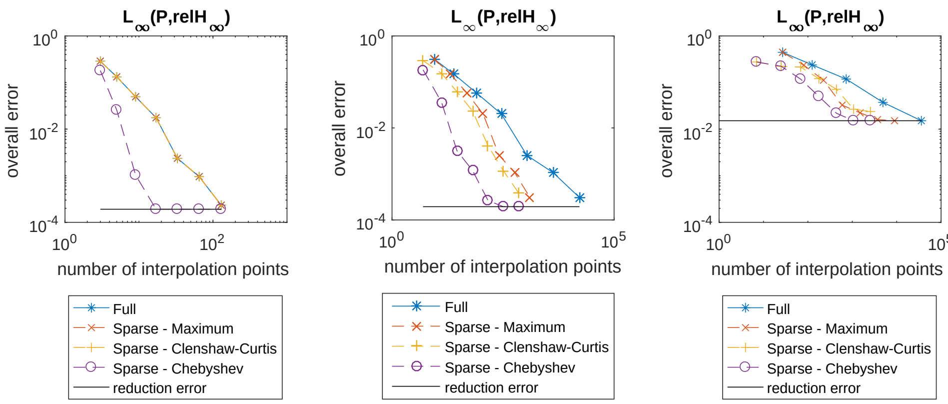
(a) Multi-input, multi-output system. (b) Single-input, single-output system.

Figure 7.11: The overall error for MIMO (left) and SISO (right) systems.

7.2.5 Dimension of Parameter Space

We vary the dimension of the parameter space from 1 to 3.

In the one-dimensional case, interpolation on full grids and sparse grids with piece-wise linear basis functions (Maximum, Clenshaw–Curtis) is the same (see Figure 7.12a). Again, sparse grid interpolation with global polynomial basis functions (Chebyshev) is more efficient.



(a) $D = 1$. (b) $D = 2$. (c) $D = 3$.

Figure 7.12: The overall error for different dimensions of the parameter space.

7.2.6 Performance of Matrix Interpolation PMOR Methods

In [BBH$^+$15] the performance of several state-of-the-art PMOR methods is compared in terms of accuracy and run-time. One of the methods is "MatrInt" [PMEL10, GPWL14], a stability-preserving PMOR method that bases on interpolation of system matrices. The second benchmark in [BBH$^+$15] is named "microthruster" and is the one-dimensional variant of the "Thermal Model" benchmark from subsection 7.1.1 with $p = p_1 \in [10^0, 10^4]$ and $p_2 = p_3 = 10^4$.

All PMOR methods that are included in this comparison, except "MatrInt", use a global reduced order basis. This in general leads to larger reduction orders, especially when the parametric full order model varies substantially over the parameter space. Furthermore, the online phase of those methods is only independent of the full order if the parameter dependence is affine.

For the "microthruster" benchmark these methods have reduction order $r = 100$, whereas "MatrInt" has reduction order $r = 10$. In this slightly unfair setting "MatrInt" achieves the worst accuracy (in the $\mathcal{L}_\infty(\mathcal{P}, \mathrm{rel}\mathcal{H}_\infty)$-norm) of all methods under consideration. However, the large overall error is not due to the reduction error but due to the interpolation error, so this casts an unfavorable light on PMOR methods that base on interpolation of system matrices.

It should be mentioned that “MatrInt” can be applied to arbitrary stable systems and does not exploit the special benchmark structure as the other methods (affine parameter dependence) and our method (symmetry of the system) do.

In this section we demonstrate that, for the aforementioned benchmark, we can achieve much better accuracy with interpolation of system matrices when taking advantage of the special structure. We do this by comparing our method, which is tailored to symmetric systems, with “MatrInt”. Moreover, our offline phase is much faster because in “MatrInt” several complete runs of the Iterative Rational Krylov Algorithm are performed and, for stability preservation, several Lyapunov equations are solved.

A new version of [BBH$^+$15] is to appear as [BBH$^+$16].

The authors of [BBH$^+$16] kindly provided their PMOR method comparison code including the implementation of “MatrInt”, which is a combination of the methods from [PMEL10, GPWL14] and written by Heiko Panzer, Matthias Geuss and Maria Cruz Varona.

Let us note that “MatrInt” guarantees stability only if the basis functions of the interpolation are non-negative, which does allow piece-wise linear interpolation on full grids. Neither sparse grids nor global polynomials can be used, which is why “MatrInt” is not suited for higher-dimensional parameter spaces.

Recall that in one-dimensional spaces, as it is the case for the “microthruster” benchmark, full grids and sparse grids coincide. We interpolate with piece-wise linear basis functions and $N = 10$ grid points. We measure the error at 577 (and not at 100 [BBH$^+$15]) points. In contrast to [BBH$^+$15], the online phase here only consists of establishing the reduced order model at an arbitrary point in parameter space without any simulation.

In this section we compute with MATLAB® R2015b 32-bit on an Acer Aspire XC-780 desktop pc with an x64-based Intel® Core$^{\text{TM}}$ i5-6400 (CPU @ 2.70GHz 2.71GHz) processor, 8.00GB RAM and 64-bit Windows® 10 operating system.

Our PMOR method yields a better accuracy than “MatrInt” (see Figure 7.13). The overall error of our method is smaller even though our reduction error (spiky valleys) is much larger. This is because our interpolation error (round hills) is significantly smaller.

At the interpolation points our method yields local reduced order models whose error is much larger than that of the local reduced order models of “MatrInt”. This is due to the fact that, for the sake of stability, we perform one-sided instead of two-sided projection and do not use the optimal frequency interpolation points from the Iterative Rational Krylov Algorithm at that specific point, but the ones from the center of the parameter space.

Using the optimal frequency interpolation points from the Iterative Rational Krylov Algorithm at every interpolation point in parameter space, as is done in “MatrInt”, can cause jumps in the reduction error of the local reduced order models as well as in the interpolation and overall errors in-between (see Figure 7.13 at $5500 < p < 6500$).

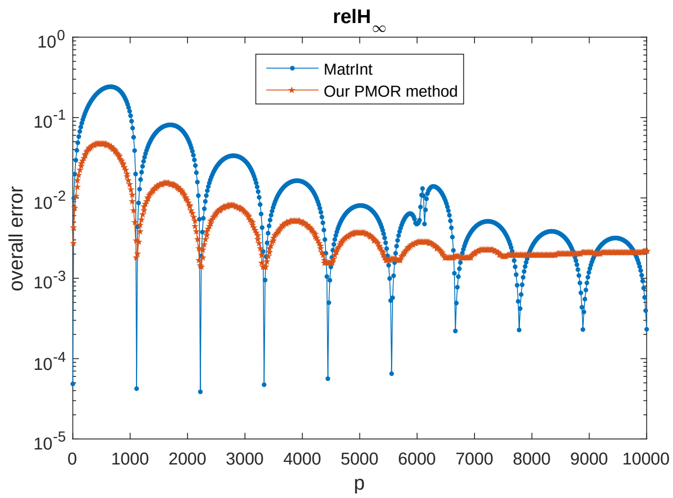


Figure 7.13: The overall error over the parameter space—comparison with [PMEL10, GPWL14] for $D = 1$.

For the symmetric "microthruster" ("Thermal Model") system our PMOR method yields a smaller overall error for the same number of interpolation points than "MatrInt", while it has an offline phase of about 1/30 length and the online phase is shorter, too (see Table 7.2).

Table 7.2: $\mathcal{L}_\infty(\mathcal{P}, \mathrm{rel}\mathcal{H}_\infty)$-error and offline phase and online phase computing times—comparison with [PMEL10, GPWL14] for $D = 1$.

	$\mathcal{L}_\infty(\mathcal{P}, \mathrm{rel}\mathcal{H}_\infty)$	offline phase	online phase (mean)
Our PMOR method	4.79e-2	3.6 s	0.00062 s
"MatrInt"	2.42e-1	115.4 s	0.00206 s

Now we vary the reduction order r and the number N of interpolation points in our PMOR method to demonstrate the influence of those two quantities on the error. We show that a matrix interpolation PMOR method can achieve similar results as other PMOR methods in [BBH$^+$15] when we properly chose the reduction order r and the number N of interpolation points.

As also mentioned in [BBH$^+$15], comparing the accuracy of PMOR methods with reduction order $r = 100$ with the accuracy of PMOR methods with reduction order $r = 10$ is slightly unfair. It seems natural to increase the reduction order of our method to be more competitive with the other methods considered in [BBH$^+$15]. However, since the overall error is dominated by the interpolation error (see Figure 7.14), increasing the number N of interpolation points (from $N = 10$ to $N = 37$) is a better strategy for decreasing the overall error. This leads to an increased offline computation time (see Table 7.3). The overall error cannot be made arbitrarily small by increasing the number N of interpolation points, since the overall error cannot

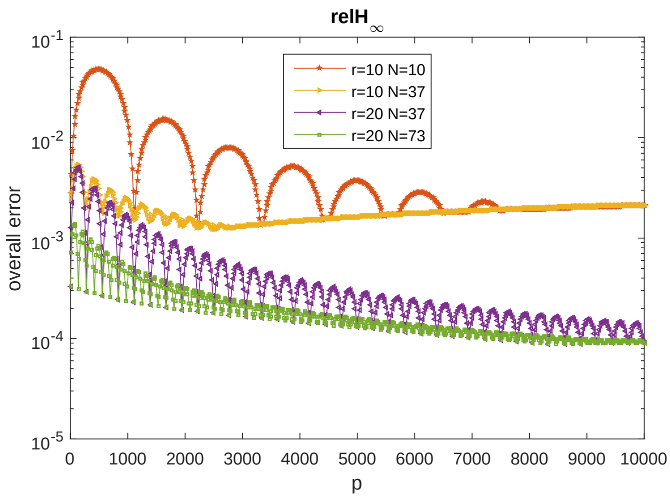


Figure 7.14: The overall error over the parameter space for different numbers of interpolation points and reduction orders.

get below the reduction error of the local reduced order models at the interpolation points. Hence, if we want to achieve better accuracies, then we need to increase the reduction order r (from $r = 10$ to $r = 20$) next. This leads to increased offline and online computation times (see Table 7.3). Even though the reduction error of the local reduced order models at the interpolation points decreases, the overall error stays about the same (see Figure 7.14). When we now further increase the number N of interpolation points (from $N = 37$ to $N = 73$), then we again get a smaller overall error (see Figure 7.14).

Table 7.3: $\mathcal{L}_\infty(\mathcal{P}, \mathrm{rel}\mathcal{H}_\infty)$-error and offline phase and online phase computing times for different number of interpolation points and reduction orders.

r	N	$\mathcal{L}_\infty(\mathcal{P}, \mathrm{rel}\mathcal{H}_\infty)$	offline phase	online phase (mean)
10	10	4.79e-2	3.6 s	0.00062 s
10	37	5.34e-3	8.2 s	0.00066 s
20	37	5.02e-3	24.0 s	0.00135 s
20	73	1.41e-3	38.9 s	0.00157 s

The experiments demonstrate that it is important to choose the reduction order and the number of interpolation points carefully to balance interpolation and reduction error and to get an overall error of certain accuracy with shortest possible run-time.

Although our PMOR method is suitable for higher-dimensional parameter spaces it performs well also in this one-dimensional setting and shows that matrix interpolation can yield similar results as the other PMOR methods considered in [BBH+15] while the reduction order r is smaller than 100 (compare Table 7.3 and Table 7.4).

Table 7.4: $\mathcal{L}_\infty(\mathcal{P}, \mathrm{rel}\mathcal{H}_\infty)$-error and offline phase computing time of methods in [BBH+15] from [BBH+15, Table 2] and [BBH+15, Table 5].

	$\mathcal{L}_\infty(\mathcal{P}, \mathrm{rel}\mathcal{H}_\infty)$	offline phase
"POD"	2.9e-3	7.83 s
"POD-Greedy"	1.3e-3	45.48 s
"TransFncInt"	9.2e-3	39.06 s
"PWH2TanInt"	3.3e-6	120.39 s
"MultiPMomMtch"	2.9e-2	4.06 s
"emWX"	1.5e-2	11.73 s
"MatrInt"	1.1e-1	110.60 s

7.2.7 Silicon Nitride Membrane Benchmark

Now we consider a benchmark with a four-dimensional parameter space. The full order model is of order $n = 60{,}020$ and we reduce it to models of order $r = 10$.

Unlike in the default benchmark, we do not use a full grid as the parameter space test grid $\mathcal{P}_{\mathrm{test}}$ since it is too expensive in four dimensions. Random points are cheaper, but they do not spread evenly and tend to cluster. A point set with low discrepancy is more suitable. Hence, we choose 100 points of a so-called HALTON *sequence* [KW97], which we scale to fit the parameter space, as the test grid. Additionally, we set the test grid to 100 other points of the same HALTON sequence to verify that those points are representative. As can be seen in Figure 7.15, the errors measured on those two test grids differ only slightly.

Like in the other benchmark, the overall error goes down until the reduction error is met. Again, interpolation with global polynomial basis functions turns out to be more efficient than interpolation with piece-wise linear basis functions.

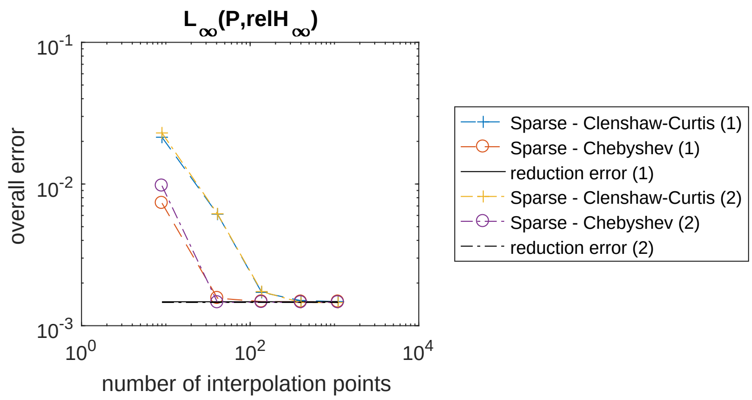


Figure 7.15: The overall error for the SiN Membrane benchmark.

7.3 Comparison with Theoretical Results

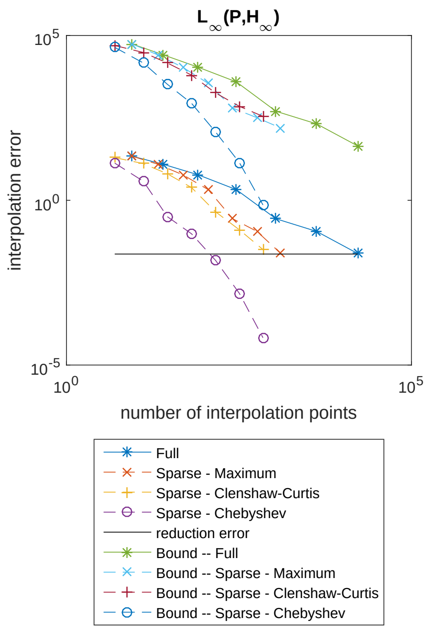
We compare the computed absolute (not relative) interpolation error with the theoretical results from chapter 5 and chapter 6 in the default benchmark setting for different types of grids.

7.3.1 Interpolation Error Bound

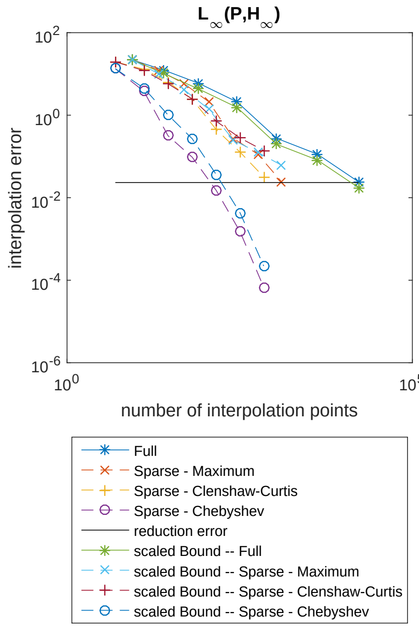
We compare the computed interpolation error with the error bound from Theorem 5.13

$$
\begin{aligned}
&\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} \\
&\leq \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \cdot \Big\{ (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}) \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot (e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} - 1) \\
&\qquad + \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty} \\
&\qquad + e^{\|\Delta\mathfrak{A}_r\|_{\mathcal{L}_\infty}} \cdot (\|\mathbf{B}_r\|_{\mathcal{L}_\infty} + \|\Delta\mathbf{B}_r\|_{\mathcal{L}_\infty}) \cdot \|\Delta\mathbf{C}_r\|_{\mathcal{L}_\infty} \Big\} .
\end{aligned}
$$

The error bound is some magnitudes larger than the actual interpolation error (see Figure 7.16a), but when we multiply the bound with a constant factor, the shapes match well (see Figure 7.16b).



(a) Interpolation error and bound.



(b) Interpolation error and scaled bound.

Figure 7.16: The interpolation error and the (scaled) interpolation error bound from Theorem 5.13 for different types of grids.

The first summand belonging to $\Delta\mathfrak{A}_r$ dominates the other two belonging to $\Delta\mathbf{B}_r$ and $\Delta\mathbf{C}_r$ (see Figure 7.17).

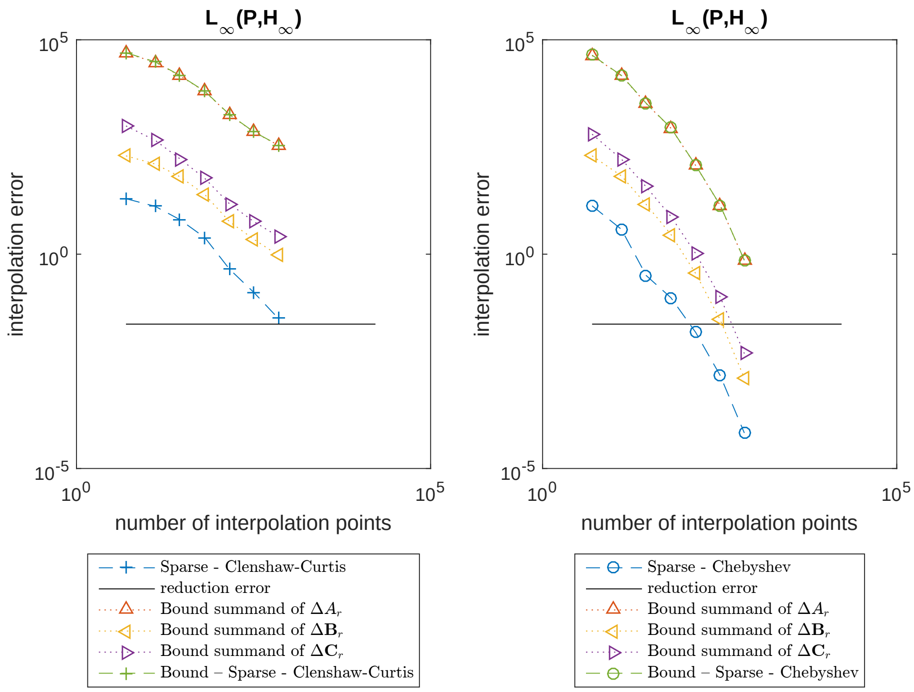


Figure 7.17: The interpolation error, the interpolation error bound from Theorem 5.13 and the bound's three summands belonging to $\Delta\mathfrak{A}_r$, $\Delta\mathbf{B}_r$ and $\Delta\mathbf{C}_r$, respectively.

Now we analyze where the large gap between the bound from Theorem 5.13 and the interpolation error stems from. The first summand of the bound belonging to $\Delta\mathfrak{A}_r$ is based on the inequality

$$\begin{aligned}&\sup_{s\in\imath\mathbb{R}} \|\mathbf{C}_r \cdot \{(s\mathbf{I}_r - \mathbf{A}_r)^{-1} - (s\mathbf{I}_r - \widetilde{\mathbf{A}}_r)^{-1}\} \cdot \widetilde{\mathbf{B}}_r\|_2 \\ &\leq \|\mathbf{A}_r^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}_r\|_2 \cdot \|\mathbf{C}_r\|_2 \cdot (e^{\|\Delta\mathfrak{A}_r\|_2} - 1)\end{aligned}$$

from Proposition 5.7 for fixed $\mathbf{A}_r$, $\mathbf{B}_r$ and $\mathbf{C}_r$, i.e., at a discrete point in parameter space.

Let us remark that we can derive the bound of Proposition 5.7 for arbitrary rectangular matrices $\widetilde{\mathbf{B}}_r$ and $\mathbf{C}_r$ from Proposition 5.7 for $\widetilde{\mathbf{B}}_r = \mathbf{C}_r = \mathbf{I}_r$:

$$\begin{aligned}
&\sup_{s\in\imath\mathbb{R}} \|\mathbf{C}_r \cdot \{(s\mathbf{I}_r - \mathbf{A}_r)^{-1} - (s\mathbf{I}_r - \widetilde{\mathbf{A}}_r)^{-1}\} \cdot \widetilde{\mathbf{B}}_r\|_2 \\
&\leq \|\mathbf{C}_r\|_2 \cdot \sup_{s\in\imath\mathbb{R}} \|(s\mathbf{I}_r - \mathbf{A}_r)^{-1} - (s\mathbf{I}_r - \widetilde{\mathbf{A}}_r)^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}_r\|_2 \\
&\leq \|\mathbf{A}_r^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}_r\|_2 \cdot \|\mathbf{C}_r\|_2 \cdot (e^{\|\Delta\mathfrak{A}_r\|_2} - 1).
\end{aligned}$$

We consider the one-dimensional setting to empirically evaluate the loss we incur by applying these inequalities. The first and the second bound differ by a factor of up to 11.5 (right plot of Figure 7.18), so we lose about one magnitude by applying the second inequality. The second bound is larger than the supremum on the left-hand side by a factor of up to $2.8e3$ (left plot of Figure 7.18). Hence, by using submultiplicativity to obtain the first inequality we lose about two orders of magnitude. This emphasizes the influence of the directions of $\widetilde{\mathbf{B}}_r$ and $\mathbf{C}_r$, which are ignored by Proposition 5.7. These directions should be taken into account both when trying to derive a better interpolation error bound for the algorithm presented in this thesis and when trying to improve the algorithm itself.

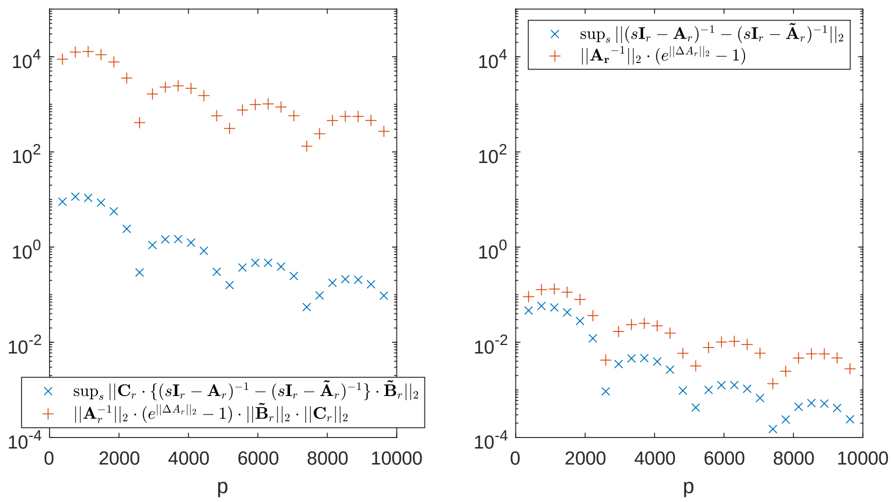


Figure 7.18: Both sides of the inequality from Proposition 5.7 for the one-dimensional setting (left) and the one-dimensional setting with $\widetilde{\mathbf{B}}_r = \mathbf{C}_r = \mathbf{I}_r$ (right).

7.3.2 Interpolation Error Convergence Rates

We compare the computed error with the convergence rates from chapter 6.

When the number of grid points N is sufficiently large (i.e., $\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty} \leq 1$ and $\|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty} \leq \|\mathbf{B}_r\|_{\mathcal{L}_\infty}$) we have for piece-wise linear basis functions under

Requirement 6.6 ($\mathfrak{A}_r \in \mathcal{W}^2_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$, $\mathbf{B}_r \in \mathcal{W}^2_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$ and $\mathbf{C}_r \in \mathcal{W}^2_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$) (see Theorem 6.8)

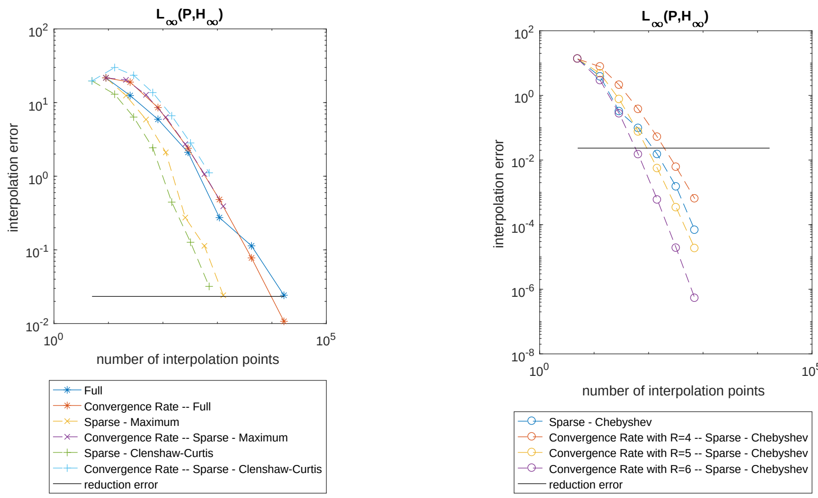
$$\begin{aligned}\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r^{(N)}\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} \leq c \cdot N^{-2} \cdot (\log N)^{3\cdot(D-1)} \cdot \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \\ \cdot \Big\{ 4r \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\mathfrak{A}_r\|_{\mathcal{W}^2_\infty} \\ + \sqrt{rm} \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{B}_r\|_{\mathcal{W}^2_\infty} \\ + 2e \cdot \sqrt{rq} \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{C}_r\|_{\mathcal{W}^2_\infty} \Big\}\end{aligned}$$

and for global polynomial basis functions under Requirement 6.11 ($\mathfrak{A}_r \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R}^{r\times r})$, $\mathbf{B}_r \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R}^{r\times m})$ and $\mathbf{C}_r \in \mathcal{C}^R_\infty(\mathcal{P}, \mathbb{R}^{q\times r})$) (see Theorem 6.14)

$$\begin{aligned}\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r^{(N)}\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} \leq c \cdot N^{-R} \cdot (\log N)^{(R+2)(D-1)+1} \cdot \|\mathbf{A}_r^{-1}\|_{\mathcal{L}_\infty} \\ \cdot \Big\{ 4r \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\mathfrak{A}_r\|_{\mathcal{C}^R_\infty} \\ + \sqrt{rm} \cdot \|\mathbf{C}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{B}_r\|_{\mathcal{C}^R_\infty} \\ + 2e \cdot \sqrt{rq} \cdot \|\mathbf{B}_r\|_{\mathcal{L}_\infty} \cdot \|\mathbf{C}_r\|_{\mathcal{C}^R_\infty} \Big\}.\end{aligned}$$

For this benchmark, the requirements $\|\Delta\mathfrak{A}_r{}^{(N)}\|_{\mathcal{L}_\infty} \leq 1$ and $\|\Delta\mathbf{B}_r^{(N)}\|_{\mathcal{L}_\infty} \leq \|\mathbf{B}_r\|_{\mathcal{L}_\infty}$ is already fulfilled for the N of the smallest level.

We cannot prove that the model of the benchmark fulfills Requirement 6.6 or Requirement 6.11, but Figure 7.19 indicates that Theorem 6.8 and Theorem 6.14 (for $R = 4$) are applicable.



(a) For piece-wise linear basis functions (see Theorem 6.8).

(b) For global polynomial basis functions (see Theorem 6.14).

Figure 7.19: Interpolation error and convergence rates.

8 Conclusions and Outlook

8.1 Conclusions

We proposed a fully-automatic parametric model order reduction method with stability guarantee for symmetric linear time-invariant dynamical systems with a higher-dimensional parameter space.

We proved that the transfer functions of the full order model and the interpolated reduced order model fulfill one-sided tangential interpolation conditions at all combinations of frequency and parameter space interpolation points.

We heavily reduced the number of interpolation points in parameter space required to reach a certain accuracy by interpolating on sparse grids instead of full grids. When the parameter dependence is very smooth, then we can further reduce the number of interpolation points by using global polynomial instead of piece-wise linear basis functions. Hence our method is more efficient in higher-dimensional parameter spaces than other stability-preserving PMOR methods.

We showed that even in a one-dimensional parameter space our method gives better results for symmetric systems than another stability-preserving PMOR method that includes interpolation of the system matrices, but is not suited for higher-dimensional parameter spaces and is mentioned in the PMOR method comparison [BBH$^+$15] but performs not so well. This indicates that interpolation of the system matrices is competitive to the other methods mentioned in [BBH$^+$15] and should be explored further as is has several advantages over the other methods which all use a global reduced order basis. For example, the reduction order is smaller and the parameter dependence can be arbitrary and need not be affine.

While deriving the interpolation error bound we pursued different approaches to relate the transfer function error to the error in the matrices that are interpolated. It turns out that our approach is much better than when using the standard theorem about perturbation of the inverse. The constant is smaller and we have no locality restriction.

Still, the upper bound for the interpolation error is some orders of magnitude too large, since submultiplicativity is too pessimistic. Nevertheless the shape of the bound is good—bound and error essentially differ by a multiplicative constant. So the course of the bound might be used together with error computations of interpolation on a coarse grid to determine the optimal number of grid points, where the interpolation error approximately equals the reduction error.

8.2 Outlook on Future Work

In the rest of this section we give directions for further investigation and mention open problems.

8.2.1 Model

Our method can simply be extended from first-order to second-order systems with symmetric positive definite system matrices. Since the standard Iterative Rational Krylov Algorithm is not applicable to second-order systems, another interpolatory model order reduction method which provides a good choice of interpolation points in frequency space is needed (see e.g., [Wya12]).

It would be nice to test our method on other benchmarks and applications, especially with a larger number of parameters. However, the symmetry requirement is a strong restriction for a system. We only found benchmarks with at most 4 parameters, except an example with 7 parameters in [BB15], which is not open source, but property of the Robert Bosch GmbH. Nevertheless, our method is also suitable for higher dimensions.

For non-symmetric systems it is not clear so far how to preserve stability in higher-dimensional parameter spaces. In [GPWL14] a stability-preserving PMOR method for general systems is presented. The idea is to transform the local stable reduced order models such that the symmetric part of the matrix $\mathbf{A}_r$ becomes symmetric negative definite. Such systems are called contractive. However, non-negative weighting functions are required to ensure that the symmetric part of the interpolated matrix $\widetilde{\mathbf{A}}_r$ is symmetric negative definite, too. So one cannot interpolate on sparse grids to deal with higher-dimensional parameter spaces.

A possible remedy to avoid the requirement of non-negative weighting functions might be as follows. First, we split $\mathbf{A}_r = \mathbf{M} + \mathbf{N}$ into a symmetric part $\mathbf{M} := (\mathbf{A}_r + \mathbf{A}_r^T)/2$ and a skew-symmetric part $\mathbf{N} := (\mathbf{A}_r - \mathbf{A}_r^T)/2$. Then we interpolate $\mathbf{M}$ and $\mathbf{N}$ separately: the matrix $-\mathbf{M}$ on the manifold of symmetric positive definite matrices (as $\mathbf{A}_r$ in this thesis) and the matrix $\mathbf{N}$ by direct matrix interpolation. So the interpolated matrices $\widetilde{\mathbf{M}}$ and $\widetilde{\mathbf{N}}$ are symmetric negative definite and skew-symmetric, respectively. The interpolant of the matrix $\mathbf{A}_r$ is then defined as $\widetilde{\mathbf{A}}_r := \widetilde{\mathbf{M}} + \widetilde{\mathbf{N}}$. Note that the symmetric part of $\widetilde{\mathbf{A}}_r$ is $(\widetilde{\mathbf{A}}_r + \widetilde{\mathbf{A}}_r^T)/2 = (\widetilde{\mathbf{M}} + \widetilde{\mathbf{M}}^T)/2 + (\widetilde{\mathbf{N}} + \widetilde{\mathbf{N}}^T)/2 = \widetilde{\mathbf{M}}$.

It feels a bit unnatural to interpolate the two parts with two different concepts, but it might be worth a try. Note that the local reduced order models need to be stable, so use e.g., Balanced Truncation, and a mass matrix $\mathbf{E}_r \neq \mathbf{I}_r$ needs to be inverted to the other side, i.e., $\mathbf{A}_r \rightsquigarrow \mathbf{E}_r^{-1} \cdot \mathbf{A}_r$ and $\mathbf{E}_r \rightsquigarrow \mathbf{I}_r$.

8.2.2 Model Order Reduction

To construct the local reduced order bases we use the $\mathcal{H}_2$-optimal frequency interpolation points of the center of the parameter space, which we compute with the Iterative Rational Krylov Algorithm, for the whole parameter space.

A combination of interpolation and approximation in frequency space might give reduced order models of a better accuracy over the whole parameter space [BG15]. But to include it into our stability-preserving parametric model order reduction method we would need a symmetric variant of the method proposed in [BG15] and ideally an extension for multi-input, multi-output systems.

To make the local reduced order models compatible, a reference basis is chosen. We take the reduced order basis at the center of the parameter space, i.e., $\mathbf{V}_{\text{ref}} = \mathbf{V}_{\ell_0}$ with ℓ_0 such that $\mathbf{p}_{\ell_0} = \mathbf{p}_{\text{center}}$. Another possibility inspired by [PMEL10, BBBG11] is to define $\mathbf{V}_{\text{ref}}$ as the first r singular vectors of $\mathbf{V}_{\text{all}} = [\mathbf{V}_1, \ldots, \mathbf{V}_N]$. Since $\mathbf{V}_{\text{all}}$ can become quite large, one should not use the final N that is used for interpolation, but a smaller one that belongs to a coarse grid.

8.2.3 Interpolation

The proposed parametric model order reduction method is not restricted to interpolation on sparse grids. Any interpolation or even approximation method can be used, but it should be suitable for higher-dimensional spaces. The preservation of stability is not affected.

For very high dimensional parameter spaces energy-norm based sparse grids might be advantageous. One can also think of inventing a completely different sparse grid. However, it is not obvious which norm to take for the cost–benefit optimization.

Currently we treat all entries of the matrices which are interpolated equally and interpolate the matrix $\mathbf{A}_r$ regardless of the matrices $\mathbf{B}_r$ and $\mathbf{C}_r$. However, minimizing the error in the transfer function is not equivalent to minimizing $\|\mathbf{A}_r - \widetilde{\mathbf{A}}_r\|$ or $\|\mathbf{A}_r^{-1} - \widetilde{\mathbf{A}}_r^{-1}\|$. It is an intriguing question whether it is possible to consider the matrices $\mathbf{B}_r$ and $\mathbf{C}_r$ when interpolating the matrix $\mathbf{A}_r$.

One possibility could be to interpolate $\log(-\mathbf{M}^T\mathbf{A}_r\mathbf{M})$ instead of $\log(-\mathbf{A}_r)$ for an invertible matrix $\mathbf{M}$. This might yield better results if the matrix $\mathbf{M}$ is chosen carefully, depending on $\mathbf{B}_r$ and $\mathbf{C}_r$.

When the dimensions have different importance, the number of interpolation points needed to reach a certain interpolation error could be made smaller by dimension adaptivity [GG03], i.e., using different refinement levels in each dimension. Before dimension adaptivity can be applied successfully a suitable error estimator that correctly portrays the interplay of the error in the different matrix entries of the three system matrices needs to be constructed.

Spacial adaptivity is only needed if the parameter dependence is very non-smooth in some regions of the parameter space. An error estimator for it might be constructed from an error estimator for dimension adaptivity.

A Definitions and Theorems from Literature

A.1 Matrix Basics

In this section we recapitulate some definitions and facts about matrices. These can be found in many textbooks, see for example [GVL89, Lan69, Ste98].

A.1.1 Properties

Definition A.1. A matrix $\mathbf{M} = [M_{ij}]_{i,j=1}^{n} \in \mathbb{C}^{n \times n}$ is

- *diagonal* if $M_{ij} = 0$ for $i \neq j$,
- *upper triangular* if $M_{ij} = 0$ for $i > j$,
- *lower triangular* if $M_{ij} = 0$ for $i < j$.

Definition A.2. A matrix $\mathbf{M} \in \mathbb{C}^{n \times n}$ is

- *symmetric* if $\mathbf{M} = \mathbf{M}^T$,
- *skew-symmetric* if $\mathbf{M} = -\mathbf{M}^T$,
- Hermitian if $\mathbf{M} = \mathbf{M}^H$,
- *unitary* if $\mathbf{M}^H\mathbf{M} = \mathbf{I}_n$,
- *normal* if $\mathbf{M}^H\mathbf{M} = \mathbf{M}\mathbf{M}^H$.

Every real symmetric matrix is Hermitian.
Every Hermitian matrix and, hence, every real symmetric matrix, is normal.

Definition A.3. A matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ is *orthogonal* if $\mathbf{M}^T\mathbf{M} = \mathbf{I}_n$.

Every real unitary matrix is orthogonal.

Definition A.4. A matrix $\mathbf{M} \in \mathbb{R}^{n \times m}$ has *orthonormal columns* if $\mathbf{M}^T\mathbf{M} = \mathbf{I}_m$.

Definition A.5. A Hermitian matrix $\mathbf{M} \in \mathbb{C}^{n \times n}$ is

- *positive definite* if its eigenvalues are positive, i.e., $\Lambda(\mathbf{M}) \subset \mathbb{R}^+$,
- *negative definite* if its eigenvalues are negative, i.e., $\Lambda(\mathbf{M}) \subset \mathbb{R}^-$.

A.1.2 Transformations

Definition A.6. Let $\mathbf{M} \in \mathbb{C}^{n\times n}$ and let $\mathbf{S} \in \mathbb{C}^{n\times n}$ be invertible. Then $\mathbf{N} = \mathbf{S}^{-1}\mathbf{M}\mathbf{S}$ is called a *similarity transformation* of $\mathbf{M}$. The matrices $\mathbf{M}$ and $\mathbf{N}$ are said to be *similar*.

Theorem A.7 ([Lan69, Theorem 2.4.1])**.** *Similar matrices have the same eigenvalues.*

Definition A.8. Let $\mathbf{M} \in \mathbb{R}^{n\times n}$ be symmetric and $\mathbf{S} \in \mathbb{R}^{n\times n}$ be invertible. Then $\mathbf{N} = \mathbf{S}^T\mathbf{M}\mathbf{S}$ is called a *congruence transformation* of $\mathbf{M}$. The matrices $\mathbf{M}$ and $\mathbf{N}$ are said to be *congruent*.

Theorem A.9 (Sylvester's law of inertia [GVL89, Theorem 8.1.12])**.** *Congruent matrices have the same number of negative, zero and positive eigenvalues, respectively.*

A.1.3 Decompositions

Theorem A.10 (singular value decomposition [Ste98, Theorem 4.27 in Chapter 1])**.** *Let* $\mathbf{M} \in \mathbb{C}^{n\times m}$. *Let* $k := \min\{n, m\}$. *Then there exist unitary matrices* $\mathbf{U} \in \mathbb{C}^{n\times n}$ *and* $\mathbf{Z} \in \mathbb{C}^{m\times m}$ *such that*

$$\mathbf{U}^H\mathbf{M}\mathbf{Z} = \begin{bmatrix} \mathbf{\Sigma} \\ \mathbf{O} \end{bmatrix} \text{ for } n \geq m$$

or

$$\mathbf{U}^H\mathbf{M}\mathbf{Z} = \begin{bmatrix} \mathbf{\Sigma} & \mathbf{O} \end{bmatrix} \text{ for } n \leq m,$$

where $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k) \in \mathbb{R}^{k\times k}$ *with* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k \geq 0$.

This decomposition is called singular value decomposition *of the matrix* $\mathbf{M}$. *The real numbers* σ_i, $i = 1, \ldots, k$, *are called the* singular values *of the matrix* $\mathbf{M}$.

Theorem A.11 (spectral decomposition [Ste98, Theorem 4.33 in Chapter 1])**.** *Let* $\mathbf{M} \in \mathbb{C}^{n\times n}$ *be* Hermitian. *Then there exists a unitary matrix* $\mathbf{U}$ *such that*

$$\mathbf{U}^H\mathbf{M}\mathbf{U} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),$$

where $\lambda_j \in \mathbb{R}$ *and* $\lambda_1 \geq \cdots \geq \lambda_n$.

This decomposition is called the spectral decomposition *of the matrix* $\mathbf{M}$. *The numbers* λ_i *are the* eigenvalues *of the matrix* $\mathbf{M}$, *since they fulfill* $\mathbf{M}\mathbf{u}_j = \lambda_j\mathbf{u}_j$ *where* $\mathbf{u}_j$ *is the* j*th column of the matrix* $\mathbf{U}$.

Definition A.12. A matrix $\mathbf{M} \in \mathbb{C}^{n\times n}$ is *diagonalizable* if there exists an invertible matrix $\mathbf{X} \in \mathbb{C}^{n\times n}$ such that $\mathbf{X}^{-1}\mathbf{M}\mathbf{X} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Theorem A.13 ([GVL89, Corollary 7.1.4])**.** *A matrix* $\mathbf{M} \in \mathbb{C}^{n\times n}$ *is normal if and only if there exists a unitary matrix* $\mathbf{U} \in \mathbb{C}^{n\times n}$ *such that* $\mathbf{U}^H\mathbf{M}\mathbf{U} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Theorem A.14 ([GVL89, Theorem 8.1.1]). *Let $\mathbf{M} \in \mathbb{R}^{n\times n}$ be symmetric. Then there exists an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{n\times n}$ such that*

$$\mathbf{Q}^T\mathbf{M}\mathbf{Q} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Theorem A.15 (CHOLESKY decomposition [Ste98, Theorem 2.7 in Chapter 3]). *A positive definite matrix $\mathbf{M} \in \mathbb{C}^{n\times n}$ can be factored uniquely in the form $\mathbf{M} = \mathbf{R}^H\mathbf{R}$, where the matrix $\mathbf{R}$ is upper triangular with positive diagonal entries.*

This decomposition is called the CHOLESKY decomposition *of the matrix* $\mathbf{M}$.

A.1.4 Projections

Theorem A.16 ([Ste98, Theorem 2.2 in Chapter 3]). *Let $\mathbf{M} \in \mathbb{C}^{n\times n}$ be positive definite and let $\mathbf{V} \in \mathbb{C}^{n\times r}$. Then $\mathbf{N} = \mathbf{V}^H\mathbf{M}\mathbf{V}$ is positive definite if and only if* $\operatorname{rank}(\mathbf{V}) = r$.

Corollary A.17. *Let $\mathbf{M} \in \mathbb{C}^{n\times n}$ be negative definite and let $\mathbf{V} \in \mathbb{C}^{n\times r}$. Then $\mathbf{N} = \mathbf{V}^H\mathbf{M}\mathbf{V}$ is negative definite if and only if* $\operatorname{rank}(\mathbf{V}) = r$.

Proof. Apply Theorem A.16 with $\mathbf{M} = -\mathbf{M}$. □

Theorem A.18 (CAUCHY interlacing theorem [Ste98, Theorem 4.34. 5. in Chapter 1]). *Let $\mathbf{M} \in \mathbb{R}^{n\times n}$ be symmetric and let $\mathbf{V} \in \mathbb{R}^{n\times(n-1)}$ have orthonormal columns. Then for $\mathbf{N} = \mathbf{V}^T\mathbf{M}\mathbf{V}$ it holds*

$$\lambda_1(\mathbf{M}) \geq \lambda_1(\mathbf{N}) \geq \lambda_2(\mathbf{M}) \geq \lambda_2(\mathbf{N}) \geq \cdots \geq \lambda_{n-1}(\mathbf{N}) \geq \lambda_n(\mathbf{M}).$$

Corollary A.19. *Let $\mathbf{M} \in \mathbb{R}^{n\times n}$ be symmetric and let $\mathbf{V} \in \mathbb{R}^{n\times r}$ have orthonormal columns. Then for $\mathbf{N} = \mathbf{V}^T\mathbf{M}\mathbf{V}$ it holds*

$$\lambda_1(\mathbf{M}) \geq \lambda_1(\mathbf{N}) \qquad \text{and} \qquad \lambda_r(\mathbf{N}) \geq \lambda_n(\mathbf{M}).$$

Proof. Apply Theorem A.18 recursively. □

A.1.5 Bounds for Matrix Exponential and Matrix Logarithm

The next lemma is needed to bound the norm of the matrix exponential by the exponential of the matrix norm.

Lemma A.20 ([HJ91, (6.5.25)]). *Let $\mathbf{M} \in \mathbb{C}^{n\times n}$. Then*

$$\|e^{\mathbf{M}}\| \leq e^{\|\mathbf{M}\|}.$$

For the matrix logarithm of a symmetric positive definite matrix and the spectral norm we prove a similar expression ourselves (see Lemma B.9), which is not only a bound, but even an equality:

$$\|\log(\mathbf{M})\|_2 = \max\left\{|\log(\|\mathbf{M}\|_2)|, |\log(\|\mathbf{M}^{-1}\|_2)|\right\}.$$

A.2 Perturbation Theory

A.2.1 Matrix Inverse

Theorem A.21 ([GVL89, Theorem 2.3.4]). *Let* $\mathbf{M}, \mathbf{F} \in \mathbb{R}^{n\times n}$ *with* $\mathbf{M}$ *invertible and* $\|\mathbf{M}^{-1}\mathbf{F}\| = \delta < 1$, *where* $\|\cdot\|$ *is a submultiplicative matrix norm. Then* $\mathbf{M}+\mathbf{F}$ *is invertible and*

$$\|\mathbf{M}^{-1} - (\mathbf{M}+\mathbf{F})^{-1}\| \leq \|\mathbf{M}^{-1}\|^2 \cdot \|\mathbf{F}\|/(1-\delta).$$

Corollary A.22. *Let* $\mathbf{M}, \mathbf{F} \in \mathbb{R}^{n\times n}$ *with* $\mathbf{M}$ *invertible and* $\|\mathbf{M}^{-1}\mathbf{F}\| \leq 1/2$, *where* $\|\cdot\|$ *is a submultiplicative matrix norm. Then* $\mathbf{M}+\mathbf{F}$ *is invertible and*

$$\|\mathbf{M}^{-1} - (\mathbf{M}+\mathbf{F})^{-1}\| \leq 2 \cdot \|\mathbf{M}^{-1}\|^2 \cdot \|\mathbf{F}\|.$$

A.2.2 Matrix Exponential

Theorem A.23 ([HJ91, Theorem 6.5.29]). *Let* $\mathbf{M}, \mathbf{F} \in \mathbb{C}^{n\times n}$ *with* $\mathbf{M}$ *normal. Then*

$$\|e^{\mathbf{M}+\mathbf{F}} - e^{\mathbf{M}}\|_2 \leq (e^{\|\mathbf{F}\|_2} - 1) \cdot \|e^{\mathbf{M}}\|_2.$$

Corollary A.24. *Let* $\mathbf{F} \in \mathbb{C}^{n\times n}$. *Then*

$$\|e^{\mathbf{F}} - \mathbf{I}\|_2 \leq e^{\|\mathbf{F}\|_2} - 1.$$

Proof. Apply Theorem A.23 with $\mathbf{M} = \mathbf{O}$. □

A.3 Miscellaneous

A.3.1 Hölder's Inequalities

Theorem A.25 (discrete Hölder's inequality [BSMM08, (1.120a)]). *Let* $p, q > 1$ *with* $1/p + 1/q = 1$. *Let* $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}_0^+$. *Then*

$$\sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} x_i^p\right)^{1/p} \left(\sum_{i=1}^{n} y_i^q\right)^{1/q}.$$

Theorem A.26 (continuous Hölder's inequality [Rou05, (1.19)]). *Let* $p, q > 1$ *with* $1/p + 1/q = 1$. *Let* $f \in \mathcal{L}_p(\Omega, \mathbb{R}^+)$ *and* $g \in \mathcal{L}_q(\Omega, \mathbb{R}^+)$. *Then*

$$\int_\Omega f(\omega) g(\omega)\ \mathrm{d}s \leq \left(\int_\Omega f(\omega)^p\ \mathrm{d}s\right)^{1/p} \left(\int_\Omega g(\omega)^q\ \mathrm{d}s\right)^{1/q}.$$

A.3.2 Laplace Transform

Definition A.27. [RR93, (5.145)] Let f be a vector-valued function. If it exists,

$$F(s) := (\mathcal{L}f)(s) = \int_0^\infty f(t)\ e^{-st}\ \mathrm{d}t, \quad s \in \mathbb{C},$$

is called the Laplace *transform* of f.

It holds $(\mathcal{L}\dot{f})(s) = sF(s) - f(0)$ [BSMM08, (15.13)].

A.3.3 Integration Formula

We have [BSMM08, bottom of page 490]

$$\int \frac{1}{(x^2+1)^2}\,\mathrm{d}x = \frac{1}{2}\left(\frac{x}{x^2+1} + \arctan x\right). \tag{A.1}$$

B Proofs

B.1 Proofs for chapter 2

B.1.1 Proofs for subsection 2.2.3

Before we can prove the actual theorem we need the following lemma.

Lemma B.1. *Let* $\mathbf{M} \in \mathbb{C}^{n\times m}$*,* $\mathbf{b} \in \mathbb{C}^m$ *and* $\mathbf{x} := \mathbf{Mb} \in \mathbb{C}^n$*. Then* $\|\mathbf{x}\|_\infty \leq \|\mathbf{M}\|_F \cdot \|\mathbf{b}\|_2$*.*

Proof. For a vector $\mathbf{x}$ we have $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2$ [GVL89, (2.2.6)]. Since $\mathbf{x}$ and $\mathbf{b}$ are vectors, we have $\|\mathbf{x}\|_2 = \|\mathbf{x}\|_F$ and $\|\mathbf{b}\|_F = \|\mathbf{b}\|_2$. Furthermore, we have $\|\mathbf{x}\|_F = \|\mathbf{Mb}\|_F \leq \|\mathbf{M}\|_F \cdot \|\mathbf{b}\|_F$ [Ste98, Theorem 4.11. in Chapter 1]. □

Theorem B.2. *For* $\mathbf{u} \in \mathcal{L}_2(\mathbb{R}_0^+;\mathbb{R}^m)$ *it holds*

$$\begin{aligned} \|\mathbf{y} - \mathbf{y}_r\|_{\mathcal{L}_2} &\leq \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_\infty} \cdot \|\mathbf{u}\|_{\mathcal{L}_2} \quad \text{and} \\ \|\mathbf{y} - \mathbf{y}_r\|_{\mathcal{L}_\infty} &\leq \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_2} \cdot \|\mathbf{u}\|_{\mathcal{L}_2}. \end{aligned}$$

Proof. We essentially follow [ABG10, Section 2.4]. Let $\mathbf{u} \in \mathcal{L}_2(\mathbb{R}_0^+;\mathbb{R}^m)$. We have

$$\widehat{\mathbf{y}}(s) - \widehat{\mathbf{y}}_r(s) = [\mathbf{H}(s) - \mathbf{H}_r(s)] \cdot \widehat{\mathbf{u}}(s). \tag{B.1}$$

Applying Lemma B.1 we get

$$\|\widehat{\mathbf{y}}(s) - \widehat{\mathbf{y}}_r(s)\|_\infty \leq \|\mathbf{H}(s) - \mathbf{H}_r(s)\|_F \cdot \|\widehat{\mathbf{u}}(s)\|_2. \tag{B.2}$$

With the Parseval relation [Wlo82, Satz 1.24] we yield

$$\begin{aligned} \|\mathbf{y} - \mathbf{y}_r\|_{\mathcal{L}_2}^2 &= \int_0^\infty \|\mathbf{y}(t) - \mathbf{y}_r(t)\|_2^2 \,\mathrm{d}t \\ &= \frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\widehat{\mathbf{y}}(s) - \widehat{\mathbf{y}}_r(s)\|_2^2 \,\mathrm{d}s \\ &\overset{\text{(B.1)}}{=} \frac{1}{2\pi} \int_{\imath\mathbb{R}} \|[\mathbf{H}(s) - \mathbf{H}_r(s)] \cdot \widehat{\mathbf{u}}(s)\|_2^2 \,\mathrm{d}s \\ &\leq \frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\mathbf{H}(s) - \mathbf{H}_r(s)\|_2^2 \cdot \|\widehat{\mathbf{u}}(s)\|_2^2 \,\mathrm{d}s \\ &\leq \sup_{s\in\imath\mathbb{R}} \|\mathbf{H}(s) - \mathbf{H}_r(s)\|_2^2 \left(\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\widehat{\mathbf{u}}(s)\|_2^2 \,\mathrm{d}s\right) \\ &= \sup_{s\in\imath\mathbb{R}} \|\mathbf{H}(s) - \mathbf{H}_r(s)\|_2^2 \left(\int_0^\infty \|\mathbf{u}(t)\|_2^2 \,\mathrm{d}t\right) \end{aligned}$$

$$= \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_\infty}^2 \cdot \|\mathbf{u}\|_{\mathcal{L}_2}^2$$

and

$$\begin{aligned}
\|\mathbf{y} - \mathbf{y}_r\|_{\mathcal{L}_\infty} &= \sup_{t\geq 0} \|\mathbf{y}(t) - \mathbf{y}_r(t)\|_\infty \\
&= \sup_{t\geq 0} \left\| \frac{1}{2\pi} \int_{\imath\mathbb{R}} [\widehat{\mathbf{y}}(s) - \widehat{\mathbf{y}}_r(s)] e^{st} \, \mathrm{d}s \right\|_\infty \\
&\leq \frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\widehat{\mathbf{y}}(s) - \widehat{\mathbf{y}}_r(s)\|_\infty \, \mathrm{d}s \\
&\overset{\text{(B.2)}}{\leq} \frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\mathbf{H}(s) - \mathbf{H}_r(s)\|_F \cdot \|\widehat{\mathbf{u}}(s)\|_2 \, \mathrm{d}s \\
&\leq \left(\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\mathbf{H}(s) - \mathbf{H}_r(s)\|_F^2 \, \mathrm{d}s \right)^{1/2} \left(\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\widehat{\mathbf{u}}(s)\|_2^2 \, \mathrm{d}s \right)^{1/2} \\
&= \left(\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\mathbf{H}(s) - \mathbf{H}_r(s)\|_F^2 \, \mathrm{d}s \right)^{1/2} \left(\int_0^\infty \|\mathbf{u}(t)\|_2^2 \, \mathrm{d}t \right)^{1/2} \\
&= \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_2} \cdot \|\mathbf{u}\|_{\mathcal{L}_2}
\end{aligned}$$

using Theorem A.26 with $p = q = 2$ for the last inequality. □

B.1.2 Proofs for subsection 2.3.1

Proof of Lemma 2.17. The number of inner grid points in a grid of resolution n is given by [BG04, (3.31)]

$$N = (2^n - 1)^D. \tag{B.3}$$

For the interpolation error of a function $f \in \overset{0}{\mathcal{W}}{}_p^2([0,1]^D, \mathbb{R})$ we have the following upper bound [BG04, Lemma 3.5.]:

$$\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q} \leq c \cdot |f|_{\mathbf{2},q} \cdot 2^{-2n}, \tag{B.4}$$

where $c = c(D) := D/6^D$.

Due to equation (B.3) we have $N \leq 2^{nD}$. Hence, it holds

$$2^{-n} \leq N^{-1/D}. \tag{B.5}$$

We now obtain

$$\begin{aligned}
\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q} &\overset{\text{(B.4)}}{\leq} c \cdot |f|_{\mathbf{2},q} \cdot 2^{-2n} \\
&\overset{\text{(B.5)}}{\leq} c \cdot |f|_{\mathbf{2},q} \cdot N^{-2/D}.
\end{aligned}$$

This is in agreement with [BG04, Lemma 3.12.], which states that $\varepsilon_q(N) = O(N^{-2/D})$. □

Proof of Lemma 2.18. The number of inner grid points in a grid of resolution n is given by [BG04, Lemma 3.6.]

$$N = \sum_{i=0}^{n-1} 2^i \cdot \binom{D-1+i}{D-1} = 2^n \cdot \left(\frac{n^{D-1}}{(D-1)!} + O(n^{D-2}) \right). \tag{B.6}$$

For the interpolation error of a function $f \in \overset{0}{\mathcal{W}}{}_p^2([0,1]^D, \mathbb{R})$ we have the following upper bound [BG04, Theorem 3.8.]:

$$\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q} \leq c_0 \cdot |f|_{\mathbf{2},q} \cdot 2^{-2n} \cdot A(D,n), \tag{B.7}$$

where $c_0 = c_0(D)$ and

$$A(D,n) := \sum_{k=0}^{D-1} \binom{n+D-1}{k} = \frac{n^{D-1}}{(D-1)!} + O(n^{D-2}). \tag{B.8}$$

Due to equation (B.6) we have $N \leq c_1 \cdot 2^n \cdot n^{D-1}$ with $c_1 = c_1(D) := 1/(D-1)!$. Hence, it holds

$$2^{-n} \leq c_1 \cdot N^{-1} \cdot n^{D-1}. \tag{B.9}$$

Similarly, due to equation (B.8) we have

$$A(D,n) \leq c_1 \cdot n^{D-1}. \tag{B.10}$$

Taking the term for $i = n-1$ of the sum in the definition of N in equation (B.6), we obtain $N \geq 2^{n-1} \geq 2^{n/2}$ for $n \geq 2$. Hence, $\log N / \log 2 \geq n/2$ and so

$$n \leq \log N \cdot 2/\log 2 = c_2 \cdot \log N \tag{B.11}$$

with $c_2 := 2/\log 2$.

We now obtain

$$\begin{aligned}
\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q} &\overset{\text{(B.7)}}{\leq} c_0 \cdot |f|_{\mathbf{2},q} \cdot 2^{-2n} \cdot A(D,n) \\
&\overset{\text{(B.9)}}{\leq} c_0 \cdot |f|_{\mathbf{2},q} \cdot c_1^2 \cdot N^{-2} \cdot n^{2\cdot(D-1)} \cdot A(D,n) \\
&\overset{\text{(B.10)}}{\leq} c_0 \cdot |f|_{\mathbf{2},q} \cdot c_1^2 \cdot N^{-2} \cdot n^{2\cdot(D-1)} \cdot c_1 \cdot n^{D-1} \\
&= c_0 \cdot c_1^3 \cdot |f|_{\mathbf{2},q} \cdot N^{-2} \cdot n^{3\cdot(D-1)} \\
&\overset{\text{(B.11)}}{\leq} c_0 \cdot c_1^3 \cdot |f|_{\mathbf{2},q} \cdot N^{-2} \cdot (c_2 \cdot \log N)^{3\cdot(D-1)} \\
&= c_D \cdot |f|_{\mathbf{2},q} \cdot N^{-2} \cdot (\log N)^{3\cdot(D-1)}
\end{aligned}$$

for $c_D = c_D(D) := c_0(D) \cdot c_1(D)^3 \cdot c_2^{3\cdot(D-1)}$.

This is in agreement with [BG04, Lemma 3.13.], which states that

$$\varepsilon_q(N) = O\big(N^{-2} \cdot (\log_2 N)^{3\cdot(D-1)}\big). \qquad \square$$

Proof of Lemma 2.20. The number of grid points in a grid of resolution n is bounded by [Feu10, Lemma 2.1.2]

$$N \leq 3^D \cdot \sum_{i=0}^{n-1} 2^i \cdot \binom{D-1+i}{D-1} = 3^D \cdot 2^n \cdot \left(\frac{n^{D-1}}{(D-1)!} + O(n^{D-2}) \right). \tag{B.12}$$

For the interpolation error of a function $f \in \mathcal{W}_p^2([0,1]^D, \mathbb{R})$ we have the following upper bound [Feu10, Lemma 2.1.16]:

$$\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q} \leq c_0 \cdot |||f|||_{\mathbf{2},q} \cdot 2^{-2n} \cdot A(D,n), \tag{B.13}$$

where $c_0 = c_0(D)$ and

$$A(D,n) := \sum_{k=0}^{D-1} \binom{n+D-1}{k} = \frac{n^{D-1}}{(D-1)!} + O(n^{D-2}). \tag{B.14}$$

Due to equation (B.12) we have $N \leq c_1 \cdot 2^n \cdot n^{D-1}$ with $c_1 = c_1(D) := 3^D/(D-1)!$. Hence, it holds

$$2^{-n} \leq c_1 \cdot N^{-1} \cdot n^{D-1}. \tag{B.15}$$

Similarly, due to equation (B.14) we have

$$A(D,n) \leq c_1 \cdot n^{D-1}. \tag{B.16}$$

Taking the term for $i = n-1$ of the sum in the definition of N in equation (B.12), we obtain $N \geq 2^{n-1} \geq 2^{n/2}$ for $n \geq 2$. Hence, $\log N / \log 2 \geq n/2$ and so

$$n \leq \log N \cdot 2/\log 2 = c_2 \cdot \log N \tag{B.17}$$

with $c_2 := 2/\log 2$.

We now obtain

$$\begin{aligned}
\|f - \widetilde{f}^{(N)}\|_{\mathcal{L}_q} &\overset{(B.13)}{\leq} c_0 \cdot |||f|||_{\mathbf{2},q} \cdot 2^{-2n} \cdot A(D,n) \\
&\overset{(B.15)}{\leq} c_0 \cdot |||f|||_{\mathbf{2},q} \cdot c_1^2 \cdot N^{-2} \cdot n^{2\cdot(D-1)} \cdot A(D,n) \\
&\overset{(B.16)}{\leq} c_0 \cdot |||f|||_{\mathbf{2},q} \cdot c_1^2 \cdot N^{-2} \cdot n^{2\cdot(D-1)} \cdot c_1 \cdot n^{D-1} \\
&= c_0 \cdot c_1^3 \cdot |||f|||_{\mathbf{2},q} \cdot N^{-2} \cdot n^{3\cdot(D-1)} \\
&\overset{(B.17)}{\leq} c_0 \cdot c_1^3 \cdot |||f|||_{\mathbf{2},q} \cdot N^{-2} \cdot (c_2 \cdot \log N)^{3\cdot(D-1)} \\
&= c_D \cdot |||f|||_{\mathbf{2},q} \cdot N^{-2} \cdot (\log N)^{3\cdot(D-1)}
\end{aligned}$$

for $c_D = c_D(D) := c_0(D) \cdot c_1(D)^3 \cdot c_2^{3\cdot(D-1)}$. □

B.2 Proofs for chapter 3

Before we can prove the actual lemmas we need the following lemma.

Lemma B.3. *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *be symmetric negative definite and let* $\mathbf{E} \in \mathbb{R}^{n\times n}$ *be symmetric positive definite. Then* $\mathbf{E}^{-1/2}\mathbf{A}\mathbf{E}^{-1/2}$ *is symmetric negative definite.*

Proof. The matrix $\mathbf{E}^{-1/2}\mathbf{A}\mathbf{E}^{-1/2}$ is symmetric, and negative definite due to Theorem A.9. □

Lemma B.4. *Let* $\mathbf{A}, \mathbf{E} \in \mathbb{R}^{r\times r}$ *be symmetric negative definite and symmetric positive definite, respectively. Then*

$$\sup_{s\in\imath\mathbb{R}} \|(s\mathbf{E} - \mathbf{A})^{-1}\|_2 = \|\mathbf{A}^{-1}\|_2.$$

Hence, the supremum is attained at $s = 0$*.*

Proof. Define $\mathbf{N} = -\mathbf{A}$, so $\mathbf{N}$ is symmetric positive definite.

Due to Corollary 2.9 the matrix $\mathbf{N}$ has a unique symmetric positive definite matrix square root $\mathbf{N}^{1/2}$.

We write

$$\begin{aligned}
(s\mathbf{E} - \mathbf{A})^{-1} &= (s\mathbf{E} + \mathbf{N})^{-1} \\
&= \mathbf{N}^{-1/2} \cdot \mathbf{N}^{1/2} \cdot (s\mathbf{E} + \mathbf{N})^{-1} \cdot \mathbf{N}^{1/2} \cdot \mathbf{N}^{-1/2} \\
&= \mathbf{N}^{-1/2} \cdot \left(\mathbf{N}^{-1/2}(s\mathbf{E} + \mathbf{N})\mathbf{N}^{-1/2}\right)^{-1} \cdot \mathbf{N}^{-1/2} \\
&= \mathbf{N}^{-1/2} \cdot \left(s\mathbf{N}^{-1/2}\mathbf{E}\mathbf{N}^{-1/2} + \mathbf{I}\right)^{-1} \cdot \mathbf{N}^{-1/2} \\
&= -\mathbf{N}^{-1/2} \cdot (s\mathbf{M} - \mathbf{I})^{-1} \cdot \mathbf{N}^{-1/2}
\end{aligned}$$

for $\mathbf{M} = -\mathbf{N}^{-1/2}\mathbf{E}\mathbf{N}^{-1/2}$. The matrix $\mathbf{M}$ is negative definite due to Lemma B.3 with $\mathbf{E} = \mathbf{N}$ and $\mathbf{A} = -\mathbf{E}$. So we have

$$\begin{aligned}
\sup_{s\in\imath\mathbb{R}} \|(s\mathbf{E} - \mathbf{A})^{-1}\|_2 &= \sup_{s\in\imath\mathbb{R}} \|-\mathbf{N}^{-1/2} \cdot (s\mathbf{M} - \mathbf{I})^{-1} \cdot \mathbf{N}^{-1/2}\|_2 \\
&\leq \sup_{s\in\imath\mathbb{R}} \left\{\|\mathbf{N}^{-1/2}\|_2 \cdot \|(s\mathbf{M} - \mathbf{I})^{-1} \cdot \mathbf{N}^{-1/2}\|_2\right\} \\
&= \|\mathbf{N}^{-1/2}\|_2 \cdot \sup_{s\in\imath\mathbb{R}} \|(s\mathbf{M} - \mathbf{I})^{-1} \cdot \mathbf{N}^{-1/2}\|_2 \\
&\overset{(5.8)}{=} \|\mathbf{N}^{-1/2}\|_2 \cdot \|\mathbf{N}^{-1/2}\|_2 \\
&= \|\mathbf{N}^{-1/2}\|_2^2 \\
&= \|\mathbf{N}^{-1}\|_2 \\
&= \|\mathbf{A}^{-1}\|_2
\end{aligned}$$

using Lemma 5.17 with $\mathbf{A} = \mathbf{M}^{-1}$ and $\mathbf{Z} = \mathbf{N}^{-1/2}$.

For $s = 0$ we have $(s\mathbf{E} - \mathbf{A})^{-1} = \mathbf{A}^{-1}$ and so $\|(s\mathbf{E} - \mathbf{A})^{-1}\|_2 = \|\mathbf{A}^{-1}\|_2$. □

Lemma B.5. *Let $\mathbf{A}, \mathbf{E} \in \mathbb{R}^{r\times r}$ be symmetric negative definite and symmetric positive definite, respectively. Then*

$$\int_{\imath\mathbb{R}} \left\|(s\mathbf{E}-\mathbf{A})^{-1}\right\|_F^2 \;\mathrm{d}s \le \pi \cdot r \cdot \|\mathbf{A}^{-1}\|_2^2 \cdot \|\mathbf{A}\|_2 \cdot \|\mathbf{E}^{-1}\|_2.$$

Proof. Define $\mathbf{N} = -\mathbf{A}$, so $\mathbf{N}$ is symmetric positive definite. Similarly to the proof of Lemma B.4, we write

$$(s\mathbf{E}-\mathbf{A})^{-1} = -\mathbf{N}^{-1/2} \cdot (s\mathbf{M}-\mathbf{I})^{-1} \cdot \mathbf{N}^{-1/2}$$

for the symmetric negative definite matrix $\mathbf{M} = -\mathbf{N}^{-1/2}\mathbf{E}\mathbf{N}^{-1/2}$. So we get

$$\begin{aligned}
\int_{\imath\mathbb{R}} \left\|(s\mathbf{E}-\mathbf{A})^{-1}\right\|_F^2 \;\mathrm{d}s &= \int_{\imath\mathbb{R}} \left\|-\mathbf{N}^{-1/2} \cdot (s\mathbf{M}-\mathbf{I})^{-1} \cdot \mathbf{N}^{-1/2}\right\|_F^2 \;\mathrm{d}s \\
&\le \int_{\imath\mathbb{R}} \left\{\|\mathbf{N}^{-1/2}\|_2^2 \cdot \|(s\mathbf{M}-\mathbf{I})^{-1} \cdot \mathbf{N}^{-1/2}\|_F^2\right\} \;\mathrm{d}s \\
&= \|\mathbf{N}^{-1/2}\|_2^2 \cdot \int_{\imath\mathbb{R}} \|(s\mathbf{M}-\mathbf{I})^{-1} \cdot \mathbf{N}^{-1/2}\|_F^2 \;\mathrm{d}s \\
&= \|\mathbf{N}^{-1}\|_2 \cdot \int_{\imath\mathbb{R}} \|(s\mathbf{M}-\mathbf{I})^{-1} \cdot \mathbf{N}^{-1/2}\|_F^2 \;\mathrm{d}s \\
&\overset{(5.13)}{=} \|\mathbf{N}^{-1}\|_2 \cdot \pi \cdot \|(-\mathbf{M})^{-1/2} \cdot \mathbf{N}^{-1/2}\|_F^2 \\
&\le \|\mathbf{N}^{-1}\|_2 \cdot \pi \cdot \|(-\mathbf{M})^{-1/2}\|_2^2 \cdot \|\mathbf{N}^{-1/2}\|_F^2 \\
&= \pi \cdot \|\mathbf{A}^{-1}\|_2 \cdot \|\mathbf{M}^{-1}\|_2 \cdot \|\mathbf{N}^{-1/2}\|_F^2
\end{aligned}$$

using Lemma 5.18 for the inequalities and equation (5.13) from Lemma 5.21 with $\mathbf{A} = \mathbf{M}^{-1}$ and $\mathbf{Z} = \mathbf{N}^{-1/2}$. We have

$$\begin{aligned}
\|\mathbf{M}^{-1}\|_2 &= \|(-\mathbf{N}^{-1/2}\mathbf{E}\mathbf{N}^{-1/2})^{-1}\|_2 \\
&= \|\mathbf{N}^{1/2}\mathbf{E}^{-1}\mathbf{N}^{1/2}\|_2 \\
&\le \|\mathbf{N}^{1/2}\|_2 \cdot \|\mathbf{E}^{-1}\|_2 \cdot \|\mathbf{N}^{1/2}\|_2 \\
&= \|\mathbf{N}^{1/2}\|_2^2 \cdot \|\mathbf{E}^{-1}\|_2 \\
&= \|\mathbf{N}\|_2 \cdot \|\mathbf{E}^{-1}\|_2 \\
&= \|\mathbf{A}\|_2 \cdot \|\mathbf{E}^{-1}\|_2.
\end{aligned}$$

Plugging this in, we achieve

$$\begin{aligned}
\int_{\imath\mathbb{R}} \left\|(s\mathbf{E}-\mathbf{A})^{-1}\right\|_F^2 \;\mathrm{d}s &\le \pi \cdot \|\mathbf{A}^{-1}\|_2 \cdot \|\mathbf{A}\|_2 \cdot \|\mathbf{E}^{-1}\|_2 \cdot \|\mathbf{N}^{-1/2}\|_F^2 \\
&\overset{(2.5)}{\le} \pi \cdot \|\mathbf{A}^{-1}\|_2 \cdot \|\mathbf{A}\|_2 \cdot \|\mathbf{E}^{-1}\|_2 \cdot r \cdot \|\mathbf{N}^{-1/2}\|_2^2 \\
&= \pi \cdot \|\mathbf{A}^{-1}\|_2 \cdot \|\mathbf{A}\|_2 \cdot \|\mathbf{E}^{-1}\|_2 \cdot r \cdot \|\mathbf{N}^{-1}\|_2 \\
&= \pi \cdot \|\mathbf{A}^{-1}\|_2 \cdot \|\mathbf{A}\|_2 \cdot \|\mathbf{E}^{-1}\|_2 \cdot r \cdot \|\mathbf{A}^{-1}\|_2.
\end{aligned}$$

□

B.3 Proofs for chapter 4

B.3.1 Proofs for subsection 4.3.1

Lemma B.6. *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *be symmetric negative definite. Let* $\mathbf{E} \in \mathbb{R}^{n\times n}$ *be symmetric positive definite and* $\mathbf{E} = \mathbf{R}^T\mathbf{R}$ *its* Cholesky *decomposition (Theorem A.15). Then* $(\mathbf{R}^T)^{-1}\mathbf{A}\mathbf{R}^{-1}$ *is symmetric negative definite.*

Proof. The matrix $(\mathbf{R}^T)^{-1}\mathbf{A}\mathbf{R}^{-1} = (\mathbf{R}^{-1})^T\mathbf{A}\mathbf{R}^{-1}$ is symmetric, and negative definite due to Theorem A.9. □

Lemma B.7. *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *be symmetric negative definite and let* $\mathbf{E} \in \mathbb{R}^{n\times n}$ *be symmetric positive definite. Then* $\Lambda(\mathbf{E}^{-1}\mathbf{A}) \subset \mathbb{R}^-$.
However, $\mathbf{E}^{-1}\mathbf{A}$ *is symmetric if and only if* $\mathbf{A}$ *and* $\mathbf{E}$ *commute.*

Proof. Let $\mathbf{E} = \mathbf{R}^T\mathbf{R}$ be the Cholesky decomposition of $\mathbf{E}$. Then $(\mathbf{R}^T)^{-1}\mathbf{A}\mathbf{R}^{-1}$ is symmetric negative definite due to Lemma B.6. We have

$$\begin{aligned}\mathbf{E}^{-1}\mathbf{A} &= (\mathbf{R}^T\mathbf{R})^{-1}\mathbf{A}\\ &= \mathbf{R}^{-1}(\mathbf{R}^T)^{-1}\mathbf{A}\\ &= \mathbf{R}^{-1}(\mathbf{R}^T)^{-1}\mathbf{A}\mathbf{R}^{-1}\mathbf{R}\end{aligned}$$

So $\Lambda(\mathbf{E}^{-1}\mathbf{A}) = \Lambda((\mathbf{R}^T)^{-1}\mathbf{A}\mathbf{R}^{-1})$ due to Theorem A.7. □

B.3.2 Proofs for subsection 4.3.2

In the next lemma we state a relation between the spectral norms of a symmetric positive definite matrix and its Cholesky factor.

Lemma B.8. *Let* $\mathbf{M}$ *be a symmetric positive definite matrix with the decomposition* $\mathbf{M} = \mathbf{R}^T\mathbf{R}$*, where* $\mathbf{R}$ *is a triangular matrix with positive diagonal entries. Then* $\|\mathbf{M}\|_2 = \|\mathbf{R}\|_2^2$.

Proof. The matrix $\mathbf{R}$ can be split into a diagonal matrix $\mathbf{D}$ with positive entries and a triangular matrix $\mathbf{N}$ with ones on the diagonal: $\mathbf{R} = \mathbf{D}\mathbf{N}$ with $\|\mathbf{N}\|_2 = 1$.

We have $\mathbf{M} = \mathbf{R}^T\mathbf{R} = (\mathbf{D}\mathbf{N})^T(\mathbf{D}\mathbf{N}) = \mathbf{N}^T\mathbf{D}^2\mathbf{N}$, so $\mathbf{D}^2 = \mathbf{N}^{-T}\mathbf{M}\mathbf{N}^{-1}$. It holds

$$\begin{aligned}\|\mathbf{R}\|_2^2 &= \|\mathbf{D}\cdot\mathbf{N}\|_2^2\\ &\le \|\mathbf{D}\|_2^2\cdot\|\mathbf{N}\|_2^2\\ &= \|\mathbf{D}\|_2^2\\ &= \|\mathbf{N}^{-T}\mathbf{M}\mathbf{N}^{-1}\|_2\\ &\le \|\mathbf{M}\|_2\cdot\|\mathbf{N}^{-1}\|_2^2\\ &= \|\mathbf{M}\|_2\end{aligned}$$

since the inverse of $\mathbf{N}$ is a triangular matrix with ones on the diagonal, too.

Furthermore, it holds $\|\mathbf{M}\|_2 = \|\mathbf{R}^T\mathbf{R}\|_2 \le \|\mathbf{R}\|_2^2$. □

The following lemma is needed to rewrite the norm of the matrix logarithm as the logarithm of the matrix norm.

Lemma B.9. *Let* $\mathbf{M} \in \mathbb{R}^{n\times n}$ *be symmetric positive definite. Then*

$$\|\log(\mathbf{M})\|_2 = \max\left\{|\log(\|\mathbf{M}\|_2)|, |\log(\|\mathbf{M}^{-1}\|_2)|\right\}.$$

Proof. Since $\mathbf{M}$ is symmetric, it is diagonalizable due to Theorem A.14. So we have

$$\lambda_i(\log(\mathbf{M})) = \log(\lambda_i(\mathbf{M})) \tag{B.18}$$

due to Theorem 2.4b. Since $\log(\mathbf{M})$ is also symmetric, it holds

$$\begin{aligned}
\|\log(\mathbf{M})\|_2 &= \max_{i=1:\,n} |\lambda_i(\log(\mathbf{M}))| \\
&\overset{(B.18)}{=} \max_{i=1:\,n} |\log(\lambda_i(\mathbf{M}))| \\
&= \max\left\{|\log(\lambda_{\min}(\mathbf{M}))|, |\log(\lambda_{\max}(\mathbf{M}))|\right\} \\
&= \max\left\{|\log(\|\mathbf{M}^{-1}\|_2^{-1})|, |\log(\|\mathbf{M}\|_2)|\right\} \\
&= \max\left\{|-\log(\|\mathbf{M}^{-1}\|_2)|, |\log(\|\mathbf{M}\|_2)|\right\} \\
&= \max\left\{|\log(\|\mathbf{M}^{-1}\|_2)|, |\log(\|\mathbf{M}\|_2)|\right\}.
\end{aligned}$$

□

B.4 Proofs for chapter 5

The matrices $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ are always required to be symmetric negative definite and we define $\mathfrak{A} = \log(-\mathbf{A})$ and $\widetilde{\mathfrak{A}} = \log(-\widetilde{\mathbf{A}})$ as before.

We show the triangle inequality for the two parametric model norms.

Lemma B.10. *It holds*

$$\begin{aligned}
\|\mathbf{H} - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} &\le \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} + \|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)}, \\
\|\mathbf{H} - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)} &\le \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)} + \|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)}.
\end{aligned}$$

Proof. The first inequality follows with the triangle inequality of the $\mathcal{H}_\infty$-norm by

$$\begin{aligned}
&\|\mathbf{H} - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} \\
&= \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{H}(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty} \\
&\le \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \left\{\|\mathbf{H}(\cdot\,;\mathbf{p}) - \mathbf{H}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty} + \|\mathbf{H}_r(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty}\right\} \\
&\le \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{H}(\cdot\,;\mathbf{p}) - \mathbf{H}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty} + \operatorname*{ess\,sup}_{\mathbf{p}\in\mathcal{P}} \|\mathbf{H}_r(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_\infty} \\
&= \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)} + \|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_\infty(\mathcal{P},\mathcal{H}_\infty)}.
\end{aligned}$$

The second inequality follows with the triangle inequality of the $\mathcal{H}_2$-norm and the triangle inequality of the $\mathcal{L}_2$-norm by

$$\begin{aligned}
&\|\mathbf{H} - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)} \\
&= \left(\int_{\mathcal{P}} \|\mathbf{H}(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2}^2 \,\mathrm{d}\mathbf{p} \right)^{1/2} \\
&\le \left(\int_{\mathcal{P}} \left\{ \|\mathbf{H}(\cdot\,;\mathbf{p}) - \mathbf{H}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2} + \|\mathbf{H}_r(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2} \right\}^2 \,\mathrm{d}\mathbf{p} \right)^{1/2} \\
&\le \left(\int_{\mathcal{P}} \|\mathbf{H}(\cdot\,;\mathbf{p}) - \mathbf{H}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2}^2 \,\mathrm{d}\mathbf{p} \right)^{1/2} + \left(\int_{\mathcal{P}} \|\mathbf{H}_r(\cdot\,;\mathbf{p}) - \widetilde{\mathbf{H}}_r(\cdot\,;\mathbf{p})\|_{\mathcal{H}_2}^2 \,\mathrm{d}\mathbf{p} \right)^{1/2} \\
&= \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)} + \|\mathbf{H}_r - \widetilde{\mathbf{H}}_r\|_{\mathcal{L}_2(\mathcal{P},\mathcal{H}_2)}. \qquad \square
\end{aligned}$$

In the rest of the section we prove the additional non-optimal interpolation error bounds.

B.4.1 Technical Lemmas

Lemma B.11. *Let $a > b > 0$ and $z_1, \ldots, z_n \ge 0$. Then*

$$\left(\sum_{i=1}^n z_i^b \right)^{1/b} \le n^{1/b-1/a} \cdot \left(\sum_{i=1}^n z_i^a \right)^{1/a}.$$

Proof. We set $p := a/b > 1$ and $q := a/(a-b) > 1$, so we have $1/p + 1/q = 1$. Applying HÖLDER's inequality (Theorem A.25) with $x_i = z_i^b$ and $y_i = 1$ for $i = 1, \ldots, n$, we get

$$\begin{aligned}
\sum_{i=1}^n z_i^b &\le \left(\sum_{i=1}^n (z_i^b)^p \right)^{1/p} \cdot \left(\sum_{i=1}^n 1^q \right)^{1/q} \\
&= \left(\sum_{i=1}^n z_i^a \right)^{b/a} \cdot n^{(a-b)/a}.
\end{aligned}$$

So

$$\begin{aligned}
\left(\sum_{i=1}^n z_i^b \right)^{1/b} &\le \left(\sum_{i=1}^n z_i^a \right)^{1/a} \cdot n^{(a-b)/(ab)} \\
&= \left(\sum_{i=1}^n z_i^a \right)^{1/a} \cdot n^{1/b-1/a}. \qquad \square
\end{aligned}$$

The next lemma helps to simplify the FROBENIUS norm of the matrix square root.

Lemma B.12. *Let $\mathbf{E} \in \mathbb{R}^{r \times r}$ be symmetric positive definite. Let $p \in [0, 1]$. Then*

$$\|\mathbf{E}^p\|_F \le r^{(1-p)/2} \cdot \|\mathbf{E}\|_F^p.$$

Proof. For $p = 0$ the claim is true, since $\|\mathbf{E}^0\|_F^2 = \|\mathbf{I}\|_F^2 = r$. For $p = 1$ it is trivially true. So now let $p \in (0, 1)$. Let $\lambda_1, \ldots, \lambda_r$ be the eigenvalues of $\mathbf{E}$. Then

$$\begin{aligned}
\|\mathbf{E}^p\|_F^2 &= \|\mathrm{diag}(\lambda_1, \ldots, \lambda_r)^p\|_F^2 \\
&= \sum_{i=1}^{r} \lambda_i^{2p} \\
&= \left\{ \left(\sum_{i=1}^{r} \lambda_i^{2p} \right)^{1/(2p)} \right\}^{2p} \\
&\le \left\{ r^{1/(2p)-1/2} \left(\sum_{i=1}^{r} \lambda_i^{2} \right)^{1/2} \right\}^{2p} \\
&= r^{1-p} \cdot \|\mathrm{diag}(\lambda_1, \ldots, \lambda_r)\|_F^{2p} \\
&= r^{1-p} \cdot \|\mathbf{E}\|_F^{2p}
\end{aligned}$$

using Lemma B.11 with $a = 2$ and $b = 2p < a$ for the inequality. □

Corollary B.13. *It holds*

$$\|(-\mathbf{A})^{-1/2}\|_F \le r^{1/4} \cdot \|\mathbf{A}^{-1}\|_F^{1/2}.$$

Proof. The claim follows from Lemma B.12 with $\mathbf{E} = -\mathbf{A}^{-1}$ and $p = 1/2$. □

For the $\mathcal{H}_\infty$-norm case we need the following lemmas.

Lemma B.14. *Let* $\lambda, \widetilde{\lambda} \in \mathbb{R}^-$. *Then*

$$\int_{\imath\mathbb{R}} \left| \frac{1}{s - \lambda} \right|^2 \cdot \left| \frac{1}{s - \widetilde{\lambda}} \right|^2 \, \mathrm{d}s = \pi \cdot \frac{1}{|\lambda| \cdot |\widetilde{\lambda}|} \cdot \frac{1}{|\lambda| + |\widetilde{\lambda}|}, \tag{B.19}$$

$$\int_{\imath\mathbb{R}} \left| \frac{1}{s - \lambda} \right|^2 \cdot \left| \frac{1}{s/\widetilde{\lambda} - 1} \right|^2 \, \mathrm{d}s = \pi \cdot \frac{|\widetilde{\lambda}|}{|\lambda|} \cdot \frac{1}{|\lambda| + |\widetilde{\lambda}|}. \tag{B.20}$$

Proof. Using the integration formula (5.12) from Lemma 5.20 we have

$$\begin{aligned}
\int_{\imath\mathbb{R}} \left| \frac{1}{s - \lambda} \right|^2 \cdot \left| \frac{1}{s - \widetilde{\lambda}} \right|^2 \, \mathrm{d}s &= \frac{1}{|\lambda|^2 \cdot |\widetilde{\lambda}|^2} \cdot \int_{\imath\mathbb{R}} \left| \frac{1}{s/\lambda - 1} \right|^2 \cdot \left| \frac{1}{s/\widetilde{\lambda} - 1} \right|^2 \, \mathrm{d}s \\
&= \frac{1}{|\lambda|^2 \cdot |\widetilde{\lambda}|^2} \cdot \pi \cdot \frac{|\lambda| \cdot |\widetilde{\lambda}|}{|\lambda| + |\widetilde{\lambda}|} \\
&= \pi \cdot \frac{1}{|\lambda| \cdot |\widetilde{\lambda}|} \cdot \frac{1}{|\lambda| + |\widetilde{\lambda}|}.
\end{aligned}$$

Using the integration formula (5.12) from Lemma 5.20 we have

$$\int_{\imath\mathbb{R}} \left| \frac{1}{s - \lambda} \right|^2 \cdot \left| \frac{1}{s/\widetilde{\lambda} - 1} \right|^2 \, \mathrm{d}s = \frac{1}{|\lambda|^2} \cdot \int_{\imath\mathbb{R}} \left| \frac{1}{s/\lambda - 1} \right|^2 \cdot \left| \frac{1}{s/\widetilde{\lambda} - 1} \right|^2 \, \mathrm{d}s$$

$$= \frac{1}{|\lambda|^2} \cdot \pi \cdot \frac{|\lambda| \cdot |\widetilde{\lambda}|}{|\lambda| + |\widetilde{\lambda}|}$$
$$= \pi \cdot \frac{|\widetilde{\lambda}|}{|\lambda|} \cdot \frac{1}{|\lambda| + |\widetilde{\lambda}|}. \qquad \square$$

Lemma B.15. *It holds*

$$\int_{\imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1}\|_F^4 \, \mathrm{d}s \leq \frac{\pi}{2} \cdot r^2 \cdot \|\mathbf{A}^{-1}\|_2^3.$$

Proof. The matrix $\mathbf{A}$ is symmetric (and negative definite) and hence orthogonally diagonalizable

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \operatorname{diag}(\lambda_1, \ldots, \lambda_r)$$

with eigenvalues $\lambda_i \in \mathbb{R}^-$. Then for $s \in \imath\mathbb{R}$

$$\mathbf{Q}^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{Q} = \operatorname{diag}\left(\frac{1}{s - \lambda_1}, \ldots, \frac{1}{s - \lambda_r}\right)$$

with $(s - \lambda_i)^{-1} \in \mathbb{C}^+$. We have

$$\begin{aligned}
\int_{\imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1}\|_F^4 \, \mathrm{d}s &= \int_{\imath\mathbb{R}} \left\| \mathbf{Q} \cdot \operatorname{diag}\left(\frac{1}{s - \lambda_1}, \ldots, \frac{1}{s - \lambda_r}\right) \cdot \mathbf{Q}^T \right\|_F^4 \, \mathrm{d}s \\
&= \int_{\imath\mathbb{R}} \left\| \operatorname{diag}\left(\frac{1}{s - \lambda_1}, \ldots, \frac{1}{s - \lambda_r}\right) \right\|_F^4 \, \mathrm{d}s \\
&= \int_{\imath\mathbb{R}} \left(\sum_{i=1}^r \left| \frac{1}{s - \lambda_i} \right|^2 \right)^2 \, \mathrm{d}s \\
&= \int_{\imath\mathbb{R}} \left(\sum_{i=1}^r \sum_{j=1}^r \left| \frac{1}{s - \lambda_i} \right|^2 \cdot \left| \frac{1}{s - \lambda_j} \right|^2 \right) \, \mathrm{d}s \\
&= \sum_{i=1}^r \sum_{j=1}^r \left(\int_{\imath\mathbb{R}} \left| \frac{1}{s - \lambda_i} \right|^2 \cdot \left| \frac{1}{s - \lambda_j} \right|^2 \, \mathrm{d}s \right).
\end{aligned}$$

Now we apply equation (B.19) from Lemma B.14 with $\lambda = \lambda_i$ and $\widetilde{\lambda} = \lambda_j$ and obtain

$$\begin{aligned}
\int_{\imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1}\|_F^4 \, \mathrm{d}s &= \sum_{i=1}^r \sum_{j=1}^r \pi \cdot \frac{1}{|\lambda_i| \cdot |\lambda_j|} \cdot \frac{1}{|\lambda_i| + |\lambda_j|} \\
&\leq \pi \cdot \frac{1}{2 \cdot |\lambda_{\max}|} \cdot \sum_{i=1}^r \sum_{j=1}^r \frac{1}{|\lambda_i| \cdot |\lambda_j|} \\
&= \frac{\pi}{2} \cdot \frac{1}{|\lambda_{\max}|} \cdot \left(\sum_{i=1}^r \frac{1}{|\lambda_i|} \right)^2
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\pi}{2} \cdot \|\mathbf{A}^{-1}\|_2 \cdot \left\|(-\operatorname{diag}(\lambda_1, \ldots, \lambda_r))^{-1/2}\right\|_F^4 \\
&= \frac{\pi}{2} \cdot \|\mathbf{A}^{-1}\|_2 \cdot \left\|(-\mathbf{Q}^T\mathbf{A}\mathbf{Q})^{-1/2}\right\|_F^4 \\
&\leq \frac{\pi}{2} \cdot \|\mathbf{A}^{-1}\|_2 \cdot \left(r^{1/4} \cdot \|(\mathbf{Q}^T\mathbf{A}\mathbf{Q})^{-1}\|_F^{1/2}\right)^4 \\
&= \frac{\pi}{2} \cdot r \cdot \|\mathbf{A}^{-1}\|_2 \cdot \|\mathbf{A}^{-1}\|_F^2,
\end{aligned}
$$

where the last inequality is due to Corollary B.13. The claim now follows with $\|\mathbf{A}^{-1}\|_F^2 \leq r \cdot \|\mathbf{A}^{-1}\|_2^2$. □

B.4.2 Bound for $\mathcal{H}_\infty$-norm at Fixed Point

Approach (IIIa): No Assumption via $\mathbf{A} - \widetilde{\mathbf{A}}$

Proposition B.16 (cf. Proposition 5.7)**.** *It holds*

$$
\begin{aligned}
&\sup_{s \in \imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_2 \\
&\leq \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2 \cdot \|\widetilde{\mathbf{A}}^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}\|_2 \cdot \|\mathbf{C}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1).
\end{aligned}
$$

Proof. Plugging in equation (5.3), we obtain

$$
\begin{aligned}
&\sup_{s \in \imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_2 \\
&= \sup_{s \in \imath\mathbb{R}} \|\mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1} \cdot (\mathbf{A} - \widetilde{\mathbf{A}}) \cdot (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \cdot \widetilde{\mathbf{B}}\|_2 \\
&\leq \sup_{s \in \imath\mathbb{R}} \Big(\|\mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1}\|_2 \cdot \|\mathbf{A} - \widetilde{\mathbf{A}}\|_2 \cdot \|(s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \cdot \widetilde{\mathbf{B}}\|_2\Big) \\
&\leq \sup_{s \in \imath\mathbb{R}} \|(s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{C}^T\|_2 \cdot \|\mathbf{A} - \widetilde{\mathbf{A}}\|_2 \cdot \sup_{s \in \imath\mathbb{R}} \|(s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \cdot \widetilde{\mathbf{B}}\|_2.
\end{aligned}
$$

Using equation (5.9) from Lemma 5.17 with $\mathbf{Z} = \mathbf{C}^T$and with $\mathbf{Z} = \widetilde{\mathbf{B}}$, respectively, we get

$$
\begin{aligned}
&\sup_{s \in \imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_2 \\
&= \|\mathbf{A}^{-1} \cdot \mathbf{C}^T\|_2 \cdot \|\widetilde{\mathbf{A}} - \mathbf{A}\|_2 \cdot \|\widetilde{\mathbf{A}}^{-1} \cdot \widetilde{\mathbf{B}}\|_2 \\
&\leq \|\mathbf{C} \cdot \mathbf{A}^{-1}\|_2 \cdot \|\mathbf{A}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1) \cdot \|\widetilde{\mathbf{A}}^{-1} \cdot \widetilde{\mathbf{B}}\|_2 \\
&\leq \|\mathbf{C}\|_2 \cdot \|\mathbf{A}^{-1}\|_2 \cdot \|\mathbf{A}\|_2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1) \cdot \|\widetilde{\mathbf{A}}^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}\|_2
\end{aligned}
$$

using inequality (5.5) from Lemma 5.15 for the first inequality. □

B.4.3 Bound for $\mathcal{H}_2$-norm at Fixed Point

Approach (I): Perturbation of the Inverse

Proposition B.17 (cf. Proposition 5.4)**.** *Let* $\|(s\mathbf{I} - \mathbf{A})^{-1} \cdot \Delta\mathbf{A}\|_F \leq 1/2$. *Then*

$$
\left(\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_F^2 \; \mathrm{d}s\right)^{1/2}
$$

$$\leq r^{3/2} \cdot \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2^{3/2} \cdot \|\widetilde{\mathbf{B}}\|_F \cdot \|\mathbf{C}\|_F \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1).$$

Proof. It holds

$$\begin{aligned}
&\int_{\imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_F^2 \; \mathrm{d}s \\
&\leq \int_{\imath\mathbb{R}} \|\mathbf{C}\|_F^2 \cdot \left\|(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\right\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \; \mathrm{d}s \\
&= \|\mathbf{C}\|_F^2 \cdot \int_{\imath\mathbb{R}} \left\|(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\right\|_F^2 \; \mathrm{d}s \cdot \|\widetilde{\mathbf{B}}\|_F^2.
\end{aligned}$$

Using Corollary A.22 with $\mathbf{M} = s\mathbf{I} - \mathbf{A}$ and $\mathbf{F} = \Delta\mathbf{A}$ and $\|(s\mathbf{I} - \mathbf{A})^{-1} \cdot \Delta\mathbf{A}\|_2 \leq 1/2$ we have

$$\begin{aligned}
\int_{\imath\mathbb{R}} \left\|(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\right\|_F^2 \; \mathrm{d}s &\leq \int_{\imath\mathbb{R}} 4 \cdot \left\|(s\mathbf{I} - \mathbf{A})^{-1}\right\|_F^4 \cdot \|\Delta\mathbf{A}\|_F^2 \; \mathrm{d}s \\
&= 4 \cdot \int_{\imath\mathbb{R}} \left\|(s\mathbf{I} - \mathbf{A})^{-1}\right\|_F^4 \; \mathrm{d}s \cdot \|\Delta\mathbf{A}\|_F^2 \\
&\leq 4 \cdot \frac{\pi}{2} \cdot r^2 \cdot \|\mathbf{A}^{-1}\|_2^3 \cdot \|\Delta\mathbf{A}\|_F^2 \\
&\leq 4 \cdot \frac{\pi}{2} \cdot r^2 \cdot \|\mathbf{A}^{-1}\|_2^3 \cdot r \cdot \|\mathbf{A}\|_2^2 \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1)^2
\end{aligned}$$

using the integral bound from Lemma B.15 for the second inequality and inequality (5.7) from Lemma 5.15 for the last inequality. □

Approach (II): Assumption: $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ commute

Proposition B.18 (cf. Proposition 5.6)**.** *If* $\mathbf{A}$ *and* $\widetilde{\mathbf{A}}$ *commute, then*

$$\begin{aligned}
&\left(\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_F^2 \; \mathrm{d}s\right)^{1/2} \\
&\leq \frac{1}{\sqrt{2}} \cdot \sqrt{\|\mathbf{A}^{-1}\|_2} \cdot \|\widetilde{\mathbf{B}}\|_F \cdot \|\mathbf{C}\|_F \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1).
\end{aligned}$$

Proof. Using Lemma 5.5 we have

$$\begin{aligned}
&\int_{\imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_F^2 \; \mathrm{d}s \\
&= \int_{\imath\mathbb{R}} \left\|\mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1} \cdot (e^{\Delta\mathfrak{A}} - \mathbf{I}) \cdot (s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \cdot \widetilde{\mathbf{B}}\right\|_F^2 \; \mathrm{d}s \\
&\leq \int_{\imath\mathbb{R}} \left(\|\mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1} \cdot (e^{\Delta\mathfrak{A}} - \mathbf{I})\|_F \cdot \|(s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \cdot \widetilde{\mathbf{B}}\|_F\right)^2 \; \mathrm{d}s \\
&\leq \int_{\imath\mathbb{R}} \left(\|\mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1}\|_F \cdot \|e^{\Delta\mathfrak{A}} - \mathbf{I}\|_2 \cdot \|(s\widetilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} \cdot \widetilde{\mathbf{B}}\|_F\right)^2 \; \mathrm{d}s \\
&\leq \int_{\imath\mathbb{R}} \left(\|\mathbf{C}\|_F \cdot \max_{i=1:\,r} \left|\frac{1}{s - \lambda_i}\right| \cdot \|e^{\Delta\mathfrak{A}} - \mathbf{I}\|_2 \cdot \max_{i=1:\,r} \left|\frac{1}{s/\widetilde{\lambda}_i - 1}\right| \cdot \|\widetilde{\mathbf{B}}\|_F\right)^2 \; \mathrm{d}s
\end{aligned}$$

$$= \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \cdot \|e^{\Delta\mathfrak{A}} - \mathbf{I}\|_2^2 \cdot \int_{\imath\mathbb{R}} \left(\left| \frac{1}{s - \lambda_{\max}} \right| \cdot \left| \frac{1}{s/\widetilde{\lambda}_{\min} - 1} \right| \right)^2 \, \mathrm{d}s$$

using Lemma 5.18 for the second inequality and Lemma 5.19 for $f(\mathbf{M}) = (s \cdot \mathbf{I} - \mathbf{M})^{-1}$ with $\mathbf{M} = \mathbf{A}$, $\mathbf{N}^T = \mathbf{C}$ and for $f(\mathbf{M}) = (s \cdot \mathbf{M}^{-1} - \mathbf{I})^{-1}$ with $\mathbf{M} = \widetilde{\mathbf{A}}$, $\mathbf{N} = \widetilde{\mathbf{B}}$ respectively, for the last inequality.

We yield the claimed inequality with $\|e^{\Delta\mathfrak{A}} - \mathbf{I}\|_2^2 \leq (e^{\|\Delta\mathfrak{A}\|_2} - 1)^2$ from Corollary A.24 and, using the integration formula (B.20) from Lemma B.14 with $\lambda = \lambda_{\max}$ and $\widetilde{\lambda} = \widetilde{\lambda}_{\min}$,

$$\begin{aligned} \int_{\imath\mathbb{R}} \left(\left| \frac{1}{s - \lambda_{\max}} \right| \cdot \left| \frac{1}{s/\widetilde{\lambda}_{\min} - 1} \right| \right)^2 \, \mathrm{d}s &= \pi \cdot \frac{|\widetilde{\lambda}_{\min}|}{|\lambda_{\max}|} \cdot \frac{1}{|\lambda_{\max}| + |\widetilde{\lambda}_{\min}|} \\ &\leq \pi \cdot \frac{|\widetilde{\lambda}_{\min}|}{|\lambda_{\max}|} \cdot \frac{1}{|\lambda_{\max}| + |\widetilde{\lambda}_{\min}|} \\ &\leq \pi \cdot \frac{1}{|\lambda_{\max}|} \\ &= \pi \cdot \|\mathbf{A}^{-1}\|_2. \end{aligned}$$

□

Approach (IIIa): No Assumption via $\mathbf{A} - \widetilde{\mathbf{A}}$

Proposition B.19 (cf. Proposition B.16). *It holds*

$$\begin{aligned} &\left(\frac{1}{2\pi} \int_{\imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_F^2 \, \mathrm{d}s \right)^{1/2} \\ &\leq \frac{1}{\sqrt{2}} \cdot \|\mathbf{A}\|_2 \cdot \sqrt{\|\widetilde{\mathbf{A}}^{-1}\|_2} \cdot \|\mathbf{A}^{-1}\|_2 \cdot \|\widetilde{\mathbf{B}}\|_F \cdot \|\mathbf{C}\|_F \cdot (e^{\|\Delta\mathfrak{A}\|_2} - 1). \end{aligned}$$

Proof. First, we bound the integrand. Using equation (5.3) we yield

$$\begin{aligned} &\left\| \mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}} \right\|_F \\ &= \left\| \mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1} \cdot (\mathbf{A} - \widetilde{\mathbf{A}}) \cdot (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \cdot \widetilde{\mathbf{B}} \right\|_F \\ &\leq \left\| \mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1} \cdot (\mathbf{A} - \widetilde{\mathbf{A}}) \right\|_F \cdot \left\| (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \cdot \widetilde{\mathbf{B}} \right\|_F \\ &\leq \left\| \mathbf{C} \cdot (s\mathbf{I} - \mathbf{A})^{-1} \right\|_F \cdot \|\mathbf{A} - \widetilde{\mathbf{A}}\|_2 \cdot \left\| (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \cdot \widetilde{\mathbf{B}} \right\|_F \\ &\leq \|\mathbf{C}^T\|_F \cdot \max_{i=1:r} \left| \frac{1}{s - \lambda_i} \right| \cdot \|\mathbf{A} - \widetilde{\mathbf{A}}\|_2 \cdot \max_{i=1:r} \left| \frac{1}{s - \widetilde{\lambda}_i} \right| \cdot \|\widetilde{\mathbf{B}}\|_F \\ &= \left| \frac{1}{s - \lambda_{\max}} \right| \cdot \left| \frac{1}{s - \widetilde{\lambda}_{\max}} \right| \cdot \|\mathbf{A} - \widetilde{\mathbf{A}}\|_2 \cdot \|\mathbf{C}\|_F \cdot \|\widetilde{\mathbf{B}}\|_F \end{aligned}$$

applying Lemma 5.18 for the second inequality and using Lemma 5.19 for $f(\mathbf{M}) = (s\mathbf{I} - \mathbf{M})^{-1}$ with $\mathbf{M} = \mathbf{A}$ and $\mathbf{N}^T = \mathbf{C}$ and with $\mathbf{M} = \widetilde{\mathbf{A}}$ and $\mathbf{N} = \widetilde{\mathbf{B}}$, respectively, for the last inequality.

Next, we bound the whole integral. We have

$$\begin{aligned}
&\int_{\imath\mathbb{R}} \|\mathbf{C} \cdot \{(s\mathbf{I} - \mathbf{A})^{-1} - (s\mathbf{I} - \widetilde{\mathbf{A}})^{-1}\} \cdot \widetilde{\mathbf{B}}\|_F^2 \,\mathrm{d}s \\
&\leq \int_{\imath\mathbb{R}} \left|\frac{1}{s - \lambda_{\max}}\right|^2 \cdot \left|\frac{1}{s - \widetilde{\lambda}_{\max}}\right|^2 \cdot \|\mathbf{A} - \widetilde{\mathbf{A}}\|_2^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \,\mathrm{d}s \\
&= \|\mathbf{A} - \widetilde{\mathbf{A}}\|_2^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \cdot \int_{\imath\mathbb{R}} \left|\frac{1}{s - \lambda_{\max}}\right|^2 \cdot \left|\frac{1}{s - \widetilde{\lambda}_{\max}}\right|^2 \,\mathrm{d}s \\
&= \|\mathbf{A} - \widetilde{\mathbf{A}}\|_2^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \cdot \pi \cdot \frac{1}{|\lambda_{\max}| \cdot |\widetilde{\lambda}_{\max}|} \cdot \frac{1}{|\lambda_{\max}| + |\widetilde{\lambda}_{\max}|} \\
&\leq \|\mathbf{A}\|_2^2 \cdot (e^{\|\mathfrak{A} - \widetilde{\mathfrak{A}}\|_2} - 1)^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \cdot \pi \cdot \frac{1}{|\lambda_{\max}| \cdot |\widetilde{\lambda}_{\max}|} \cdot \frac{1}{|\lambda_{\max}| + |\widetilde{\lambda}_{\max}|} \\
&\leq \|\mathbf{A}\|_2^2 \cdot (e^{\|\mathfrak{A} - \widetilde{\mathfrak{A}}\|_2} - 1)^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2 \cdot \pi \cdot \frac{1}{|\lambda_{\max}| \cdot |\widetilde{\lambda}_{\max}|} \cdot \frac{1}{|\lambda_{\max}|} \\
&= \pi \cdot \|\widetilde{\mathbf{A}}^{-1}\|_2 \cdot \|\mathbf{A}^{-1}\|_2^2 \cdot \|\mathbf{A}\|_2^2 \cdot (e^{\|\mathfrak{A} - \widetilde{\mathfrak{A}}\|_2} - 1)^2 \cdot \|\mathbf{C}\|_F^2 \cdot \|\widetilde{\mathbf{B}}\|_F^2
\end{aligned}$$

using the integration formula (B.19) from Lemma B.14 and, for the second inequality, inequality (5.5) from Lemma 5.15. □

List of Algorithms

2.1 Iterative Rational Krylov Algorithm. 20

4.1 PMOR for symmetric LTI systems - Offline phase (reduction). . . . 44
4.2 PMOR for symmetric LTI systems - Online phase (interpolation). . . 44

7.1 Symmetric initialization of Iterative Rational Krylov Algorithm. . 94

List of Figures

2.1 Input–state–output model. 13
2.2 The two-dimensional bi-linear basis function $\phi_{(2,1),(1,1)} = \phi_{2,1} \cdot \phi_{1,1}$. . 22
2.3 One-dimensional piece-wise linear bases of level 3 with boundary. . . 23
2.4 Subspaces belonging to grids without boundary for $D = 2$ and $n = 3$. 25
2.5 Subspaces belonging to full grids without boundary for $D = 2$ up to $n = 3$. 25
2.6 Subspaces belonging to sparse grids without boundary for $D = 2$ up to $n = 3$. 25
2.7 Points of full grids without boundary for $D = 2$ up to $n = 4$. 26
2.8 Points of sparse grids without boundary for $D = 2$ up to $n = 4$. . . . 26
2.9 Subspaces belonging to sparse grids with nodal boundary for $D = 2$ up to $n = 3$. 28
2.10 Subspaces belonging to grids with hierarchical boundary for $D = 2$ and $n = 3$. 28

4.1 Flow chart of the method. 45
4.2 Parameter dependence of reduced order basis. 49

5.1 Interpolation scheme for system matrix $\mathbf{A}_r(\mathbf{p})$. 62

7.1 Sparse grid with boundary “Maximum” for $D = 2$. 96
7.2 Sparse grid with boundary “Clenshaw–Curtis” for $D = 2$. 96
7.3 Sparse grid with boundary “Maximum” for $D = 3$. 97
7.4 Sparse grid with boundary “Clenshaw–Curtis” for $D = 3$. 97
7.5 Sparse grid with boundary “Clenshaw–Curtis” for $D = 2$. 97
7.6 Sparse grid with boundary “Chebyshev” for $D = 2$. 97
7.7 The overall and interpolation error for interpolation on full grids and different types of sparse grids. 98
7.8 The overall error for without and with elimination of $\mathbf{E}_r$. 99
7.9 Reduction error for different reduction order r. 99
7.10 The overall error for different reduction orders. 100
7.11 The overall error for MIMO and SISO systems. 100
7.12 The overall error for different dimensions of the parameter space. . . 101
7.13 The overall error over the parameter space—comparison with [PMEL10, GPWL14] for $D = 1$. 103
7.14 The overall error over the parameter space for different numbers of interpolation points and reduction orders. 104
7.15 The overall error for the SiN Membrane benchmark. 105

7.16 The interpolation error and the (scaled) interpolation error bound for different types of grids. 106
7.17 The interpolation error, the interpolation error bound and the bound's three summands. 107
7.18 Both sides of the inequality from Proposition 5.7 for the one-dimensional setting and the one-dimensional setting with $\mathbf{B}_r = \mathbf{C}_r = \mathbf{I}_r$. 108
7.19 Interpolation error and convergence rates. 109

List of Tables

2.1 Interpolation error of interpolation on full and sparse grids in dependence on the number of grid points. 26
2.2 Extrema of CHEBYSHEV polynomials. 30

3.1 Exponential and logarithm mappings for matrix manifolds of interest. 42

4.1 Overview of preservation of requirements for well-defined error measures. 52

7.1 Number of grid points for sparse grids with different boundary resolution. 96
7.2 $\mathcal{L}_\infty(\mathcal{P}, \mathrm{rel}\mathcal{H}_\infty)$-error and offline phase and online phase computing times—comparison with [PMEL10, GPWL14] for $D = 1$. 103
7.3 $\mathcal{L}_\infty(\mathcal{P}, \mathrm{rel}\mathcal{H}_\infty)$-error and offline phase and online phase computing times for different number of interpolation points and reduction orders. 104
7.4 $\mathcal{L}_\infty(\mathcal{P}, \mathrm{rel}\mathcal{H}_\infty)$-error and offline phase computing time of methods in [BBH$^+$15] from [BBH$^+$15, Table 2] and [BBH$^+$15, Table 5]. 105

Bibliography

[ABG10] Athanasios C. Antoulas, Christopher A. Beattie, and Serkan Gugercin. Interpolatory Model Reduction of Large-Scale Dynamical Systems. In Javad Mohammadpour and Karolos M. Grigoriadis, editors, *Efficient Modeling and Control of Large-Scale Systems*, pages 3–58. Springer US, 2010. 15, 17, 18, 19, 52, 121

[AF08] David Amsallem and Charbel Farhat. Interpolation Method for Adapting Reduced-Order Models and Application to Aeroelasticity. *AIAA Journal*, 46(7):1803–1813, 2008. 39, 46

[AF11] David Amsallem and Charbel Farhat. An Online Method for Interpolating Linear Parametric Reduced-Order Models. *SIAM Journal on Scientific Computing*, 33(5):2169–2198, 2011. 1, 3, 4, 39, 40, 43, 46, 47, 48, 51

[AFPA07] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Geometric Means in a Novel Vector Space Structure on Symmetric Positive-Definite Matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347, 2007. 51

[Ant05] Athanasios C. Antoulas. *Approximation of Large-Scale Dynamical Systems.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 2005. 2, 11, 13, 52

[BB09] Ulrike Baur and Peter Benner. Modellreduktion für parametrisierte Probleme durch balanciertes Abschneiden und Interpolation. *at - Automatisierungstechnik*, 57(8):411–419, 2009. 40, 47

[BB15] Angelika Bruns and Peter Benner. 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. 112

[BBBG11] Ulrike Baur, Christopher A. Beattie, Peter Benner, and Serkan Gugercin. Interpolatory Projection Methods for Parameterized Model Reduction. *SIAM Journal on Scientific Computing*, 33(5):2489–2518, 2011. 1, 19, 37, 39, 40, 46, 113

[BBH+15] Ulrike Baur, Peter Benner, Bernard Haasdonk, Christian Himpe, Immanuel Maier, and Mario Ohlberger. Comparison of methods for parametric model order reduction of instationary problems. Preprint MPIMD/15-01, Max Planck Institute Magdeburg, February 2015. Available from `http://www.mpi-magdeburg.mpg.de/preprints/`. 2, 3, 5, 101, 102, 103, 104, 105, 111, 141

[BBH+16] Ulrike Baur, Peter Benner, Bernard Haasdonk, Christian Himpe, Immanuel Martini, and Mario Ohlberger. Comparison of methods for parametric model order reduction of instationary problems. In Peter Benner, Albert Cohen, Mario Ohlberger, and Karen Willcox, editors, *Model Reduction and Approximation: Theory and Algorithms*. SIAM, 2016. To appear. 3, 102

[BBM+06] José M. Badía, Peter Benner, Rafael Mayo, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí, and Afredo Remón. Balanced Truncation Model Reduction of Large and Sparse Generalized Linear Systems. Chemnitz Scientific Computing Preprints 06-04, Fakultät für Mathematik, TU Chemnitz, November 2006. 3, 17

[BF14] Peter Benner and Lihong Feng. A robust algorithm for parametric model order reduction based on implicit moment matching. In Alfio Quarteroni and Gianluigi Rozza, editors, *Reduced Order Methods for Modeling and Computational Reduction*, volume 9 of *MS&A - Modeling, Simulation and Applications*, chapter 6, pages 159–186. Springer, 2014. 39, 46

[BG04] Hans-Joachim Bungartz and Michael Griebel. Sparse grids. *Acta Numerica*, 13:147–269, 2004. 3, 4, 8, 10, 20, 21, 24, 25, 26, 122, 123

[BG09] Christopher A. Beattie and Serkan Gugercin. A trust region method for optimal $\mathcal{H}_2$ model reduction. In *Proceedings of the 48th IEEE Conference on Decision and Control*, pages 5370–5375, December 2009. 20

[BG12] Christopher A. Beattie and Serkan Gugercin. Realization-independent $\mathcal{H}_2$-approximation. In *Proceedings of the 51st IEEE Conference on Decision and Control*, pages 4953–4958, December 2012. 19, 20

[BG15] Peter Benner and Sara Grundel. Model order reduction for a family of linear systems with applications in parametric and uncertain systems. *Applied Mathematics Letters*, 39:1–6, 2015. 49, 113

[BGW15] Peter Benner, Serkan Gugercin, and Karen Willcox. A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems. *SIAM Review*, 57(4):483–531, 2015. 3, 13, 37

[BKVW10] Angelika Bunse-Gerstner, Dorota Kubaliska, Georg Vossen, and Daniel Wilczek. h_2-norm optimal model reduction for large scale discrete dynamical MIMO systems. *Journal of Computational and Applied Mathematics*, 233(5):1202–1216, 2010. 19

[BNR00] Volker Barthelmann, Erich Novak, and Klaus Ritter. High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics*, 12(4):273–288, 2000. 4, 10, 20, 30

[BSMM08] Ilja N. Bronstein, Konstantin A. Semendjajew, Gerhard Musiol, and Heiner Mühlig. *Taschenbuch der Mathematik*. Harri Deutsch, 7th edition, 2008. 118, 119

[DVW10] Joris Degroote, Jan Vierendeels, and Karen Willcox. Interpolation among reduced-order matrices to obtain parameterized models for design, optimization and probabilistic analysis. *International Journal for Numerical Methods in Fluids*, 63(2):207–230, 2010. 39, 46

[ECP+11] Rudy Eid, Rosa Castañé-Selga, Heiko Panzer, Thomas Wolf, and Boris Lohmann. Stability-preserving parametric model reduction by matrix interpolation. *Mathematical and Computer Modelling of Dynamical Systems*, 17(4):319–335, 2011. 40, 46, 47

[FBG12] Garret Flagg, Christopher A. Beattie, and Serkan Gugercin. Convergence of the Iterative Rational Krylov Algorithm. *Systems & Control Letters*, 61(6):688–691, 2012. 19

[Feu10] Christian Feuersänger. *Sparse Grid Methods for Higher Dimensional Approximation*. Dissertation, Institut für Numerische Simulation, Universität Bonn, September 2010. 21, 29, 31, 124

[GAB08] Serkan Gugercin, Athanasios C. Antoulas, and Christopher A. Beattie. $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. *SIAM Journal on Matrix Analysis and Applications*, 30(2):609–638, 2008. 3, 4, 17, 19

[Gar13] Jochen Garcke. Sparse Grids in a Nutshell. In Jochen Garcke and Michael Griebel, editors, *Sparse Grids and Applications*, volume 88 of *Lecture Notes in Computational Science and Engineering*, pages 57–80. Springer, 2013. extended version with python code `http://garcke.ins.uni-bonn.de/research/pub/sparse_grids_nutshell_code.pdf`. 3, 20, 21

[GBP+14] Matthias Geuss, Daniel Butnaru, Benjamin Peherstorfer, Hans-Joachim Bungartz, and Boris Lohmann. Parametric model order reduction by sparse-grid-based interpolation on matrix manifolds for multidimensional parameter spaces. In *Proceedings of the 13th European Control Conference*, pages 2727–2732, June 2014. 3, 40, 46

[GG03] Thomas Gerstner and Michael Griebel. Dimension-adaptive tensor-product quadrature. *Computing*, 71:65–87, 2003. Archives for Scientific Computing. 113

[GG13] Jochen Garcke and Michael Griebel, editors. *Sparse Grids and Applications*, volume 88 of *Lecture Notes in Computational Science and Engineering*. Springer, 1st edition, 2013. 3

[GPWL14] Matthias Geuss, Heiko Panzer, Thomas Wolf, and Boris Lohmann. Stability preservation for parametric model order reduction by matrix interpolation. In *Proceedings of the 13th European Control Conference*, pages 1098–1103, June 2014. 3, 40, 46, 47, 101, 102, 103, 112, 139, 141

[Gri98] Michael Griebel. Adaptive sparse grid multilevel methods for elliptic PDEs based on finite differences. *Computing*, 61:151–179, 1998. 31

[GVL89] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, USA, 2nd edition, 1989. 7, 83, 115, 116, 117, 118, 121

[Hig08] Nicholas J. Higham. *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008. 11, 12, 50

[HJ91] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991. 117, 118

[HJ12] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 2nd edition, 2012. 7

[HP05] Diederich Hinrichsen and Anthony J. Pritchard. *Mathematical Systems Theory I. Modelling, State Space Analysis, Stability and Robustness*, volume 48 of *Texts in Applied Mathematics*. Berlin: Springer, 2005. 15

[Kli07] Andreas Klimke. Sparse Grid Interpolation Toolbox – user's guide. Technical Report IANS report 2007/017, University of Stuttgart, 2007. 95, 96

[KW97] Ladislav Kocis and William J. Whiten. Computational investigations of low-discrepancy sequences. *ACM Transactions on Mathematical Software*, 23(2):266–294, June 1997. 105

[KW05] Andreas Klimke and Barbara Wohlmuth. Computing expensive multivariate functions of fuzzy numbers using sparse grids. *Fuzzy Sets and Systems*, 154(3):432–453, 2005. 95

[Lan69] Peter Lancaster. *Theory of Matrices*. Academic Press, 1969. 115, 116

[Maz14] Marco Mazzucco. A reduced-order matrices fitting scheme with Log-Euclidean metrics for fast approximation of dynamic response of parametric structural systems . *Computer Methods in Applied Mechanics and Engineering*, 269:1–19, 2014. 1, 39, 46, 51

[MC] MORwiki-Community. MOR Wiki - Model Order Reduction Wiki. `http://www.modelreduction.org`. 93, 94

[Moo81] Bruce C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, 26(1):17–32, February 1981. 17

[MS05] Volker Mehrmann and Tatjana Stykel. Balanced Truncation Model Reduction for Large-Scale Systems in Descriptor Form. In Peter Benner, Danny C. Sorensen, and Volker Mehrmann, editors, *Dimension Reduction of Large-Scale Systems*, volume 45 of *Lecture Notes in Computational Science and Engineering*, pages 83–115. Springer Berlin Heidelberg, 2005. 14

[Obe] Oberwolfach Model Reduction Benchmark Collection. Available at `http://portal.uni-freiburg.de/imteksimulation/downloads/benchmark/`. 93

[PMEL10] Heiko Panzer, Jan Mohring, Rudy Eid, and Boris Lohmann. Parametric Model Order Reduction by Matrix Interpolation. *at - Automatisierungstechnik*, 58(8):475–484, 2010. 39, 46, 48, 101, 102, 103, 113, 139, 141

[Rou05] Tomáš Roubíček. *Nonlinear Partial Differential Equations with Applications*, volume 153 of *International Series of Numerical Mathematics*. Birkhäuser Basel, 2005. 118

[RR93] Michael Renardy and Robert C. Rogers. *An Introduction to Partial Differential Equations*, volume 13 of *Texts in Applied Mathematics*. Springer, 1993. 8, 9, 118

[Rud74] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill series in higher mathematics. McGraw-Hill, 2nd edition, 1974. 9

[Smo63] Sergey A. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Doklady Akademii Nauk SSSR*, 148:1042–1045, 1963 [Russian]; English translation in: *Soviet Mathematics Doklady*, 4:240–243, 1963. 3

[Ste98] Gilbert W. Stewart. *Matrix Algorithms*. Society for Industrial and Applied Mathematics, 1998. 115, 116, 117, 121

[VGA08] Paul Van Dooren, Kyle A. Gallivan, and Pierre-Antoine Absil. $\mathcal{H}_2$-optimal model reduction of MIMO systems. *Applied Mathematics Letters*, 21(12):1267–1273, 2008. 19

[Wlo82] Joseph Wloka. *Partielle Differentialgleichungen*. Mathematische Leitfäden. B. G. Teubner, 1982. 121

[Wya12] Sarah Wyatt. *Issues in Interpolatory Model Reduction: Inexact Solves, Second-order Systems and DAEs.* PhD thesis, Virginia Polytechnic Institute and State University, 2012. 112

[Zaa67] Adriaan Cornelis Zaanen. *Integration.* North-Holland, Amsterdam, completely revised edition of "An Introduction to the Theory of Integration", 2nd edition, 1967. 9

[ZCL13] Taishan Zeng, Jian Chen, and Chunyuan Lu. A Tangential Interpolation Algorithm for Optimal H2 Model Reduction with Stability Guarantee. In *Proceedings of the 5th International Conference on Computational and Information Sciences*, pages 955–958, June 2013. 19

[Zen91] Christoph Zenger. Sparse Grids. In Wolfgang Hackbusch, editor, *Parallel Algorithms for Partial Differential Equations*, volume 31 of *Notes on Numerical Fluid Mechanics*, pages 241–251, 1991. 3, 24

Index

affine parameter dependence, **39**
(asymptotically) stable, **14**

Balanced Truncation, 17
basis
  hierarchical, **23**
  nodal, **22**

CHOLESKY decomposition, 44, **117**
closed under conjugation, **19**
columns
  orthonormal, **115**
congruence transformation, **116**
congruent matrices, **116**

de-hierarchization, **31**
decomposition
  CHOLESKY, 44, **117**
  singular value, **116**
  spectral, **116**
diagonal matrix, **115**
diagonalizable matrix, **116**
dimension
  of a grid space, **24**

efficient, **1**
eigenvalues, 116
equivalent systems, **15**, 40, 41, 50
extrema of the CHEBYSHEV polynomials, **30**

FOM, *see* full order model
FROBENIUS norm, **8**, 16
full grid spaces, **24**
full order, **15**
full order model, **15**
  parametric, **35**
function of a matrix, **12**

$\mathcal{H}_2$-norm, **16**, 36
$\mathcal{H}_\infty$-norm, **16**, 36
HALTON sequence, 105
HARDY space, **11**
hat function, **22**
HERMITIAN matrix, **115**
hierarchical basis, **23**
hierarchical difference spaces, **23**
hierarchical increments, **23**
hierarchical surplus, **24**
hierarchization, **31**
higher-dimensional, **3**

IRKA, *see* Iterative Rational KRYLOV Algorithm
Iterative Rational KRYLOV Algorithm, 17, **19**

$\mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$-norm, **36**
$\mathcal{L}_\infty(\mathcal{P}, \mathcal{H}_\infty)$-norm, **36**
LAPLACE transform, 14, **118**
LEBESGUE–BOCHNER spaces, **9**
linear time-invariant (LTI) (continuous-time) system, **13**
lower triangular matrix, **115**
LTI, *see* linear time-invariant

matrices
  congruent, **116**
  similar, **116**
matrix
  diagonal, **115**
  diagonalizable, **116**

HERMITIAN, **115**
lower triangular, **115**
negative definite, **115**
normal, **115**
orthogonal, **115**
p-norm of a, **8**
positive definite, **115**
skew-symmetric, **115**
symmetric, **115**
unitary, **115**
upper triangular, **115**
matrix p-norms, **8**
matrix exponential, **12**
matrix logarithm
principal, **12**
matrix norm, 7
matrix square root
principal, **12**
MIMO, *see* multi-input, multi-output
Model Order Reduction Wiki, 93
moment matching, 18
multi-index, **8**
multi-input, multi-output (MIMO) system, **14**

negative definite matrix, **115**
nodal basis, **22**
norm
FROBENIUS, **8**, 16
$\mathcal{H}_2$, **16**, 36
$\mathcal{H}_\infty$, **16**, 36
$\mathcal{L}_2(\mathcal{P}, \mathcal{H}_2)$, **36**
$\mathcal{L}_\infty(\mathcal{P}, \mathcal{H}_\infty)$, **36**
p-norm of a matrix, **8**
p-norm of a vector, **7**
SCHATTEN p-norm, **7**
spectral, **8**
normal matrix, **115**

Oberwolfach Model Reduction Benchmark Collection, 93
one-sided projection, **16**
order, **14**
orthogonal matrix, **115**
orthogonal Procrustes problem, 41
orthonormal columns, **115**
output equation, **14**

p-norm of a matrix, **8**
p-norm of a vector, **7**
parameter dependence
affine, **39**
parametric full order model, **35**
parametric reduced order model, **36**
poles, **14**
positive definite matrix, **115**
(principal) matrix logarithm, **12**
(principal) matrix square root, **12**
projection
one-sided, **16**
two-sided, **16**
projection framework, **15**

reduced order basis, **16**
reduced order model, **15**
parametric, **36**
reduction order, **15**
ROB, *see* reduced order basis
ROM, *see* reduced order model

SCHATTEN p-norm, **7**
similar matrices, **116**
similarity transformation, **116**
single-input, single-output (SISO) system, **14**
singular value decomposition, 41, **116**
singular values, **116**
SISO, *see* single-input, single-output
skew-symmetric matrix, **115**
SOBOLEV spaces, **10**, 23, 29
SOBOLEV spaces with bounded mixed derivatives, **10**, 21
space of continuous functions, **11**
space of R-times continuously differentiable functions, **11**
spaces
full grid, **24**
sparse grid, **24**
sparse grid spaces, **24**

spectral decomposition, **116**
spectral norm, **8**
stable
  asymptotically, **14**
state equation, **14**
submultiplicativity, **8**
symmetric matrix, **115**
system
  linear time-invariant (LTI) (continuous-time), **13**
  multi-input, multi-output (MIMO), **14**
  single-input, single-output (SISO), **14**
systems
  equivalent, **15**, 40, 41, 50

transfer function, **14**
  of the parametric full order model, **35**
  of the parametric reduced order model, **36**
  of the reduced order model, **15**
trust-region, 20
two-sided projection, **16**

unitary matrix, **115**
upper triangular matrix, **115**

vector
  p-norm of a, **7**