

# Parameterized Model Order Reduction with Applications to Thermal Systems

***Citation for published version (APA):***

Lou, D. (2021). *Parameterized Model Order Reduction with Applications to Thermal Systems*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Electrical Engineering]. Technische Universiteit Eindhoven.

***Document status and date:***

Published: 24/11/2021

***Document Version:***

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

***Please check the document version of this publication:***

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

***General rights***

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

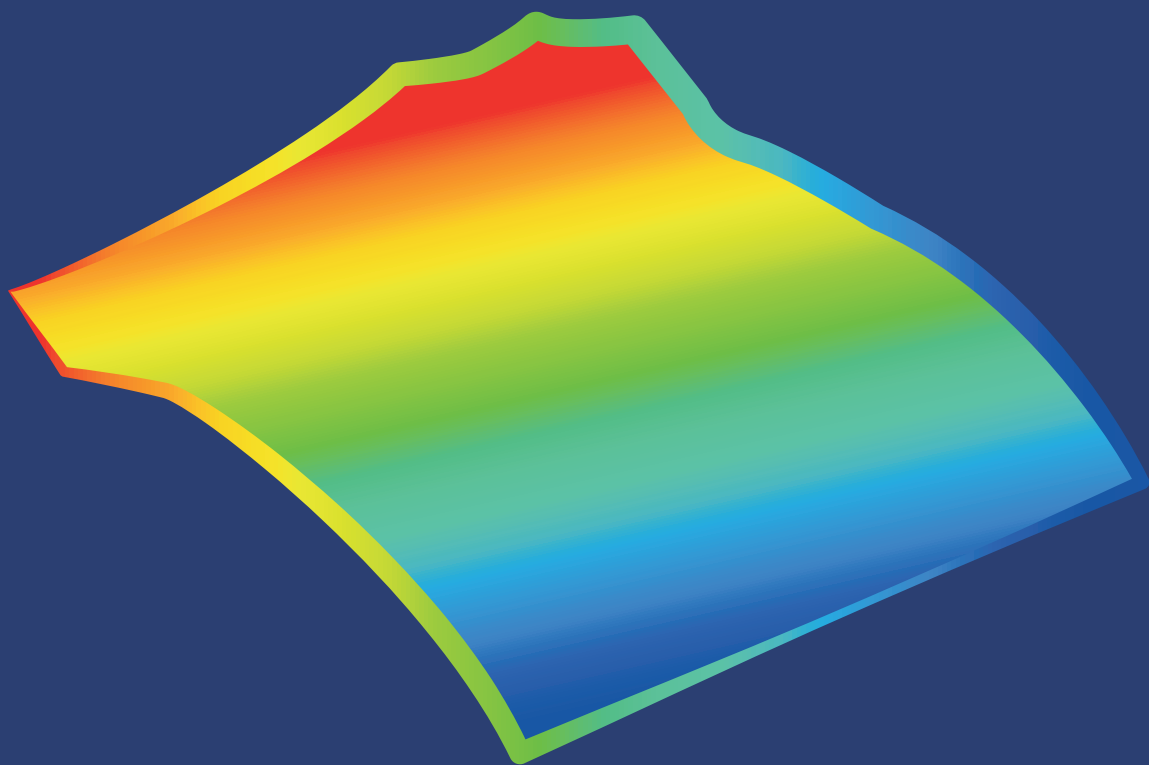
[www.tue.nl/taverne](http://www.tue.nl/taverne)

***Take down policy***

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.



# **Parameterized Model Order Reduction with Applications to Thermal Systems**

**Daming Lou**

# **PARAMETERIZED MODEL ORDER REDUCTION WITH APPLICATIONS TO THERMAL SYSTEMS**

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de  
Technische Universiteit Eindhoven, op gezag van de  
rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een  
commissie aangewezen door het College voor  
Promoties, in het openbaar te verdedigen op  
woensdag 24 november 2021 om 13:30 uur

door

Daming Lou

geboren te Hangzhou, China

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr.ir. P.H.N. de With
1 <sup>e</sup> promotor:	prof.dr. S. Weiland
copromotor:	dr. L. Ozkan
leden:	prof.dr.ir. N. van de Wouw
	prof. dr. A.J. van der Schaft (Rijksuniversiteit Groningen)
	prof. dr.-Ing. J. Lunze (Ruhr University Bochum)
adviseurs:	dr.ir. M. van de Wal (ASML)
	dr. J. M. L. Maubach

Het onderzoek dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening



# Parameterized model order reduction with applications to thermal systems

Daming Lou



ADVANCED THERMAL CONTROL CONSORTIUM

This work is part of the research programme *Advanced Thermal Control Consortium*, which is financed by Thermo Fisher Scientific Inc, ASML Netherlands B.V., IBS Precision Engineering B.V. and Philips Innovation Services.

A catalogue record is available from the Eindhoven University of Technology Library  
ISBN: 978-90-386-5402-7

Copyright © 2021 by Daming Lou.

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilised in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the copyright owner.



# Summary

## **Parameterized model order reduction with applications to thermal systems**

Daming Lou

Many applications in engineering and applied sciences are dedicated to the challenge of how to adjust a system, or parts of a system, so as to improve its functionality. Notably, for a thermal or a thermal relevant system, it often requires a high fidelity dynamical model which, in turn, tends to be a complex mathematical model that requires substantial computational resources for simulation, prediction or model-based control. This complexity raises the question how we can approximate or reduce a large-scale model of a thermal system, deconstruct it in several submodels while preserving its inherent interconnection structure and physical relevance. In this thesis, we look at how model reduction techniques can preserve critical physical parameters for different engineering scenarios while approximating the input-output properties of thermal systems.

Part I of this research considers a moment-matching based method for LPV systems with time-independent parameters. Laplace variables and physical parameters are distinguished such that the real physical interpretation of parameters is preserved in the reduced-order model. Furthermore, we introduce a local error bound to characterize the mismatch induced by the expansion points. When an affine LPV system with a large number of parameters is considered, two parameter reduction techniques are proposed that aim to reduce the dimension of the parameter space. The first one considers the Hankel-norm approximation based on the upper bound of the system Gramians for the parameter space. In the second method, the sensitivity function of the transfer function is employed to truncate the least essential parameters. The problem of parameter calibration using real measurement data is explored. We give a well-defined equivalence between error objective functions using the reduced-order model and the full-order parametric system. Besides, an adaptive strategy is developed which selects optimal parameter expansion points on the basis of measured data. The moment-matching method is extended to uncertain systems, where the nominal part of the system is preserved in the reduced-order model for both LTI and affine LPV systems.

Part II focuses on thermal systems. The concept of equilibria for control systems is compared to equilibria of thermodynamic systems. The stability of a heat conduction system is proved using irreversible entropy as a suitable Lyapunov function. The research topic is concluded by investigating the dissipation properties of thermal systems using the entropy as the thermal energy notion.

Part III of the research presents dedicated tooling and software for applying parameterized model reduction of large-scale systems. The tools are applied to and demonstrated for several large-scale systems that have been provided by industrial partners in this project.



# Abbreviations

---

MOR	model order reduction
pMOR	parameterized model order reduction
ROM	reduced order model
FOM	full order model
POD	proper orthogonal decomposition
PDE	partial differential equation
ODE	ordinary differential equation
FEM	finite element method
LTI	linear time-invariant
LPV	linear parameter-varying
LMI	linear matrix inequalities
SVD	singular value decomposition
MIMO	multi-input multi-output
SISO	single-input single-output
MEMS	microelectromechanical systems
TSCM	transfer-sensitivity covariance matrix
LFT	linear fractional transformations
MVT	mean value theorem
VLSI	very-large-scale integration
paraMOR	parameterized model order reduction



# Contents

---

<b>Summary</b>	<b>v</b>
<b>Abbreviations</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background and motivation . . . . .	1
1.2 Research questions . . . . .	7
1.3 Thesis outline . . . . .	11
1.4 Lists of publications . . . . .	13
<b>I Parameterized model order reduction and applications</b>	<b>15</b>
<b>2 Parameterized model reduction of thermal systems</b>	<b>17</b>
2.1 Introduction . . . . .	17
2.2 Problem description . . . . .	19
2.3 Parameterized model order reduction . . . . .	21
2.4 Error estimation . . . . .	27
2.5 Example: A parametric thermal model . . . . .	31
2.6 Conclusion . . . . .	34
<b>3 Model reduction for LPV systems through parameter projection</b>	<b>35</b>
3.1 Introduction . . . . .	35
3.2 Preliminaries and notation . . . . .	37
3.3 Hankel-norm reduction . . . . .	41
3.4 Simulation results . . . . .	46
3.5 Conclusion . . . . .	50
<b>Appendices</b>	
3.A Proof of upper bound of the Lyapunov function . . . . .	51
3.B Proof of the upper bound of Hankel-singular values . . . . .	52
3.C Proof of upper bound with affine function $f(\theta)$ . . . . .	52

<b>4</b>	<b>Parameter calibration using adaptive parameterized model reduction</b>	<b>53</b>
4.1	Introduction . . . . .	53
4.2	Problem formulation . . . . .	55
4.3	Error equivalent parameter calibration problem . . . . .	58
4.4	Moment-matching for parameter calibration . . . . .	66
4.5	Numerical example: the PSA setup . . . . .	69
4.6	Conclusion . . . . .	73
<b>5</b>	<b>Model reduction of LFT systems using moment-matching method</b>	<b>75</b>
5.1	Introduction . . . . .	75
5.2	Preliminaries . . . . .	77
5.3	Model reduction of LFT systems . . . . .	83
5.4	Study cases . . . . .	90
5.5	Conclusion . . . . .	95
<b>II</b>	<b>Stability and dissipation analysis for thermal systems</b>	<b>99</b>
<b>6</b>	<b>Thermodynamic equilibrium and stability analysis of thermal systems</b>	<b>101</b>
6.1	Thermodynamics fundamentals . . . . .	103
6.2	Stability of thermodynamic systems . . . . .	107
6.3	Numerical example . . . . .	115
6.4	Conclusion . . . . .	117
<b>Appendices</b>		
6.A	Derivation of entropy balance equation . . . . .	117
<b>7</b>	<b>Dissipative properties of thermal systems</b>	<b>119</b>
7.1	Introduction . . . . .	119
7.2	Fundamentals . . . . .	121
7.3	Dissipative properties of thermal systems . . . . .	125
7.4	Dissipative properties of interconnected systems . . . . .	127
7.5	Numerical example: 2-D heat conduction . . . . .	132
7.6	Conclusion . . . . .	135



<b>III Tooling of parameterized model order reduction</b>	<b>137</b>
<b>8 Tooling of parameterized model order reduction for large-scale systems</b>	<b>139</b>
8.1 Overview of paraMOR tooling . . . . .	140
8.2 Structure of the paraMOR . . . . .	142
8.3 Programming principles . . . . .	142
8.4 Case studies . . . . .	143
8.5 Conclusion . . . . .	148
<b>9 Conclusions and future research directions</b>	<b>149</b>
9.1 Conclusions . . . . .	149
9.2 Future research directions . . . . .	150
<b>Bibliography</b>	<b>153</b>
<b>Acknowledgments</b>	<b>167</b>
<b>Curriculum Vitae</b>	<b>169</b>



# Introduction

---

Numerical simulation of dynamical systems has become an indispensable tool for studying and reproducing complex physical phenomena. This tool has thoroughly transformed and driven the development and research in science and engineering. Still considerable computational resources are often required either due to the inherent complexity and large-scale nature or due to the ultra-precision required. One approach to overcome this challenge is through model reduction; the principle is to construct a lower dimensional model that accurately represents the dynamic response of the original system while demanding less computation time. For linear and non-parametric dynamical systems, the development of model reduction techniques has reached a level of maturity. However, a further challenge arises in the question of how to achieve model reduction of a system whose dynamical behaviour depends on a set of parameters. This is what motivates extension of model order reduction techniques to parametric model order reduction where the purpose is to retain physical parameters in the reduced model. In particular, for systems that involve thermal phenomena and thermodynamics this is vital to the system performance. Thus, the main thread of this research follows the question: how can we efficiently and effectively approximate a large-scale model of thermal systems while the physical parameters are kept in the reduced model? Furthermore, for thermal systems which are composed of coupled subsystems, we investigate the equilibrium, stability and dissipation behaviours of these thermal systems.

---

## 1.1 Background and motivation

Especially in the high-tech sector where the development of multi-physical systems plays a key role, thermal effects have been identified as one of the dominant

factors affecting the performance [41, 151]. Systems that involve such thermal effects are ranging from transmission electron microscopes, 3D metrology machines, to lithography systems as depicted in Figure 1.1. Among these applications, there is a persistent trend that future developments and technological innovations lead to systems that are more sophisticated, more interconnected, more complex. In particular, their functioning and the understanding of their technical properties require a deeper understanding of how multiple physical domains are (or need to) integrate into technical or engineered designs. For the high-tech industry and depending on the specific application, this often requires a combined and integrated understanding of electrical, mechanical, thermodynamic, chemical, optical behaviour. This has already led to more optimized accuracy, productivity, and performance, and is a persistent challenge for more technological and scientific innovation.



**Figure 1.1:** High-precision systems: Transmission electron microscope with ultra resolution up to 60 pm (a) and the latest lithography machine which the resolution can be achieved to 13 nm (c). The 3D metrology system for measuring complex surfaces and free form optics with resolution of 1.6 nm (b).

One typical example of thermal effects in high precision systems is their thermo-mechanical behaviour. For systems that involve such phenomena, relatively small temperature fluctuations will result in small deflections and deformations of materials, leading to mechanical deviations and therefore thermally induced position errors. Conversely, mechanical deformations may lead to increased mechanical stress in materials and corresponding variations in temperature. Although

these errors are often negligible, for a high precision system that requires nanometer or even subnanometer scale accuracy, they may have a significant impact on the performance. The system design, control synthesis and observer design to compensate such errors are challenging due to its inherent complexity and its multi-physical aspect. This is even more so when the goal is to attenuate not only the thermally induced errors in its steady-state but also its transient error dynamics. On top of the difficulties mentioned above, other challenges arise in experiment design, measurement data acquisition and system simulation for such high precision systems. For instance, the thermally induced error is seldom directly accessible or measurable through sensors, usually have slower time scales than the time scale at which the system operates, the influence of thermal effects in systems is hard to isolate in experiments where actuation through heating and cooling requires accessibility to systems, and if it is possible it commonly requires ultra-precision sensors to observe [109]. Besides, in real applications, there are a limited number of sensors that can be installed in the system. The challenge is how to smartly choose the locations such that the deformations can be measured with a limited amount of sensors.

### 1.1.a Advanced Thermal Control Consortium

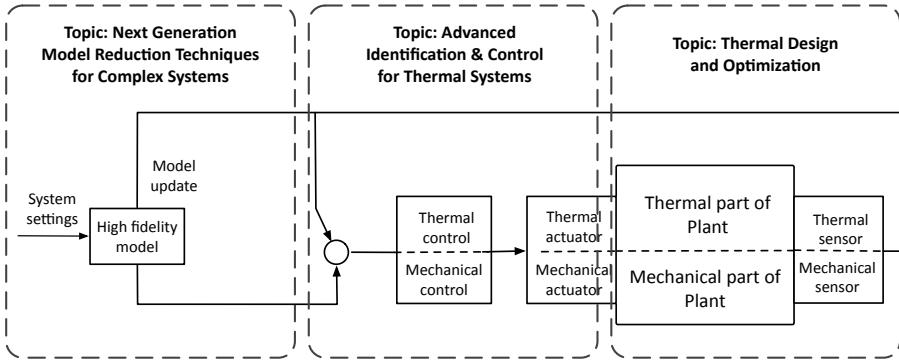
Apart from the academic interest in this topic, the need for improved thermal control has been echoed in industry. To achieve new levels of precision in production processes, the active control of thermal effects is required [178]. In 2013, a group of leading industrial partners and academia (TU Eindhoven & TU Delft) from the Dutch high tech sector initiated a cooperation project: *Advanced Thermal Control Consortium* (ATC) [44]. This consortium focuses on technical strategies and solutions to enable the next generation of industry capability. The ATC consortium officially started in late 2015, and three PhD projects were funded each covering a part of the research into high precision thermo-mechanical systems (c.f. Figure 1.2). This included

1. Thermal Design and Topology Optimization,
2. Advanced Identification & Control for Thermal Systems,
3. Next Generation Model Reduction Techniques for Complex Systems.

The first project is about the topology optimization at the design stage. Different from the first-principle based modelling, the second research topic is using the data-driven method to construct the model for control synthesis of thermal systems. The last one, elucidated in this thesis, focuses on the reduction of high precision thermal system models while preserving physical relevance.

### 1.1.b Motivation and Focus

The next generation tools are needed to develop high precision machines for which experimentation of physical systems can be inaccessible or is too expensive to perform. To enable studying the underlying physical behaviour of these machines,



**Figure 1.2:** Block diagram overview of a high precision thermo-mechanical systems, showing the three research topics of the ATC.

a new generation of tools for coupling the corresponding multi-physical models and for the efficient numerical analysis and simulation of the resulting models are needed.

In real life, a high precision machine consists of thousands of interconnected or coupled subsystems. Generally only few subsystems of the machine, i.e., those that are most critical to the system performance, need a design upgrade or a deeper understanding of their physical properties to enhance or improve their engineered behaviour. For these subsystems, spatial discretization methods (e.g., finite element method) often yield a large-scale mathematical model due to the high precision requirement and the complex physical nature. In addition, for performance enhancement, these models frequently need to be extended to incorporate physical behaviour (such as thermal properties) to broaden their scope. Invariably, this leads to more complexity in the model. Complexity as indicated by the number of finite elements, by the number of first order differential equations and complexity in terms of additional physical laws that are incorporated in the model. Consequences of increased complexity consist of slower simulations, more complicated (or impossible) model-based designs, slower monitoring, the infeasibility of real-time and online computations, etc. One way to remedy the increase of complexity amounts to simplify these models and substitute complex models by simpler ones that exclusively focus on the intended use of the model. This is particularly relevant in circumstances where the reduction of a full, complete model is not desired or not possible in terms of computation time or memory requirement. Furthermore, the correct coupling will enable a modular structure in which individual reduced order models for elementary parts are useful in a repetitive structure in which the coupling of multiple reduced order model elements constitutes the complete system. But how to correctly couple a large number of models such as to result in a reliable and thermal relevant model is a challenge.

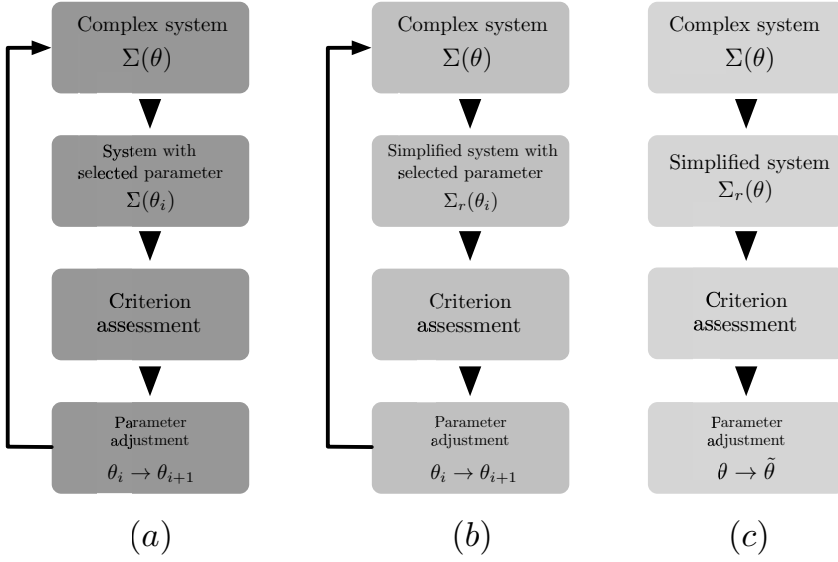
With the increasing need for higher fidelity, system simulations require to include more details leading to a large-scale and complex mathematical model. Moreover, during the design phase, this mathematical model regularly needs to be iterated, leading to multiple realizations for all possible scenarios to find the *best* perform-

ance. The large-scale model together with its parameter values, inputs, initial conditions and boundary conditions, define the simulation and are the basis for optimization, calibration, design decisions, controller synthesis, etc. Furthermore, storing such large-scale model and data poses another challenge for the memory space. Indeed, additional challenges that may be solved by deploying more computational resources: by using more powerful processors, adding more memory storages and exploiting parallelism. However, this is not always realistic or feasible, since the unlimited matching of computational resources is not possible, and discards possibilities and challenges for more intelligent solutions to meet computational demands. Moreover, it can be problematic for a system that needs to operate real-time whilst the computational resource is limited.

Model order reduction techniques have been considered as one of the few efficient and effective means to alleviate the computational burden. The primary principle is to construct a low dimensional and efficient model that still approximates the original full-order model with high fidelity. In such a way, the original full-order model is substituted by the reduced model, so as to enable fast computation for system simulation, design optimization, calibration, prediction, etc. Many successful applications ranging from very-large-scale integration (VLSI) [136], aircraft design [166] to computational fluid dynamics [48] have been based on the implementation of model reduction order techniques. However, despite the advancement of model reduction order techniques for linear time-invariant (LTI) systems, there is still a lack of efficient model reduction order (MOR) for approximating parameter dependent systems where the system structure and physical interpretations of parameters can be preserved in the reduced order model. The latter is critical to enable efficient design iterations.

A class of models explicitly incorporating parameter variation can be represented by parameter-dependent systems. More precisely, a parameter-dependent system denotes a collection of LTI descriptions where each LTI description characterizes a specific instance of the parameter values for the parameter-dependent system. A non-parametric reduced model can approximate the original system well around the frozen parameters with small variations. Instead, if parameter changes are more substantial or cover larger ranges, then the reduced model no longer suffices and dynamics may change substantially with parametric changes of the original model. Then, subsequently, it necessitates the re-generation of a new reduced model. Although the reduced order model has the computational advantage over the full order model, the repeated generation of the reduced model is still computationally costly, in particular for applications where parameters are changed for every iteration.

A typical example is parameter calibration for a complex system  $\Sigma(\theta)$ , in the parameter vector  $\theta$  which can be any realization of the system. In practice, the model built by engineers has some mismatch with the actual system, e.g., the numerical value of parameters are inconsistent with the *true* ones, say  $\theta_{\text{true}}$  in the real physical system. This can be due to production variation or different operating conditions. A classical scheme of parameter calibration starts with performing a high-quality measurement for which the data is assumed to reflect the *true* behaviour of the dynamical system. Then a parameter  $\theta_i$  of this dynamical system is



**Figure 1.3:** The classical scheme for parameter calibration (a), the improved scheme using (non-parameterized) simplified model (b) and the proposed scheme with parameterized simplified model (c). With a parameter-dependent simplified model, the computational burden of the scheme at the right is reduced in comparison to the classical scheme at the left and the existing one in the middle.

selected for calibration, in the  $i$ -th iteration ( $i = 1, 2, \dots$ ) and the parameter is adjusted  $\theta_i \rightarrow \theta_{i+1}$  when the criterion assessment with the model is close enough to the measurement data. This iteration stops when  $\theta_i$  is close enough to  $\theta_{\text{true}}$ . As the number of parameters increases and higher accuracy requirements are imposed, the process can be tedious and time-consuming. This scheme is improved if a simplified model  $\Sigma_r(\theta_i)$  approximates the dynamics of the original system at the sampling point  $\theta_i$ . In such a way, the calibration can be carried out with the simplified model instead of the full-order model. This improved scheme has been implemented in many physical systems and applications [91, 88, 149]. However, this scheme is time-consuming if there are many parameters for calibration and each parameter requires several iterations. The proposed scheme is that if a simplified model  $\Sigma_r(\theta)$  which not only captures the dynamics of the original system  $\Sigma(\theta)$  but also preserves the parameter variation in the simplified model, such that the calibration can be performed without repeating the criterion assessment on the full order model for every parameter. The classical scheme, the improved one and the proposed one for parameter calibration are illustrated in Figure 1.3. The challenge described in this example is also observed in many other applications where the model is of high complexity, and the assessment requires to iterate for specific parameters with the model, for example in process engineering, the identifiability of a large-scale chemical model is crucial for system analysis and the preservation of the identifiability remains an open question.

Another example is in uncertain systems where the computational challenge raised



by parameter variations is observed. Besides the parametric uncertainties described above, the mismatch between the model and the real physical system can contain unmodelled dynamics or nonlinearities. This type of mismatch can be represented as uncertain variables. For such a large-scale uncertain system, the system analysis, controller synthesis and model based risk assessment which have to iterate all possible scenarios are computationally costly. Therefore, parametrized model order reduction (pMOR) is needed to speed up the computation while it has remained an open question for large-scale uncertain systems.

## 1.2 Research questions

To sum up, motivated by the industrial need for the further development of high precision systems, fast and reliable computation for numerical simulation, design optimization, parameter calibration and system analysis is needed. In this thesis, we investigate the physically relevant model reduction techniques for large-scale and complex systems. Specifically, we focus on the system that is composed of several subsystems, and these subsystems are coupled with thermal phenomena.

On these premises we raise the research question:

How can we efficiently approximate large-scale models of thermal systems composed out of several submodels while preserving its inherent interconnection structure and physical relevance?

With this research question, we can split it into three subquestions and attempt to answer each of them:

- a). Can we develop efficient model approximation techniques for large-scale systems in which distinguished physical parameters retain their physical meanings? More precisely, for a given parameter-dependent system, we aim to develop a **pMOR** method such that the parameters in the parametrized reduced model remain the same interpretations as in the original parameter-dependent model.
- b). Can we establish both thermally relevant and correct coupling and decoupling of modules of systems for composition and decomposition of models? To do so, modules need to be relevant from a thermodynamic point of view. The objective is to find a **thermal coupling/decoupling** strategy for two thermal models such that the coupled/decoupled system is also a thermal relevant system in the sense that the first and the second laws of thermodynamics are satisfied.
- c). Can we develop an efficient numerical tool which allows performing parameterized model order reduction for an industry-approved application? We focus on the implementation of parameterized model order reduction method

in two aspects: the complexity of the algorithm and the memory requirement for the data. Particularly, the **pMOR tool** needs to be operated independently and efficiently with a regular PC. The goal of this tool is to enable the parameterized reduction for an industrial relevant application as described in Figure 1.3, and the computational time of the whole procedure is expected to be shortened by a factor of 5 or more. The configuration of the system: state number  $\approx 5 \times 10^5$ , multiple-input multiple-output (MIMO) and parameter dimension  $\geq 20$ .

### 1.2.a The state of art and subquestion on pMOR

In the literature, pMOR techniques can be roughly classified into two main categories: local basis approaches and global basis approaches [29]. The scheme in the first category uses interpolation methods, including manifold interpolation [7, 47, 38, 6, 9, 43, 154, 179, 40], transfer function interpolation [23, 22] and matrix interpolation [56, 112, 121, 138, 130, 8]. The local basis approaches rely on a set of non-parametric models  $\{\mathbf{G}_1, \dots, \mathbf{G}_\ell\}$  which is the parameter-dependent system  $\mathbf{G}(\theta)$  for every parameter vector  $\{\theta_1, \dots, \theta_\ell\}$  of the parameter space  $\theta_i \in \Theta$ . For each  $i = 1, \dots, \ell$  a pair of projection matrices  $\{V_i, W_i\}$  of dimension  $n \times r$  is generated via any projection-based method [21]. Then, with this local information  $\{V_i, W_i\}_{i=1}^\ell$ , a parameterized reduced order model can be constructed and the dimension is  $r$  for the reduced model and  $n, n \gg r$  for the original model, respectively. A collection of reduced order models  $\{\hat{G}_1, \dots, \hat{G}_\ell\}$  is determined which are subsequently interpolated over the parameter space  $\Theta$  through the points  $\theta_1, \dots, \theta_\ell$ , leading to an interpolated reduced order model  $\hat{G}(\theta)$ .

The advantage of this approach is that it imposes less requirements for the parameter dependency in a sense that it can be non-affine, nonlinear or even non-analytical in the parameter space as long as the local system realization is available. However, as multiple realizations are necessary for capturing the local information, this is memory demanding for large-scale models. This is even more so when the number of parameters increases. Additionally, the computational complexity of the manifold interpolation methods commonly requires  $\mathcal{O}(nr)$  for the online operation of the reduced model, one solution which overcomes this problem requests a conservative condition of the system matrices [40]. In applications using transfer function interpolation method, manipulated poles will be introduced for the parametric transfer.

For the second category, the global basis approaches consider a common basis that is constructed by a pair of projection matrices  $\{V, W\}$ , where  $\{V, W\}$  of dimension  $n \times r$  are directly computed from the parameter-dependent system  $\mathbf{G}(\theta)$  and the matrices contain parameter information regarding the parameter space. The scheme of global basis approaches avoids most of the problems encountered in local basis approaches, but it requires parameter affine structure for the state space realization. In the case where the parametric dependency is non-affine, the system matrices can be approximated into parameter affine structure using proper methods, e.g., discrete empirical interpolation method (DEIM) [52, 123, 134, 35]. The global approaches include parametric balanced model order reduction method

[173, 155, 156], parametric proper orthogonal decomposition (POD) [42, 132], and reduced basis method [20, 80, 94, 99, 97, 98]. However, these methods have some limitations regarding various aspects. For a model with high state dimension  $\geq 10^4$ , the parametric balanced model order reduction method does not scale well with the dimension of the original model and the number of parameters since it needs to solve the Lyapunov equation for every parameter. In the case that the system will be exited with various inputs, the parametric POD method commonly cannot provide a robust high fidelity reduced model or needs to generate a large number of snapshots and leads to a high dimensional projection matrices, since the projection matrices are dependent on inputs. The reduced basis methods are successful in solving parametrized partial differential equations (PDEs) problems which provide more information than the parameterized ordinary differential equations (ODEs), especially boundary conditions. In this thesis, we assume that the PDEs are not provided in the default setting where simulation software is used for modelling and the spatial discretization is performed.

Another well known global basis approach is multivariable Padé approximations, or also called moment-matching based methods [37, 53, 64, 69, 28, 177, 71, 87, 86, 170]. The idea is to construct a projection space where the (mixed) *moments* of the original system are collected. In such a way, the reduced model generated through this projection space can incorporate the parameter information. Here the *moment* can be interpreted as the coefficient of the Taylor series of a parametric transfer function. Based on this idea, a number of methods and algorithms are developed which differ in the way the moments are computed, the projection space is constructed and the number of moments are matched. Approaches based on implicit moment-matching provides more numerical stability in comparison to explicit approach. The two-sided method can provide twice coincided matched moments in the reduced model. Even though these moment-matching based methods can provide numerically efficient procedures to generate parametric reduced order models that well approximate the dynamical behaviours, parameter interpretations in the reduced model are not consistent with the ones in the original model. Since parts of parameters are not taken into account in the expansion points directly, the resultant moments will mix the frequency expansion points and physical parameter expansion points and the parametric reduced model cannot be interpreted as the same physical parameter dependent system as the original one. Such models are not preferable for subsequent operations, e.g., optimization or parameter calibration cannot be performed on the original parameter vector. Thus, for a large-scale affine LPV system, it is necessary to develop a physical relevant pMOR method such that parameters are preserved in the reduced model and the physical interpretation is consistent with the ones in original model.

Apart from the parameter-dependent systems  $G(\theta)$  mentioned above, another class of systems that demands repeated evaluation of the model over a space are uncertain systems, where the stability analysis or control synthesis needs to evaluate every point in the uncertainty set. In the literature, we observe that all existing work on model order reduction of uncertain systems is based on balanced truncations [26, 168, 11, 25] which is not suitable for large-scale uncertain systems.

### 1.2.b The state of art and subquestion on thermal coupling

As we discussed in the above, the thermal effects are critical for high-precision system performance. Regularly, these systems consist of a large number of components and multi-physics interactions. Although many energy-based methods and tools, e.g., [163, 5], have been developed for modelling and control synthesis of such multi-physics systems, a rigorous integration of the second law of thermodynamics is often lacking. The entropy-based methods [54, 63] are developed using entropy-balance equation, as it permits to include the irreversible process. However, the proposed entropy function focuses on the dynamics of a closed system, and the stability or dissipation criterion for composite systems is not considered. In the field of control systems, conservation laws and energy functions are often used to prove stability of equilibria of systems, to analyze their behavior, and for purposes of control and optimization. The role of Lyapunov functions in autonomous systems, together with their generalizations in the direction of open dynamical systems, passive and dissipative dynamical systems, bond-graph theory and port-Hamiltonian systems are fundamental in modern systems theory [101, 171, 172]. The essence of these methodologies lies in the observation that the efficiency of many physical processes is limited by their energy storage and the amount of energy that has been supplied by its environment. Indeed, Carnot's principle claims the limited efficiency of heat engines; in Hamiltonian and Lagrangian mechanics, the recoverable energy is always bounded by the Hamiltonian or Lagrangian function; in an ideal reversible thermodynamical process, all heat can be converted into work. In the last decades, non-equilibrium thermodynamics has been developed as a research field, aimed at describing physics of thermal processes beyond and away from thermodynamic equilibrium. In particular, it aims at incorporating the time-course of intensive variables such as temperature and pressure and to generalize the concept of entropy to thermal states that are not in equilibrium. Starting with Onsager's reciprocal relations [126, 127] in 1931, important contributions on non-equilibrium thermodynamics have been made by Denbigh [61] on steady state principles, De Groot's work on linear irreversible processes [55, 54] and Prigogine's minimum entropy production principle [139] that has been extended to the theory of entropy generation minimization by Bejan [27]. This theory claims that the maximum efficiency of a thermodynamic system is achieved while the dissipated energy is minimal.

In the literature, we also observe that quadratic Lyapunov functions have been widely used to provide insight in the qualitative behavior of systems. Although in [59], a quadratic function of temperature is proposed to satisfy Lyapunov stability criteria, these functions do not naturally inherit the energy attributes of thermodynamic systems. Another well known Lyapunov candidate representing entropy generation such as in [165], is the product of the thermodynamic force and flux, which leads to a temperature dependent coefficient for Fourier's law. This formulation does not fit in with the conventional modelling framework in engineering where the heat conduction coefficient is chosen as temperature independent constant. Thus, for thermal and thermo-relevant systems, it is essential to seek for a physical meaningful Lyapunov functions that prove useful for both composition and decomposition properties of thermal systems that consist of interconnected

components. This part of work is given in Chapter 6 and Chapter 7.

### 1.2.c The state of art and subquestion on pMOR tools

For an easy use of pMOR techniques in applications, software tools are necessary. The challenges of model reduction implementation for large-scale systems mainly are from two aspects: storage and computation complexity. For a model with state number larger than  $\geq 10^4$ , a simple operation such as importing a system matrix to MATLAB already leads to high storage requirements. To overcome this difficulty, we may exploit the sparsity of the system representation which is commonly seen in very-large-scale integration (VLSI) models and thermal models. For the second challenge, some operators for a large-scale matrix such as inverse operator and eigenvalue operator often are computational expensive or even prohibitive. Instead of directly performing such operators one can rely on some matrix factorization techniques.

For the purpose of model reduction, there are several toolboxes including MESS (Matrix Equations and Sparse Solvers) [144], DPA (Dominant Pole Algorithm)[141], Emgr (Empirical Gramian Framework) [93], MORLAB (Model Order Reduction LABoratory) [31], MORPACK (Model Order Reduction PACKage) [105] and sssMOR(sparse state space and Model Order Reduction toolbox)[51]. These toolboxes differ in methods, languages, operating environment and applications. Inspired by the implementation of MESS which incorporates with sparsity, a dedicated toolbox based on Chapter 2 is developed for large-scale models with a regular PC. An overview of the toolbox is given in Chapter 8.

## 1.3 Thesis outline

The remainder of this thesis is structured into three parts: In Part I, we focus on the parameterized model order reduction for different parameter-dependent dynamical systems that preserve their physical interpretation. Additionally, we also investigate the scheme of integrating the proposed approximation methods with parameter calibration applications. In Part II, we study the stability and dissipation properties of thermal systems. Specifically, the Lyapunov stability and the dissipative inequality are discussed for different dynamical systems. Numerical tooling is developed and summarized in Part III, based on the proposed parameterized model order reduction methods in Part I.

### Part I: Parameterized model order reduction and application

Chapter 2 proposes a general framework for physical parameter preservation reduction of large-scale linear parameterized systems. An implicit algorithm is proposed that establishes a perfect match of moments in the desired set of frequency points and the desired set of parameter samples. The simplified systems preserve the parameter-dependence as good as the original systems. Additionally,

an a-prior error bound has been derived that represents the local accuracy of the transfer function of the reduced-order model nearby interpolation points. The materials in this chapter are based on the conference paper [113].

Chapter 3 studies a novel reduction methodology for linear parameter-varying (LPV) systems, in which the dimension of the parameter space is truncated. A notion of Hankel norm for LPV systems is introduced to characterize the system Gramian with parametric variations. By analyzing the influence of parameter space on the system Gramian, we evaluate the system behaviour and construct the upper bound of the Hankel-norm for the given parametric variation. The reduction method is developed based on finding a reduced parameter space where the approximation error, defined as the Hankel-norm difference between the full-order LPV system and the reduced-order LPV system, is minimal. The materials in this chapter are based on [150].

Chapter 4 applies parameterized model order reduction to a class of calibration problems that involve large-scale models. The proposed scheme simultaneously reduces the computation time for calibration problems and guarantees the equivalent Hermite interpolation condition between the calibration problem using the full-order and approximated models. Based on the equivalent condition, an adaptive parameterized model reduction algorithm is proposed to minimize the calibration error, in which measurement data is considered in choosing the expansion points for the approximated model.

Chapter 5 explores a model reduction problem of uncertain systems represented in linear fractional transformation (LFT), which is a broader class of parameter-dependent systems. Specifically, two types of uncertain systems are studied: uncertain LTI systems and uncertain LPV systems. The concept of the moment for uncertain LTI systems and uncertain LPV systems is used to identify the critical dynamics of the uncertain system with respect to the uncertainties. We extend the moment-matching method to uncertain LTI systems and uncertain LPV systems. The proposed methods reduce the complexity of uncertain systems and preserve the LFT structure and the uncertainty properties of the systems.

## **Part II: Stability and dissipation analysis for thermal systems**

Chapter 6 investigates the stability properties of a class of thermal systems, which is governed by ordinary differential equations. A conceptual distinction is made between thermodynamic equilibrium and equilibria of autonomous systems. An entropy generation based Lyapunov function is proposed and proves the stability of thermodynamic equilibria. The materials in this chapter are based on [114].

Chapter 7 further studies the dissipation properties of thermal systems, in which the boundary condition is relaxed from time-independence to time-dependence. We introduce the concept of entropy for irreversible systems. We show that this concept satisfies the dissipative inequality. In addition, we extend to a thermal complex system and an energy-based upper bound and lower bound are derived. Moreover, this chapter provides a result of time evolution profiles for the proposed storage function with different boundary conditions.

## **Part III: Tooling of parameterized model order reduction**

Chapter 8 presents a software solution with a toolbox called paraMOR, which stands for **parameter Model Order Reduction**, for model order reduction of large-scale LTI and LPV systems. Here large-scale refers to systems with a state dimension of  $10^5$  that are assumed linear may be either parametric or non-parametric. The underlying programming principle is to maximize the memory storage usage efficiency and deliver the high-fidelity reduced-order model. An overview of functionalities and the structure of this toolbox is given. We briefly introduce the supported matrix decomposition methods and orthogonal methods. The chapter includes two case studies.

Finally, this dissertation is concluded in Chapter 9. Conclusions, limitations and some suggestions for future work are stated.

## 1.4 Lists of publications

### Proceedings and congress contributions

D. Lou, S. Weiland. *Stability analysis of Thermodynamic systems: Heat Conduction in Solids*. IFAC World Congress, Berlin, Germany, 2020.

S. Schouten, D. Lou, S. Weiland. *Model reduction for linear parameter-varying systems through parameter projection*. IEEE CONFERENCE ON DECISION AND CONTROL, NICE, FRANCE, 2019.

D. Lou, S. Weiland. *Parametric model order reduction for large-scale and complex thermal systems*. IEEE European Control Conference, Limassol, Cyprus, 2018.

D. Lou, S. Weiland, *Parameter calibration using Krylov subspace methods for large-scale systems*. 38th Benelux Meeting on Systems and Control, Lommel, Belgium, 2019. (abstract and presentation)

D. Lou, S. Weiland, *Error analysis for parametric model order reduction using Krylov subspace method*. 37th Benelux Meeting on Systems and Control, Soesterberg, Netherlands, 2018 (abstract and presentation)

D. Lou, S. Weiland, *Model reduction for industrial complex systems*. VDMA Electronics, Micro and Nano Technologies Symposium, Frankfurt, Germany, 2018 (invited talk)

D. Lou, S. Weiland, *Parametric model order reduction for large-scale and complex systems using Krylov subspace methods*. Model Reduction of Parameterized Systems IV Conference, Nantes, France, 2018 (abstract and presentation)

D. Lou, S. Weiland, *Parametric model order reduction for large-scale and thermal systems*. 36th Benelux Meeting on Systems and Control, Spa, Belgium, 2017 (abstract and presentation)

### In preparation

D. Lou, S. Weiland. *Parameter estimation using Krylov subspace methods for large-scale complex systems*. (To be submitted)

D. Lou, S. Weiland. *Model reduction for uncertainty systems using moment-matching method*. (To be submitted)

**Toolbox**

D. Lou, *paraMOR - Parameterized Model Order Reduction using Moment-Matching toolbox for Matlab*. (Distributed among partners of ATC consortium. Available upon request.)



## **Part I**

# **Parameterized model order reduction and applications**



# Parameterized model reduction of thermal systems

---

**I**n this chapter, we consider the parameterized model order reduction (pMOR) problem for a large-scale and complex thermal system. The main principle behind this method is that any change of the physical parameters in the high-fidelity model can be updated directly in the simplified model. For deriving the parametric reduced order model, a Krylov subspace method is employed, which yields the relevant subspaces of the projected state. The Krylov subspace is constructed via the moment-matching method, which captures the dynamics of the system in the frequency domain. Therefore, we first define the moments for parametric systems such that the frequency variable and parameter variables are distinguishable. Within this formulation and the projection operator, the first moments of the low order model are set to match the corresponding moments of the original model. Additionally, a prior upper bound of the error induced by the approximation is derived.

---

## 2.1 Introduction

Numerical simulations of complex dynamical systems are an indispensable tool in studying thermodynamic phenomena. However, for complex thermal systems where ultra-high precision simulations are required, the finite element method (FEM) commonly yields large-scale models. These models demand considerable computational resources. Therefore, model order reduction techniques are employed to reduce the computational complexity by replacing the high-order dynamic model with a low-order one. For the class of LTI systems, many reduction techniques have reached a relatively high level of maturity [13]. For systems with

uncertain or time-varying parameters there is a persistent need of novel reduction techniques. Since creating a new reduced model for every parameter value is inefficient and computationally costly, there is a strong need for pMOR techniques as introduced by [29]. We distinguish among reduction techniques in which both the state dimension and parameter vector dimension are reduced, and techniques in which state dimension is reduced while the parameter vector keeps its physical relevance. In addition, we distinguish among time-varying and time-invariant parametric dependence of high-fidelity models. In either of these cases, the objective is to find a low-cost, but accurate, parameterized model.

In the last decades, MOR techniques based on Padé approximation [12] have been used as powerful tools for large-scale system simulation, in particular for very-large-scale system integration simulation [84] and microelectromechanical simulation tasks [130]. The first algorithm based on Padé techniques approximation technique was an asymptotic waveform evaluation in [136], which uses explicit moment-matching. For a SISO system, the Arnoldi algorithm, also called the one-sided moment-matching technique [152] was proposed to match  $r$  moments of a transfer function by a  $r$ th-order approximation. Meanwhile, two-sided moment matching, known as Lanczos method [67], was introduced to perform matching of  $2r$  moments by a  $r$ th-order approximation. Then the block Arnoldi [68] and block Lanczos method [36] were proposed to solve the same problem for MIMO systems.

However, for large-scale systems which are parameter dependent, the aforementioned methods are not suitable anymore for obtaining a low order model while maintaining the physical interpretation of the parameters after the MOR procedure. Therefore, there is a need to develop such MOR techniques that allows the reduction of parameter-varying systems. Hence, a new branch of MOR is known as parametric model order reduction (pMOR).

The work of pMOR using Krylov subspace methods was introduced in [170], where the state evaluation matrix linearly depends on a single parameter. In consideration of the physical systems that can be described with multiple parameters, an extensive pMOR method which matches the coefficients of multivariate Taylor series is generalized in [53]. This method replaces the product of complex independent variables  $s$  and the parameter  $\theta$  in the transfer function  $H(s, \theta)$  by a set of redefined parameters, where the moments of the reduced model are equivalent to the corresponding moments of the original model with respect to the new expansion points. As a consequence, the reduced order model  $\hat{H}$  matches moments, but changes the meaning of the parameter vector  $\theta$  to a newly defined parameter  $\hat{\theta}$  unequal to  $\theta$ . That is, with these techniques  $\hat{H}(s, \hat{\theta})$  is the resulting approximate system with  $\hat{\theta} \neq \theta$ . In our opinion, this has the serious disadvantage that subsequent operation, optimization or calibration of parameters  $\theta$  cannot be performed on the original parameter vector  $\theta$ .

A second issue amounts to establishing error bounds. Contrary to singular value decomposition based truncation, the pMOR techniques based on Krylov subspaces have no prior error bound or global error bound.

In this chapter, we explore MOR techniques for parametric models with the purpose to match multiple moments of the transfer function. We introduce a separation between the physical parameters  $\theta$  and the frequency points  $s$  such that, unlike previous work, the physical interpretation of the parameters in the reduced model is maintained. Also, simulation results demonstrate the benefits of this separation. Moreover, an upper bound of the approximation error between the ROM and the FOM is derived, which allows making *an priori* estimate of the misfit between FOM and ROM, without calculating the exact reduced model.

The remainder of the chapter is structured as follows: In Section 2.2, a condensed problem description for parametric model order reduction is presented. The proposed method for parametric model order reduction with moment matching of multiple frequencies and multiple parameters is introduced in Section 2.3. The analytic expression of the upper bound of this method is given in Section 2.4. Simulation results are shown in Section 2.5. Finally, Section 2.6 concludes this work.

## 2.2 Problem description

The main goal of the reduction techniques proposed in this chapter is to preserve parameters in the system as symbolic quantities in the reduced order model. In fact, there are many key applications where this parametric model reduction technique is of crucial interest. These include design optimizations (where the parameter vector represents design parameters) in which the design-loop can substantially benefit from simplified parametric models, calibration of parameters through simplified models or control design for parameter dependent models. In either of these cases, a change in parameters or optimization of parameters does not require to repeat the reduction procedure, but simply the evaluation of the parametric model. The aim is to infer a reduced order model with smaller state dimension, but which remains explicit in the parameter vector  $\theta$ .

Consider a linear parameter-varying model with parameter  $\theta \in \Theta$  described as

$$\Sigma(\theta) := \begin{cases} E(\theta)\dot{x}(t) = A(\theta)x(t) + B(\theta)u(t) \\ y(t) = C(\theta)x(t) \end{cases} \quad (2.1)$$

where  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$  and  $y(t) \in \mathbb{R}^q$  denote, respectively, the state vector, the input and the output. Here,  $n$  is the dimension of the state space,  $\ell$  is the dimension of the parameter space. The state-space matrices are functions  $A : \mathbb{R}^\ell \rightarrow \mathbb{R}^{n \times n}$ ,  $E : \mathbb{R}^\ell \rightarrow \mathbb{R}^{n \times n}$ ,  $B : \mathbb{R}^\ell \rightarrow \mathbb{R}^{n \times m}$  and  $C : \mathbb{R}^\ell \rightarrow \mathbb{R}^{q \times n}$ . Before explaining the model reduction, we first state several assumptions.

**Assumption 2.1** *i) The state space matrices*

$$\begin{aligned} E(\theta) &= E_0 + \sum_{i=1}^{\ell} \theta_i E_i \\ A(\theta) &= A_0 + \sum_{i=1}^{\ell} \theta_i A_i \\ B(\theta) &= B_0 + \sum_{i=1}^{\ell} \theta_i B_i \\ C(\theta) &= C_0 + \sum_{i=1}^{\ell} \theta_i C_i \end{aligned}$$

*are affine functions of  $\theta \in \mathbb{R}^{\ell}$ . ii) For all  $\theta \in \Theta$ , all generalized eigenvalues of  $\sigma I - A(\theta)$  have negative real parts. iii)  $\theta \in \Theta \subset \mathbb{R}^{\ell}$  is not varying with time.*

With the time-invariant  $\theta$ , the transfer function of (2.1) is meaningful and is given as follows:

$$H(s, \theta) = C(\theta) [sE(\theta) - A(\theta)]^{-1} B(\theta). \quad (2.2)$$

The system has McMillan degree  $n$  if, for at least one  $\theta \in \Theta$ ,  $E(\theta)$ ,  $A(\theta)$ ,  $B(\theta)$ ,  $C(\theta)$  is controllable and observable. We consider a projection-based method for generating the reduced order parametric model. That is, let  $V \in \mathbb{R}^{n \times r}$  and  $W \in \mathbb{R}^{n \times r}$  be full rank matrices with  $r \ll n$ . Define the projection spaces

$$\mathcal{V} := \text{im}(V); \quad \mathcal{W} := \text{im}(W); \quad (2.3)$$

i.e. two  $r$ -dimensional subspaces of  $\mathbb{R}^n$ . Let  $\Pi_{\mathcal{V}}$  and  $\Pi_{\mathcal{W}}$  denote the (canonical) projections of  $\mathbb{R}^n$  onto  $\mathcal{V}$  and  $\mathcal{W}$ , respectively. That is  $\Pi_{\mathcal{V}} = V(V^T V)^{-1} V^T$  and  $\Pi_{\mathcal{W}} = W(W^T W)^{-1} W^T$ . Define the reduced state

$$x_r := (V^T V)^{-1} V^T x \quad (2.4)$$

It clearly satisfies

$$V x_r = \Pi_{\mathcal{V}} x \quad (2.5)$$

for any  $x \in \mathbb{R}^n$ , so that  $V x_r$  is an approximation of the state variable  $x$  in (2.1) incurring the error

$$\|x - V x_r\| = \|(I - \Pi_{\mathcal{V}})x\| = \|\Pi_{\mathcal{V}^{\perp}} x\| \quad (2.6)$$

which is the Euclidean length of the projection of  $x$  onto  $\mathcal{V}^{\perp}$ . We call  $\Pi_{\mathcal{V}} x$  the state projection. Moreover,  $x_r$  defined in (2.4) minimizes the error (2.6) in the sense that

$$\|x - V x_r\| \leq \|x - V x'_r\|$$

for any vector  $x'_r \in \mathbb{R}^r$ . Thus,  $x_r$  in (2.4) defines the *optimal* coordinates in the basis defined by the columns of  $V$  of the projection of  $x$  onto  $\mathcal{V}$ . Clearly, substituting  $V x_r$  for  $x$  in (2.1) incurs a residual error

$$\rho := E(\theta) V \dot{x}_r - A(\theta) V x_r - B(\theta) u \quad (2.7)$$

The projection  $\Pi_{\mathcal{W}}$  is meant to require that

$$\Pi_{\mathcal{W}} \rho(t) = 0 \quad \text{for all } t \geq 0. \quad (2.8)$$

We call this the residual projection. When combining state and residual projection the state space representation of the reduced model is then given by

$$\Sigma_r(\theta) := \begin{cases} E_r(\theta)\dot{x}_r(t) = A_r(\theta)x_r(t) + B_r(\theta)u(t) \\ y(t) = C_r(\theta)x_r(t) \end{cases} \quad (2.9)$$

where the reduced state space matrices are

$$\begin{aligned} E_r(\theta) &= W^T E(\theta) V, & A_r(\theta) &= W^T A(\theta) V \\ B_r(\theta) &= W^T B(\theta), & C_r(\theta) &= C(\theta) V. \end{aligned} \quad (2.10)$$

It is important to observe that, under Assumption 2.1, i) all matrices in (2.10) are again affine functions of  $\theta$  that can be computed directly in terms of the expansion matrices in Assumption 2.1. That is,

$$\begin{aligned} E_r(\theta) &= W^T E(\theta) V \\ &= W^T E_0 V + \sum_{i=1}^{\ell} W^T E_i V \theta_i = E_{r,0} + \sum_{i=1}^{\ell} E_{r,i} \theta_i \end{aligned}$$

with similar expansions for the other matrices. Thus,  $E_r, A_r, B_r, C_r$  are affine in  $\theta$  where the expansion coefficients can be derived directly in terms of the expansions of  $E, A, B, C$  and are independent of  $\theta$ .

The transfer function of the parametric ROM is

$$H_r(s, \theta) = C_r(\theta) [sE_r(\theta) - A_r(\theta)]^{-1} B_r(\theta) \quad (2.11)$$

If  $W = V$ , this is called an (ordinary) Galerkin projection. If  $W \neq V$ , this is a Petrov-Galerkin projection.

## 2.3 Parameterized model order reduction

In this section, the moment-matching method is elaborated first for an LTI system. In the second part, we present a parametric moment-matching method which separates the frequency variable  $s$  and the parameter  $\theta$ .

### 2.3.a Moment-matching

Suppose that an LTI system is given according to

$$\begin{cases} E\dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \quad (2.12)$$

with  $x(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^m$  and  $y(t) \in \mathbb{R}^q$ . The matrix pencil  $A - sE$  is assumed to be *regular*. In this way, (2.12) has the transfer function

$$H(s) = C(sE - A)^{-1} B \in \mathbb{C}^{q \times m}. \quad (2.13)$$

Given a region  $\mathcal{R}(s_0) := \{s \mid |s - s_0| < R\}$  around the point  $s_0 \in \mathbb{C}$ , on which the complex function  $H(s)$  is analytic it admits, on the region, a Taylor series about the point  $s_0$  given by the infinite power series

$$\begin{aligned} H(s) &= H(s_0) + \frac{d}{ds}H(s) \Big|_{s=s_0} (s - s_0) + \frac{1}{2} \frac{d^2}{ds^2}H(s) \Big|_{s=s_0} (s - s_0)^2 \dots \\ &= \sum_{i=0}^{\infty} \underbrace{\frac{1}{i!} \frac{d^i}{ds^i}H(s) \Big|_{s=s_0}}_{:=m_i(s_0)} (s - s_0)^i \end{aligned} \quad (2.14)$$

where this infinite series converges point-wise in  $s$  for all  $s \in \mathcal{R}(s_0)$ . The coefficients  $m_i(s_0)$  defined in (2.14)

$$m_i(s_0) := \frac{1}{i!} \frac{d^i}{ds^i}H(s) \Big|_{s=s_0} \quad (2.15)$$

are referred to as the (*analytic*) moments of (2.13) at  $s_0$ . The moments  $m_i(s_0)$  can be computed by Cauchy's integral formula

$$m_i(s_0) = \frac{1}{2\pi j} \oint \frac{H(s)}{(s - s_0)^{i+1}} ds \quad (2.16)$$

where the integration is counterclockwise along a path enclosing  $s_0$  and lying in the region  $\mathcal{R}(s_0)$  where the power series converges. With such an expansion, the moment  $m_i(s_0)$  can be determined from the state space representation (2.12) as follows

**Theorem 2.1** *If the matrix pencil  $sE - A$  of (2.12) is regular and the transfer function  $H(s)$  is analytic in the region of  $\mathcal{R}(s_0)$ , then the  $i$ th (analytic) moment about point  $s_0$  is*

$$m_i(s_0) = C[-(s_0E - A)^{-1}E]^i (s_0E - A)^{-1}B.$$

For the proof we refer to [18]. □

A similar procedure can be applied to negative power expansions of the form

$$H(s) = \sum_{i=-\infty}^0 m_i(s)(s - s_0)^i$$

where the domain of convergence lies outside a region of the form  $\mathcal{R}(s_0)$  for  $R$  sufficiently large, i.e.,  $\{s \in \mathbb{C} \mid |s - s_0| > R\}$ . We will not further elaborate on these Laurent expansions.

The problem that we solve in this chapter amounts to finding the reduced order model (2.9) with transfer function (2.11) in such a way that the moments  $m_i(s_0)$  of the full order model match the moments  $m_{i,r}(s_0)$  of the reduced order model at a user-defined frequency point  $s_0$  and for a user defined number of moments  $i = 1, \dots, r$ . We construct the approximated model as in (2.9) and (2.11), where



the projection matrices  $V, W$  are generated via Krylov subspace methods. For any positive integer  $\omega$ , the Krylov subspace is defined as

$$\mathcal{K}_\omega(M, f) = \text{colspan}\{f, Mf, \dots, M^{\omega-1}f\} \quad (2.17)$$

where

$$\begin{aligned} f &= (s_0 E - A)^{-1} B \in \mathbb{R}^{n \times m} \\ M &= (s_0 E - A)^{-1} E \in \mathbb{R}^{n \times n}. \end{aligned}$$

and where  $\text{colspan}$  is the operator that produces the linear span of the columns of all matrices in its arguments. Then  $\mathcal{K}_\omega(M, f)$  is a subspace of  $\mathbb{C}^n$  which has dimension  $\leq \omega \times m$  whenever  $s_0$  is non-real. If  $s_0$  is real,  $\mathcal{K}_\omega(M, f)$  is a subspace of  $\mathbb{R}^n$ . This is sometimes called the right Krylov subspace associated with the system (2.12).

**Theorem 2.2** *If the matrix  $V$  in (2.3), is a basis of Krylov subspace  $\mathcal{K}_\omega(M, f)$  and  $W$  in (2.3) is chosen  $W = V$ , then the first  $\omega$  moments around  $s_0$  of the original and the reduced model match*

$$m_i(s_0) = m_{r,i}(s_0), \quad i = 0, 1, \dots, \omega - 1,$$

here  $m_i(s_0), m_{r,i}(s_0)$  denote the  $i$ th moment of the full order model and the corresponding reduced model generated as in (2.9), respectively.

**Proof:** Based on Theorem 2.2, the 0th moment of the reduced model is

$$m_{r,0} = C_r(s_0 E_r - A_r)^{-1} B_r = CV(s_0 W^T E V - W^T A V)^{-1} W^T B.$$

The vector  $(s_0 E - A)^{-1} B$  can be written as a linear combination of the columns of the projection matrix  $V$

$$\exists \mathbf{r}_0 \in \mathbb{C}^{\omega \times m}, (s_0 E - A)^{-1} B = V \mathbf{r}_0.$$

Therefore,

$$\begin{aligned} W^T B &= W^T [(s_0 E - A)(s_0 E - A)^{-1}] B = W^T (s_0 E - A) V \mathbf{r}_0 \\ &= (s_0 W^T E V - W^T A V) \mathbf{r}_0 \end{aligned}$$

With this, the  $m_{r,0}$  is

$$\begin{aligned} m_{r,0}(s_0) &= CV(s_0 W^T E V - W^T A V)^{-1} W^T B \\ &= CV(s_0 W^T E V - W^T A V)^{-1} (s_0 W^T E V - W^T A V) \mathbf{r}_0 \\ &= C V \mathbf{r}_0 = C(s_0 E - A)^{-1} B \\ &= m_0(s_0) \end{aligned}$$

By repeating these steps in the same manner, the proof can be continued until

$$m_{r,\omega-1}(s_0) = m_{\omega-1}(s_0). \quad \square$$

Loosely speaking, the Krylov subspace derived above contains the information of the input-sided moments, defined by  $(E, A, B)$  only. In a similar way, the output-sided moments involve the triple  $(E, A, C)$  as follows.

$$\mathcal{K}_\omega(M^T, l) = \text{colspan}\{l, M^T l, \dots, (M^T)^{\omega-1} l\} \quad (2.18)$$

where

$$\begin{aligned} l &= (s_0 E - A)^{-T} C^T \in \mathbb{R}^{n \times q} \\ M^T &= (s_0 E - A)^{-T} E^T \in \mathbb{R}^{n \times n}, \end{aligned}$$

and  $\mathcal{K}_\omega(M^T, l)$  has column dimension  $\leq q \times \omega$ . If the reduced model is constructed via Petrov-Galerkin projection  $W \neq V$ , we have the following theorem

**Theorem 2.3** *If the matrices  $W$  and  $V$  with  $q = m$  in (2.3) are the basis of the Krylov subspace  $\mathcal{K}_\omega(M, f)$  and  $\mathcal{K}_\omega(M^T, l)$ , then the first  $2\omega - 1$  moments around  $s_0$  of the original and the reduced model match*

$$m_i(s_0) = m_{r,i}(s_0), \quad i = 0, \dots, 2\omega - 1.$$

**Proof:** According to the Proof of Theorem 2.2, the first  $\omega$  moments with respect to the  $\mathcal{K}_\omega(M, f)$  coincide. The matrix  $(s_0 E - A)^{-T} C^T$  can be written as a linear combination of the columns of  $W$

$$\exists \mathbf{l}_0 \in \mathbb{C}^{\omega \times q}, (s_0 E - A)^{-T} C^T = W \mathbf{l}_0$$

Following the previous proof, we have

$$\begin{aligned} CV(s_0 W^T EV - W^T AV)^{-1} W^T EV &= C[(s_0 E - A)^{-1}(s_0 E - A)] W^T EV \\ &= \mathbf{l}_0^T W^T EV = C(s_0 E - A)^{-1} EV \end{aligned}$$

Then, from the Theorem 2.2, the  $m_{r,\omega-1}(s_0)$  is

$$\begin{aligned} &V[(s_0 W^T EV - W^T AV)^{-1} W^T EV]^{\omega-1} (s_0 W^T EV - W^T AV)^{-1} W^T B \\ &= V \mathbf{r}_\omega = [(s_0 E - A)^{-1} E]^{\omega-1} (s_0 E - A)^{-1} B \end{aligned}$$

Thus, for the moment  $m_\omega(s_0)$  is

$$\begin{aligned} m_\omega(s_0) &= C[(s_0 E - A)^{-1} E][(s_0 E - A)^{-1} E]^{\omega-1} (s_0 E - A)^{-1} B \\ &= C_r[(s_0 E_r - A_r)^{-1} E_r] V[(s_0 E_r - A_r)^{-1} E_r]^{\omega-1} (s_0 E_r - A_r)^{-1} B_r \\ &= m_{r,\omega-1}(s_0). \end{aligned} \quad (2.19)$$

As in the previous proof, the step can be repeated until  $m_{r,2\omega-1}(s_0) = m_{2\omega-1}(s_0)$ .

□

Notice that for the case the input number is not equal to the output  $q \neq m$  is not considered here.

**Remark 2.1** The construction of projection matrices  $V, W$  is for a single expansion point. For the case of multiple expansion points  $\{s_1, s_2, \dots, s_N\}$  and multiple orders  $r_1, \dots, r_N$ , we define

$$\begin{aligned} V &= \text{im}[V_1, V_2, \dots, V_N] \\ W &= \text{im}[W_1, W_2, \dots, W_N]. \end{aligned}$$

Each  $V_i, W_j$  can be constructed via (2.17) and (2.18) with expansion points and Krylov orders  $(s_i, \omega_i)$  and  $(s_j, \omega_j)$ , respectively.

**Remark 2.2** If the complex expansion points  $s_0$  is chosen, the method above leads to complex-valued matrices  $V, W$ . As a result, the system matrices of the ROM contain complex numbers which are physically undesirable. One solution is to add  $s_0$ 's complex conjugate as the complex pair  $[s_0, s_0^*]$ . This will result in complex-valued matrices  $V_c, W_c$ . The real-valued projection matrices are defined as the

$$\begin{aligned} V &= [\text{real}(V_c), \text{imag}(V_c)], \\ W &= [\text{real}(W_c), \text{imag}(W_c)]. \end{aligned}$$

### 2.3.b Multi-parameter and multi-frequency moment matching

The main ingredient of the multi-parameter and multi-frequency moment matching method is to expand the transfer function into a Taylor series at both a desired frequency and a desired parameter  $(s_0, \theta_0)$ . This work has been generalized in [53]. Using the Assumption 2.1, we obtain

$$H(s, \theta) = C(\theta)[sE(\theta) - A(\theta)]^{-1}B(\theta) \quad (2.20)$$

where the expansion of  $sE(\theta)$  involves products of  $s$  with  $\theta$ . One remedy for this problem is to define virtual frequency points  $\hat{s}_i := s \cdot \theta$ . This has been proposed in [53]. Potentially, combining the  $s$  with  $\theta_i$  as a auxiliary variable  $\hat{s}_i$  can decrease the complexity of the power series. In such a way, the computational burden of the projection matrices  $V$  and  $W$  is reduced. However, with such conversion the reduced model matches the first moments of the original model with respect to  $\hat{s}_i$  and the physical interpretation of the parameters  $\theta$  in the reduced model is no longer preserved. Moreover, the Taylor expansion of the new parametric transfer function includes terms such as  $(\frac{\partial^2 H}{\partial \hat{s}_i \partial \hat{s}_{i+1}})$  which are partial derivatives of composite functions rather than the original one. As a consequence, also the interpretation of the moments as Markov parameters defined in Theorem 2.1 changes under this redefinition of the frequencies and parameters.

We first give the Taylor series derivation of a single frequency point and a single parameter expansion point. Assume that the transfer function (2.2) about  $(s_0, \theta_0)$  is analytic in a region  $\mathcal{R}(s_0) \times \Theta(\theta_0)$  where  $\mathcal{R}(s_0) := \{s \mid |s - s_0| < R_s\}$  and  $\Theta(\theta_0) := \{\theta \mid |\theta - \theta_0| < R_\theta\}$ , respectively. Then the Taylor series of  $H(s_0, \theta_0)$ ,  $\ell = \dim(\theta) = 1$  given by infinite power series is

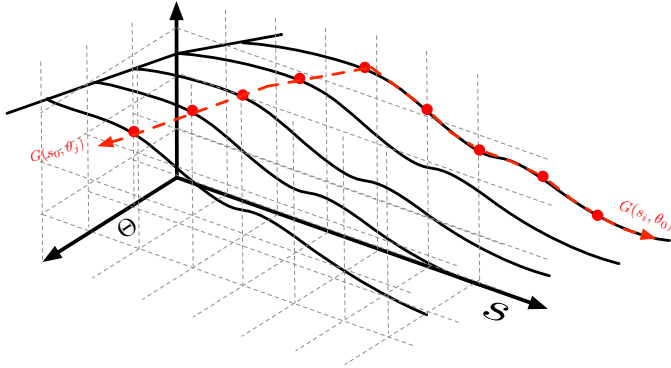
$$\begin{aligned}
 H(s, \theta) &= H(s, \theta) \Big|_{(s=s_0, \theta=\theta_0)} \\
 &+ \frac{\partial}{\partial s} H(s, \theta) \Big|_{(s=s_0, \theta=\theta_0)} (s - s_0) + \frac{\partial}{\partial \theta} H(s, \theta) \Big|_{(s=s_0, \theta=\theta_0)} (\theta - \theta_0) \\
 &+ \frac{1}{2} \frac{\partial^2}{\partial^2 s} H(s, \theta) \Big|_{(s=s_0, \theta=\theta_0)} (s - s_0)^2 + \frac{1}{2} \frac{\partial^2}{\partial^2 \theta} H(s, \theta) \Big|_{(s=s_0, \theta=\theta_0)} (\theta - \theta_0)^2 \\
 &+ \frac{1}{2} \left[ \frac{\partial^2}{\partial s \partial \theta} + \frac{\partial^2}{\partial \theta \partial s} \right] H(s, \theta) \Big|_{(s=s_0, \theta=\theta_0)} (s - s_0)(\theta - \theta_0) + \dots \\
 &= \sum_{I=0, J=0}^{\infty} m_{\{I, J\}}(s_0, \theta_0) (s - s_0)^I (\theta - \theta_0)^J,
 \end{aligned}$$

and where

$$m_{\{I, J\}}(s_0, \theta_0) = \frac{1}{(I + J)!} \frac{\partial^{I+J}}{\partial s^I \partial \theta^J} H(s_0, \theta_0)$$

for all  $I \geq 0$  and  $J \geq 0$ .

The above derivation presents the Taylor series for a single frequency point and a single parameter point. By defining two sets  $\mathcal{S} := \{s_1, \dots, s_k\} \subset \mathbb{C}$  and  $\Theta := \{\theta_1, \dots, \theta_\ell\} \subset \mathbb{R}^\ell$ , we can extend this to multiple-frequency and multiple-parameter expansion points, as depicted in Figure 2.1.



**Figure 2.1:** Multi-frequency and multi-parameter expansion

The (full rank) parametric projection matrices  $\{V, W\}$  that satisfy (2.3) follows

$$\mathcal{V} = \text{im}[v(s_1, \theta_1), \dots, v(s_k, \theta_1), \dots, v(s_1, \theta_\ell), \dots, v(s_k, \theta_\ell)] \quad (2.21)$$

$$\mathcal{W} = \text{im}[w(s_1, \theta_1), \dots, w(s_k, \theta_1), \dots, w(s_1, \theta_\ell), \dots, w(s_k, \theta_\ell)] \quad (2.22)$$

where

$$\begin{aligned} v(s_i, \theta_j) &= \mathcal{K}_{\sigma_{i,j}^v}(M_{i,j}^v, F_{i,j}) \\ w(s_i, \theta_j) &= \mathcal{K}_{\sigma_{i,j}^w}(M_{i,j}^w, L_{i,j}) \end{aligned}$$

with

$$\begin{aligned} F_{i,j} &= (s_i E(\theta_j) - A(\theta_j))^{-1} B_{\theta_j} \\ M_{i,j}^v &= -(s_i E(\theta_j) - A(\theta_j))^{-1} E_{\theta_j} \\ L_{i,j} &= (s_i E(\theta_j) - A(\theta_j))^{-T} C_{\theta_j}^T \\ M_{i,j}^w &= -(s_i E(\theta_j) - A(\theta_j))^{-T} E_{\theta_j}^T. \end{aligned}$$

Here  $B_{\theta_j} = \frac{d}{d\theta}(B(\theta_j))$ ,  $E_{\theta_j} = \frac{d}{d\theta}(E(\theta_j))$ ,  $C_{\theta_j} = \frac{d}{d\theta}(C(\theta_j))$ .  $\sigma_{i,j}^v, \sigma_{i,j}^w$  represent the number of preserved moments for each pair  $(s_i, \theta_j)$  in  $v(s_i, \theta_j)$  and  $w(s_i, \theta_j)$ , respectively. As in (2.11), the transfer function of the reduced model is

$$H_r(s, \theta) = C_r(\theta)[sE_r(\theta) - A_r(\theta)]^{-1} B_r(\theta). \quad (2.23)$$

which coincides in as many coefficients of its moments about  $(s_i, \theta_j)$  as possible for a given order. That is, we have

**Theorem 2.4** *Consider the transfer function  $H(s, \theta)$  in (2.20). If the matrices  $W, V$  are the basis of Krylov subspace  $\mathcal{W}, \mathcal{V}$  as defined in (2.21) and (2.22), this leads to the transfer function  $H_r(s, \theta)$  of the reduced model in (2.9). Then the Hermite interpolation condition is satisfied as*

$$m_{\{I_i, J_j\}}(s_i, \theta_j) = m_{r, \{I_i, J_j\}}(s_i, \theta_j)$$

for all frequency points  $s_i \in \{s_1, \dots, s_k\}$  and parameter points  $\theta_j \in \{\theta_{11}, \dots, \theta_{\ell}\}$

$$\begin{aligned} I_i &= 0, \dots, \sigma_{i,j}^v + \sigma_{i,j}^w - 1 \\ J_j &= 0, \dots, \sigma_{i,j}^v + \sigma_{i,j}^w - 1. \end{aligned}$$

Here  $m_{\{I_i, J_j\}}(s_i, \theta_j)$ ,  $m_{r, \{I_i, J_j\}}(s_i, \theta_j)$  denotes the  $\{I_i, J_j\}$ th moment of the original model  $\Sigma(\theta)$  in (2.1), the reduced model  $\Sigma_r(\theta)$  in (2.9), respectively.

It can be proved using the similar proof approach of Theorem 2.3.

## 2.4 Error estimation

Guarantees on the fidelity of the reduced model remain an important question for the previously described moment matching procedure. One major issue is establishing a prior estimate of the approximation error. In the following we will derive an error estimate between  $H(s, \theta)$  and  $H_r(s, \theta)$  which provide a local accuracy at points  $(s_0, \theta_0) \in \mathcal{S} \times \Theta$  and their neighborhoods. That is, the error estimate of  $H(s, \theta) - H_r(s, \theta)$  will have local validity in  $(s_0, \theta_0)$  and is based on the truncation order of the Taylor series.

### 2.4.a The remainder in Taylor series

Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be continuous on the interval  $[a, b]$  and differentiable on the open interval  $(a, b)$ . Then there exists  $c$  with  $a < c < b$  such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \quad (2.24)$$

The above result is the Mean Value Theorem (MVT) which is a fundamental theorem of calculus. Taylor's formula can be viewed as a generalization of the MVT. In particular, let  $f$  be a function that is continuous at  $a$ , then Taylor's theorem:

**Theorem 2.5** *Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a function that is infinitely differentiable at the point  $a$ . Then  $f(x)$  can be written as the power series*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x - a) + \cdots + \frac{f^{(N)}(a)}{N!}(x - a)^N + \cdots \quad (2.25)$$

$$= \sum_{i=0}^{\infty} \frac{f^{(i)}(a)}{i!}(x - a)^i \quad (2.26)$$

and the  $N$ th-degree Taylor polynomial is obtained

$$f_N(x) = f(a) + \sum_{i=1}^N \frac{f^{(i)}(a)}{i!}(x - a)^i \quad (2.27)$$

Thus we can write Taylor's theorem

$$f(x) = f_N(x) + R_N(x) \quad (2.28)$$

where  $R_N(x) := f(x) - f_N(x) = \sum_{i>N} \frac{f^{(i)}(a)}{i!}(x - a)^i$  is the remainder term which quantifies the difference between the function  $f$  and the  $N$ th-degree Taylor polynomial in terms of the magnitude of the  $(N + 1)$ st higher order derivative of  $f$ .

**Theorem 2.6** *Suppose that  $f : \mathbb{R} \rightarrow \mathbb{R}$  is defined on a closed interval  $I$  that has  $a$  in its interior and  $f^{(N+1)}$  exists on the same interval for  $N \in \mathbb{Z}_+$ . Then for each  $x \neq a$ ,  $x \in I$ , there is a  $\xi \in (a, x)$  such that*

$$R_N(x) = \frac{f^{(N+1)}(\xi)}{(N + 1)!}(x - a)^{N+1}. \quad (2.29)$$

Thus, there exists a  $\xi \in I$ , with depending on whether  $x < a$  or  $x > a$  such that the remainder  $R_N(x)$  in (2.28) satisfies (2.29). A key observation is that when  $N = 1$ , this reduces to the ordinary MVT. The above derivation (2.28) (2.29) can be referred to [2]. Following this strategy, the key is to observe the generalization of Rolle's theorem.

**Proof:** Indeed, under the condition that the power series of  $f$  exists (and thus  $f$  is infinitely differentiable), and both  $x, a \in I$ , then, by definition,

$$R_N(x) := \sum_{i>N} \frac{f^{(i)}(x)}{i!} (x-a)^i$$

for every  $x \in I$ . Then apply MVT to this function by first noting that  $R_N$  is continuous on  $I$  (since  $f$  is) and, secondly, by noting that  $R_N$  is differentiable on  $I$  (since  $f$  is) and thirdly, by noting that  $R'_N(x) = \frac{f^{(N+1)}(x)}{(N+1)!}$ . The result then follows by applying MVT to  $R_N(x)$  and gives (2.29).  $\square$

Furthermore, we can find the upper bound of the remainder term in Theorem (2.6) using Taylor's inequality.

**Theorem 2.7** *Under the conditions of Theorem (2.6), let  $I_a \subseteq I$  be any interval that has  $a$  in its interior. We define*

$$\mathcal{M} := \sup_{\xi \in I_a} \left| \frac{f^{(N+1)}(\xi)}{(N+1)!} \right| \quad (2.30)$$

Then

$$|f(x) - f_N(x)| = |R_N(x)| \leq \mathcal{M}|x-a|^{N+1} \quad (2.31)$$

holds for all  $x \in I_a$ .

**Proof:** Use Theorem (2.6) and (2.29) to infer that, there exists  $\xi \in \mathcal{L}(x, a)$

$$\begin{aligned} |f(x) - f_N(x)| &= |R_N(x)| = \left| \frac{f^{(N+1)}(\xi)}{(N+1)!} (x-a)^{N+1} \right| \\ &\leq \sup_{\xi \in I_a} |x-a|^{N+1} = M|x-a|^{N+1} \end{aligned}$$

$\square$

Hence an upper bound of the error between the  $N$ th degree Taylor approximation  $f_N$  and  $f$  is obtained and is a prior bound, only depending on  $f$  itself and the considered interval  $I$ . In the following, we generalize Theorem 2.7 to a multivariable function  $f : \mathbb{C}^\ell \rightarrow \mathbb{C}^{m \times p}$  for MIMO systems with  $\ell$  expansion points,  $m$  inputs and  $q$  outputs.

**Theorem 2.8** *Suppose  $f : \mathbb{C}^\ell \rightarrow \mathbb{C}^{m \times p}$  is  $N+1$  times differentiable on a set  $I \subset \mathbb{C}^\ell$  that contains a vector  $a \in \mathbb{C}^\ell$  in its interior. Let  $I_a \subset I$  be a closed and bounded subset of  $I$ . Then the truncated  $N$ th order Taylor polynomial around  $a \in \mathbb{C}^\ell$  at order  $N$  is*

$$f_N(x) = f(a) + \sum_{|i|=0}^N \frac{[D^{|i|}f]_i(a)}{i!} (x-a)^i$$

where  $i = (i_1, \dots, i_\ell)$  is a multi-index,  $|i| = i_1 + \dots + i_\ell$  is its "length", and  $i! = i_1! \cdots i_\ell!$  is the factorial of  $i$ .  $(x-a)^i = (x_1-a_1)^{i_1} \cdots (x_\ell-a_\ell)^{i_\ell}$  is a multivariate scalar polynomial

and the  $k$ th partial derivative of  $f$  at a point  $a \in \mathbb{C}^\ell$  is denoted by

$$[D^{|k|}f]_k := \left. \frac{\partial^{k_1} \dots \partial^{k_\ell}}{\partial x_1^{k_1} \dots \partial x_\ell^{k_\ell}} f(x) \right|_{x=a} \quad (2.32)$$

where the multi-index is  $k = (k_1, \dots, k_\ell)$  and define  $|k| = \sum_i k_i$  as its cardinality.

When ranging over all multi-indices  $k = (k_1, \dots, k_\ell)$  with  $|k| = K$ , then  $[D^{|k|}f]_k$  can be interpreted as a order  $|k|$  tensor, i.e., the multi-linear functional

$$[D^{|k|}f] : \underbrace{\mathbb{C}^\ell \times \mathbb{C}^\ell \times \dots \times \mathbb{C}^\ell}_{|k| \text{ copies}} \rightarrow \mathbb{C} \quad (2.33)$$

whose  $k = (k_1, \dots, k_\ell)$ th entry is  $[D^{|k|}f]$  evaluated at  $x = a$  defined in (2.32).

Thus,  $[D^0f](a) = f(a)$  is the evaluation of  $f$  at the point  $a$ ,  $[D^1f](a) = \nabla f(a)$  is the gradient of  $f$  at the point  $a$ ,  $[D^2f](a) = \nabla^2 f(a)$  is the Hessian matrix of  $f$  at the point  $a$  and

$$[D^{|N|}f](v_1, \dots, v_N) := \sum_{|k|=0}^N [D^{|N|}f]_k (v_1 \otimes \dots \otimes v_N). \quad (2.34)$$

Then, the  $N$ th order multivariable approximation of  $f$  is given below.

**Theorem 2.9** For a function  $f : \mathbb{C}^\ell \rightarrow \mathbb{C}^{m \times p}$  is  $N+1$  times differentiable on a set  $I \subset \mathbb{C}^\ell$  that contains a vector  $a \in \mathbb{C}^\ell$  in its interior. The  $N$ th order multivariable approximation or multivariable Taylor polynomial of  $f$  at  $a \in \mathbb{C}^\ell$  is

$$\begin{aligned} f_N(x) &= [D^0f](a) + [D^1f]([x-a]) + \dots \\ &\quad + \frac{1}{N!} [D^{|N|}f] \left( \underbrace{[x-a], \dots, [x-a]}_{N \text{ copies}} \right). \end{aligned} \quad (2.35)$$

Here the tensor definition of  $[D^0f], [D^1f], \dots, [D^Nf]$  is given in (2.34).

Define the error by the residual which has the similar form as the single variable one (2.28)

$$R_N(x) := f(x) - f_N(x), \quad (2.36)$$

with

$$R_N(x) = \frac{(x-a)^{N+1}}{2\pi i} \int \frac{f(\xi)}{(\xi-a)^{N+1}(\xi-x)} d\xi$$

Let, as (2.30)

$$\mathcal{M} := \sup_{\xi \in I_a} \left\| \frac{D^{|N+1|}f(\xi)}{(N+1)!} \right\| \quad (2.37)$$

where the norm is the Frobenius norm of a multi-linear operator and the supremal



is taken over expansion points  $\xi \in I_a \subset I$ . We claim that

$$\|f(x) - f_N(x)\| \leq \mathcal{M} \|x - a\|^{N+1}; \forall x \in I_a \quad (2.38)$$

**Theorem 2.10** *Suppose  $H : \mathbb{C} \times \mathbb{C}^\ell \rightarrow \mathbb{C}$  is  $N + 1$  times differentiable on a set  $I \subset \mathbb{C}^\ell$  that contains  $(s_0, \theta_0)$  in its interior. Let  $H_N$  denote the  $N$ th order Taylor polynomial of  $H$  about  $(s_0, \theta_0)$ . Then*

$$\|H(s, \theta) - H_N(s, \theta)\| \leq \mathcal{M} \left\| \begin{bmatrix} s - s_0 \\ \theta - \theta_0 \end{bmatrix} \right\|^{N+1} \quad (2.39)$$

holds for all  $(s, \theta) \in I_{(s_0, \theta_0)}$ , where

$$\mathcal{M} := \sup_{(\xi_s, \xi_\theta) \in I_{(s_0, \theta_0)}} \left\| \frac{D^{N+1} H(\xi_s, \xi_\theta)}{(N+1)!} \right\| \quad (2.40)$$

is the supremum over the induced norm of the tensor  $[D^{N+1} H]$  over  $I_{(s_0, \theta_0)}$ .

Hence, (2.39) is the error bound for a multivariable function for a MIMO system.

The function  $H(s, \theta)$  above can be interpreted as the transfer function of the FOM in (2.2) and  $H_N(s, \theta)$  can be considered as  $N$ th ROM (2.11) at points  $(s_0, \theta_0)$ , respectively. With the desired reduced order  $N$  and expansion points  $(s_0, \theta_0)$ , the Theorem (2.10) gives a prior upper bound on the reduction error using the moment matching method. This means that the reduction error at each expansion point can be evaluated before constructing the ROM. We can minimize the reduction error and find the optimal ROM by adjusting the reduced order  $N$  and expansion points  $(s_0, \theta_0)$ . However, the upper bound (2.40) is, by no means, easy to compute.

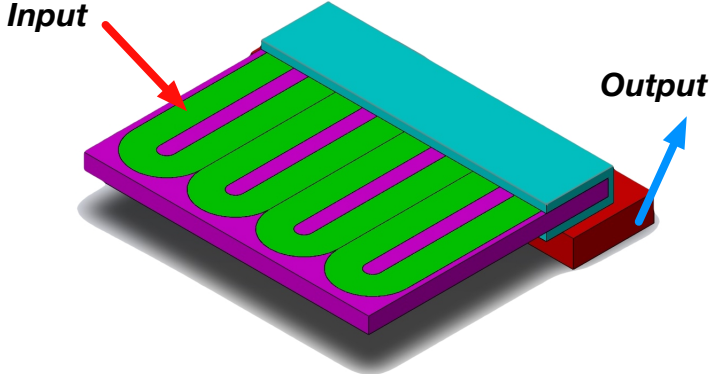
## 2.5 Example: A parametric thermal model

In this section, we demonstrate some numerical results of the proposed method. To illustrate the procedure and technique, we consider the example of a linear motor (see Figure 2.2). The coils of the linear motor generate a magnetic flux that operate along the track. Meanwhile, the coil also creates a heat flux which influences the temperature and, with that the stability of the thermal system. The aim is to predict the transient thermal behavior of the linear motor. Let  $\Omega \subset \mathbb{R}^3$  be the spatial configuration space of the motor. Consider the heat equation on  $\Omega$  as given by

$$c_p \rho \cdot \dot{T} = \nabla \cdot (\kappa \nabla T) + \dot{q} \quad \text{in } \Omega \quad (2.41)$$

$$T(\varphi, 0) = T_0(\varphi) \quad T_\partial(\varphi, t) \quad \text{on } \partial\Omega \quad (2.42)$$

where  $T(\varphi, t)$  denotes the temperature of the linear motor at position  $\varphi \in \Omega$  and at time  $t \geq 0$ . The initial temperature  $T_0(\varphi)$  and the boundary condition  $T_\partial(\varphi, t)$  for



**Figure 2.2:** Linear motor with single-input and single-output

all  $t \geq 0$  are given, respectively.  $c_p$  is the material specific heat capacity,  $\rho$  denotes the material density and  $\kappa$  describes the thermal conductivity. The heat flux at the coil of the linear motor is spatially distributed and is considered as the input of the system. The surface temperature where the sensor locates is chosen as the output.

Using a finite element (FE) discretization of  $\Omega$  and parameterizing the material properties, a parameter-dependent system is generated

$$\begin{cases} E(\theta)\dot{T} = AT + Bu \\ y = CT \end{cases}, \quad (2.43)$$

where, with some abuse of notation,  $T(t) \in \mathcal{R}^n$  denotes the vector of all temperatures in  $n = 1560$  finite elements and where  $\theta$  denotes the heat capacity of the material in the coil (indicated in green color). We treat the heat capacity  $c_p$  as the uncertain parameter. It enters the matrix  $E$  in (3.35) as an affine function in the sense that

$$E(\theta) = E_0 + \theta E_1$$

where  $\theta = c_p$ . The state  $T$  represents spatial-temporal information and is given as

$$T(t) = \text{col}(T(x_1, t), \dots, T(x_{1560}, t))$$

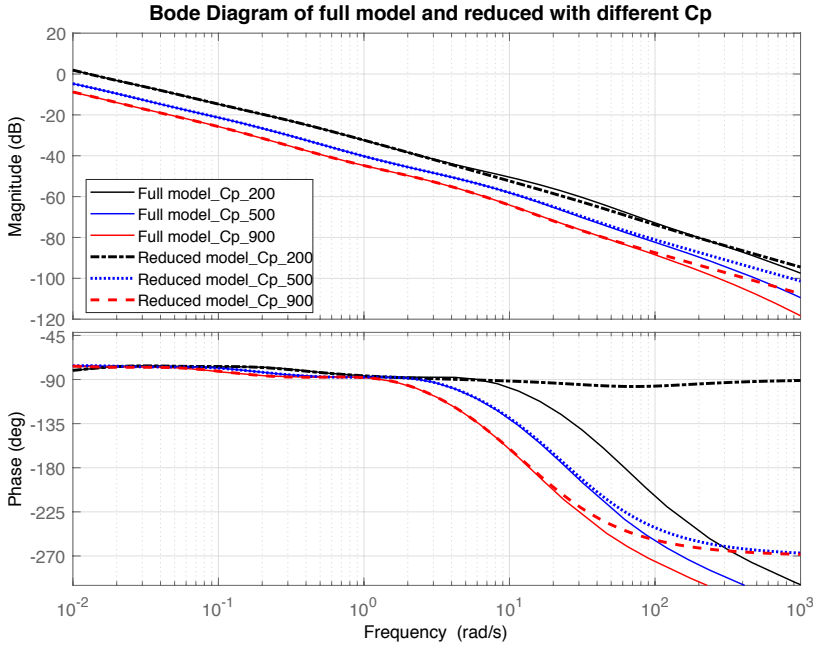
with  $x_i$  a representative point in the  $i$ th finite element in the configuration of Figure 2.2. Thus, the state dimension  $n = 1560$ .

The design goal for this linear motor is to find a suitable material (i.e. with suitable heat capacity  $c_p$ ) which leads to desirable thermo-dynamical properties of the linear motor. We consider an operating range  $c_p \in [200, 900]$  in units  $[J/kg.K]$  and aim to reduce the complexity of the lumped and parametrized model to more manageable proportions. The operating frequency range of the thermal behavior is mainly located at low frequency (near  $1 \sim 10$  Hz). Under these conditions thereupon, we choose frequencies  $s_0 \in \mathcal{S} := \{0, 1 + i, 1 - i\}$  and parameter  $\theta_0 \in \Theta := \{200, 500, 900\}$ .

Following the proposed method, we obtain the reduced model with  $r = 12$  orders which includes 6 expansion points  $[0, 200]$ ,  $[0, 500]$ ,  $[0, 900]$ ,  $[1 \pm i, 200]$ ,  $[1 \pm i, 500]$ ,  $[1 \pm i, 900]$  (The complex conjugate pairs are considered as one pair of expansion points). We preserve two moments for each point and obtain the reduced model as

$$\begin{cases} E_r(\theta)\dot{x}_r = A_r x_r + B_r u \\ y = C_r x_r \end{cases}, \quad \text{with } \dim(x_r) = 12 \quad (2.44)$$

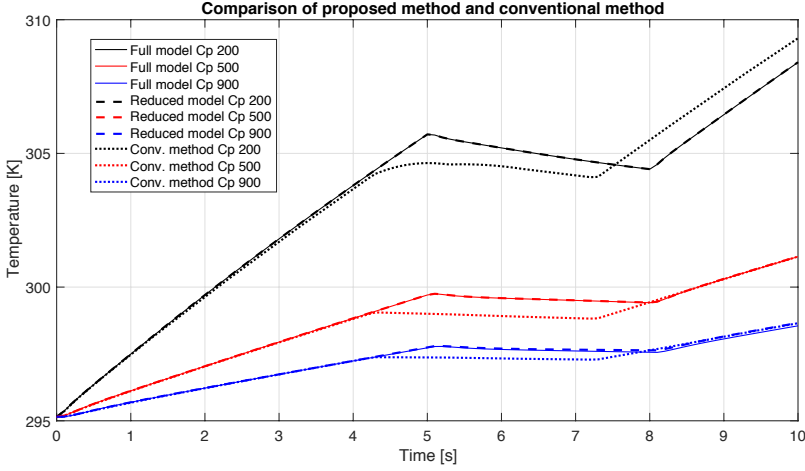
here  $E_r(\theta) = E_{r,0} + c_p E_{r,1}$ . As demonstrated in Figure 2.3, the solid lines represent the full model with different values of  $c_p$ . The dashed lines denote the reduced order model with corresponding  $\theta$ . The frequency range  $[0.1, 10]$ [rad/sec] in the Bode plot for all three models are matched, and there are some mismatches from 20 [rad/sec].



**Figure 2.3:** Bode diagram of full model and reduced model with  $c_p = [200, 500, 900]$ [J/kg · K].

To show the benefits of the separation of frequencies and parameters in the construction of the projection matrices  $V$  and  $W$ , we also provide a comparison of time domain simulations of the ROM using the proposed method and the ROM using the method proposed in [53] and the FOM, the COMSOL simulation results which are considered as reference are also included. By applying the same input sequences  $u(t)$ ,  $t > 0$ , Figure (2.4) shows the outputs of all 9 models with different heat capacities.

In Figure (2.4), the solid lines represent the reference of all models with different  $\theta$ . The dashed lines are the results of the proposed method, the dotted lines are



**Figure 2.4:** Comparison of full model and reduced model with different heat capacities  $c_p = [200, 500, 900][J/kg \cdot K]$  in time domain simulation.

methods from previous work [170] (we choose the same expansion points and same number of moments as the proposed method). Obviously, the proposed method has the better performance which is very close to the reference. The previous method also can match most of the thermal behavior except for the transient parts.

## 2.6 Conclusion

This chapter considers the moment matching problem for linear parameterized systems. An explicit algorithm is proposed that establishes a perfect match of moments in a desired set of frequency points and a desired set of parameter samples. An error bound has been derived that represents the local accuracy of the transfer function of the reduced order model nearby the points  $\mathcal{S} \times \Theta$ . Simulation results from both time domain and frequency domain show that the proposed method delivers good matching and outperforms the previous work. An important open question is how to choose the optimal expansion points with or without prior knowledge.

# Model reduction for LPV systems through parameter projection

---

**F**or affine linear parameter-varying (LPV) systems, this chapter develops a method that reduces the dimension of its parameter space. The method first transforms these affine LPV systems into a parameter-dependent form. Then, leveraging a newly defined parameter-dependent Hankel norm, the parameter reduction can be achieved by analyzing the importance of each parameter in this norm and computing the corresponding singular values. Simulation results of an academic example and a thermal model are discussed before drawing conclusions.

---

## 3.1 Introduction

In the past decade, the class of linear parameter-varying (LPV) systems has been developed and established as a reliable and very efficient model class for specific classes of nonlinear systems, representing parametric uncertainty and gain scheduling purposes. Many successful applications ranging from very-large-scale integration (VLSI) to aircraft designs [29, 116, 62] have been based on implementations of the LPV framework. The inherent complex nature of physical systems often results in high dimensional models with large dimensional state spaces and large dimensional parameter spaces. Typically, the dimension of the parameter space grows as the complexity of the system increases. In practice, it is often necessary to evaluate system performance over substantial ranges of parameter values. We find this theme in problems where design parameters need to be tuned,

calibration problems, geometrical optimization, optimization of material properties in circuits [53] and microelectromechanical systems (MEMS) devices [21] and robustness analyses in control systems. To have a reasonable computational complexity in terms of synthesis and simulation, model order reduction for parametrised systems (pMOR) is often required.

In the aforementioned applications, a high dimensional state space model is often derived from a high resolution spatial discretization of partial differential equations (PDEs) that define the model. Typically, if high precision is required, this process results in many first order ordinary differential equations (ODEs) approximating the solutions of the PDE. Moreover, the physical constraints and design parameters of such high-fidelity models lead to large dimensional parameter and state spaces. A number of relevant approximation problems can then be phrased as follows:

1. the *state reduction problem* involves the reduction of the dimension of the state space, while preserving accuracy of the input-output behavior and the physical meaning of the parameters.
2. the *state and parameter reduction problem* involves the simultaneous reduction of the dimension of the state space and the dimension of the parameter space.
3. the *parameter reduction problem* involves the reduction of the dimension of the parameter space only.

In most pMOR work, the primary goal has been to solve the *state reduction problem*. A number of such methods have been proposed, mainly focusing on sampling techniques [130, 28, 42]. On this topic, two influential survey papers [28, 29] have appeared. Methods based on moment matching of the parametrised transfer function are of particular interest, but tend to become complex with the number of matched moments. For example, in [28], the dimension of the reduced model increases exponentially as the number of parameter interpolation points and moments increase. One solution to this problem is to interpolate both the state space and the parameter state within the predefined variation range [113]. Even though it is not always stated, simple sampling schemes imply or assume static dependence on the parameter. Techniques which explicitly deal with time-varying parameters are more involved [174]. All these works are confined to state reduction without considering the problem to reduce the number of parameters. Approaches which consider parameter reduction either require typical trajectories of the parameter to be known a priori [140], lack interpretation, are limited in application [158] or only focus on static state reduction [110]. The development of more general parameter reduction techniques can significantly improve the efficiency of simulations, often without loss of generality, as was shown in [72]. This leads to the second challenging problem: the *state and parameter reduction problem*. In [72] a two-step approach is introduced. First, parameter reduction is employed to find a low-dimensional parameter space. The second step amounts to constructing a state reduction via moment-matching. However, the reduced rank regression method used in the first step only quantifies the relation between the parameters and the outputs, and is limited by the type of input excitation used.

Besides, the system dynamics are not taken into account over ranges of parameters. Given the fact that parameter spaces are usually determined by the physics and the design constraints, it is relevant to explore the correlation between the parameter space and system theoretical properties such as reachability and observability of parametrised systems. In [110], a joint Gramian based method is proposed such that the parameter and state reduction can be performed via singular value decomposition (SVD), and the results show the joint Gramian based method can be a viable alternative for reducing the parameter complexity. Nevertheless, the system structure or parameter dependency is not exploited in this work which potentially could lead to a different formulation.

Our intention, therefore, is to focus on the *parameter reduction problem*. We exploit the relationship between the parameter space and system Gramian for an affine linear parameter-varying system. A projection-based method is proposed as the means to reduce the dimension of the parameter space using Hankel-norm approximation. In doing so, some definitions of system norms for LPV systems are introduced, aimed at characterizing approximation errors in a consistent manner. The methods in this paper are developed for time-invariant parameters with an extension of the Gramian based approach to the time-varying rate-bounded case.

This chapter is organised as follows. In Section 3.2 we provide a brief introduction to LPV systems and introduce the system norms that will be used in this paper. After a formalisation of the problem, an argument for affine Gramian reduction is presented. Next, the cross-correlation among the parameter space in frequency domain is introduced in Section 3.3 as well as the approximation error for parameter-dependent systems. In Section 3.4, the results are illustrated in an academic example and in a real application of a thermal model which consists of several interconnected components. Conclusions and future work are presented in Section 3.5.

## 3.2 Preliminaries and notation

Consider a system  $\Sigma(\theta)$  defined by the following LPV state-space representation

$$\Sigma(\theta) : \begin{cases} \dot{x}(t) = A(\theta)x(t) + B(\theta)u(t) \\ y(t) = C(\theta)x(t) + D(\theta)u(t) \end{cases}, \quad (3.1)$$

where  $x(t) \in \mathbb{R}^n$  is the state variable,  $u(t) \in \mathbb{R}^m$  and  $y(t) \in \mathbb{R}^q$  denote the input and output of the system, respectively. Furthermore,  $\theta$  represents the parameter that is assumed to reside in the parameter space  $\Theta \subseteq \mathbb{R}^\ell$ . Usually, an important distinction is made between systems where  $\theta$  is time-variant and time-invariant. The state space matrices are assumed to have affine dependence on  $\theta$ , i.e., the

system matrices  $A(\theta), B(\theta), C(\theta), D(\theta)$  are given by

$$\begin{aligned} A(\theta) &= A_0 + A_1\theta_1 + \cdots + A_\ell\theta_\ell, \\ B(\theta) &= B_0 + B_1\theta_1 + \cdots + B_\ell\theta_\ell, \\ C(\theta) &= C_0 + C_1\theta_1 + \cdots + C_\ell\theta_\ell, \\ D(\theta) &= D_0 + D_1\theta_1 + \cdots + D_\ell\theta_\ell \end{aligned} \quad (3.2)$$

where  $A_i, B_i, C_i, D_i$  are real matrices of compatible dimensions, and where  $\theta_i$  denotes the  $i$ th coefficient/component of the vector  $\theta$ . Without loss of generality the system is assumed to be scaled such that  $\theta_i \in [0, 1], i = 1, \dots, \ell$ . Since  $\theta$  does not vary with time both the transfer function

$$G(\theta, s) = C(\theta)(sI - A(\theta))^{-1}B(\theta) + D(\theta); \quad s \in \mathbb{C}$$

and the impulse response

$$H(\theta, t) = \begin{cases} 0 & ; t < 0 \\ D(\theta) & ; t = 0 \\ C(\theta) \cdot e^{A(\theta)t} \cdot B(\theta) + D(\theta) & ; t > 0 \end{cases} \quad (3.3)$$

are well defined for all  $\theta \in \Theta$ . In addition, the system in (3.1) is assumed to be stable  $\forall \theta \in \Theta$  in the sense that  $A(\theta)$  is Hurwitz matrix for all  $\theta \in \Theta$ .

Let  $\mathcal{T}$  be a subspace of  $\mathbb{R}^\ell$  of dimension  $r < \ell$ . A projector onto  $\mathcal{T}$  is a function  $\Pi_{\mathcal{T}} : \mathbb{R}^\ell \rightarrow \mathbb{R}^\ell$  with the property that  $\text{im}(\Pi_{\mathcal{T}}) = \mathcal{T}$  and  $\Pi_{\mathcal{T}}^2 = \Pi_{\mathcal{T}}$ . In particular, if  $\{\phi_1, \dots, \phi_r\}$  is a basis of  $\mathcal{T}$  and  $T = [\phi_1, \dots, \phi_r] \in \mathbb{R}^{\ell \times r}$  then  $\Pi_{\mathcal{T}} = T(T^{-1}T)^{-1}T^T$  is a matrix representation of a projector onto  $\mathcal{T}$ . We call  $I - \Pi_{\mathcal{T}}$  the complementary projector onto  $\ker(\Pi_{\mathcal{T}})$ . It is easily seen that  $I - \Pi_{\mathcal{T}} = \Pi_{\mathcal{T}^\perp}$  with  $\mathcal{T}^\perp$  the orthogonal complement of  $\mathcal{T}$  in  $\mathbb{R}^\ell$ . Thus, for any  $\theta \in \mathbb{R}^\ell$  we have

$$\theta = \Pi_{\mathcal{T}}\theta + (I - \Pi_{\mathcal{T}})\theta = \theta_{\mathcal{T}} + \theta_{\mathcal{T}^\perp} \quad (3.4)$$

where we define

$$\theta_{\mathcal{T}} := \Pi_{\mathcal{T}}\theta \quad \text{and} \quad \theta_{\mathcal{T}^\perp} := (I - \Pi_{\mathcal{T}})\theta = \Pi_{\mathcal{T}^\perp}\theta. \quad (3.5)$$

Note that  $\theta_{\mathcal{T}} \in \mathbb{R}^\ell$  and  $\theta_{\mathcal{T}^\perp} \in \mathbb{R}^\ell$  and that, in the basis  $\{\phi_1, \dots, \phi_r\}$  of  $\mathcal{T}$  one can represent  $\theta_{\mathcal{T}}$  as

$$\theta_{\mathcal{T}} = T(T^T T)^{-1}T^T \theta = T\theta_r \quad (3.6)$$

where  $\theta_r \in \mathbb{R}^r$  are the coordinates of  $\theta_{\mathcal{T}}$  in the basis of  $\{\phi_1, \dots, \phi_r\}$  of  $\mathcal{T}$ . So,  $\theta_r$  has dimension  $r < \ell$ ,  $\theta_{\mathcal{T}}$  has dimension of  $\ell$ . With  $\theta$  decomposed as  $\theta = \theta_{\mathcal{T}} + \theta_{\mathcal{T}^\perp}$  any such projector defines a second system  $\Sigma(\theta_{\mathcal{T}})$  that we view as a parameter-reduced system. Specifically, with  $r = \dim(\mathcal{T})$  set

$$\Sigma_r(\theta_r) := \Sigma(\theta_{\mathcal{T}}) = \Sigma(T\theta_r). \quad (3.7)$$



Then,  $\Sigma_r(\theta_r)$  has representation

$$\Sigma_r(\theta_r) : \begin{cases} \dot{x}_r(t) = A_r(\theta_r)x_r(t) + B_r(\theta_r)u(t) \\ y_r(t) = C_r(\theta_r)x_r(t) + D_r(\theta_r)u(t) \end{cases} \quad (3.8)$$

where

$$\begin{aligned} A_r(\theta_r) &= A(T\theta_r), & B_r(\theta_r) &= B(T\theta_r), \\ C_r(\theta_r) &= C(T\theta_r), & D_r(\theta_r) &= D(T\theta_r) \end{aligned}$$

and simply observe that, because  $\{A(\theta), B(\theta), C(\theta), D(\theta)\}$  are affine functions of  $\theta$  (3.2), also  $\{A_r(\theta_r), B_r(\theta_r), C_r(\theta_r), D_r(\theta_r)\}$  are affine functions of  $\theta_r$  that can therefore be written as explicit expansions

$$\begin{aligned} A_r(\theta_r) &= A(T\theta_r) = A_r^0 + A_r^1\theta_r^1 + \dots + A_r^r\theta_r^r \\ B_r(\theta_r) &= B(T\theta_r) = B_r^0 + B_r^1\theta_r^1 + \dots + B_r^r\theta_r^r \\ C_r(\theta_r) &= C(T\theta_r) = C_r^0 + C_r^1\theta_r^1 + \dots + C_r^r\theta_r^r \\ D_r(\theta_r) &= D(T\theta_r) = D_r^0 + D_r^1\theta_r^1 + \dots + D_r^r\theta_r^r \end{aligned} \quad (3.9)$$

with  $A_r^0 = A_0, B_r^0 = B_0, C_r^0 = C_0, D_r^0 = D_0$  in (3.2). Note that the number of parameters is reduced from  $\ell$  in (3.2) to  $r$  above in (3.9). More specifically, in the basis  $\{\phi_1, \dots, \phi_r\}$  and under the assumption that  $\{\phi_1, \dots, \phi_r, \phi_{r+1}, \dots, \phi_\ell\}$  of  $\mathcal{T}$  the system and  $\theta_r^i$  as the  $i$ th component of  $\theta_r$ , i.e.,

$$\theta_r^i = \langle \theta_r, e_i \rangle, \quad \text{for } e_i \text{ the } i\text{-th unit in } \mathbb{R}^r \quad (3.10)$$

here  $\langle \cdot, \cdot \rangle$  denoting the inner product in  $\mathbb{R}^r$ . More specifically, if  $T = (\phi_1, \dots, \phi_\ell)$  and the collection of vectors  $\{\phi_1, \dots, \phi_\ell\}$  spans  $\mathcal{T}$ , then

$$\begin{aligned} A_r(\theta_r) &= A(T\theta_r) = A_0 + A_1\langle T\theta_r, e_1 \rangle + \dots + A_\ell\langle T\theta_r, e_\ell \rangle \\ &= A_0 + A_1\langle \theta_r, T^T e_1 \rangle + \dots + A_\ell\langle \theta_r, T^T e_\ell \rangle. \end{aligned}$$

Now, let  $T_{ij}$  denote the  $(i, j)$ th entry of  $T$  ( $i = 1, \dots, \ell, \quad j = 1, \dots, r$ ). Then the above expansion reads

$$A_0 + A_r^1\theta_r^1 + \dots + A_r^r\theta_r^r$$

where

$$A_r^j = A_1T_{1,j} + \dots + A_\ell T_{\ell,j} \quad \text{for } j = 1, 2, \dots, r.$$

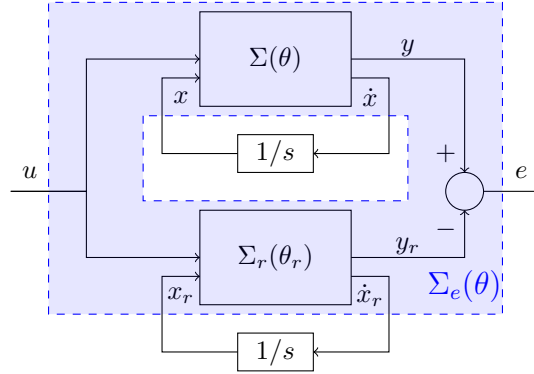
Similarly,  $B_r^0 = B_0, C_r^0 = C_0$  and  $D_r^0 = D_0$

$$\begin{aligned} B_r^j &= B_1T_{1,j} + \dots + B_\ell T_{\ell,j} \\ C_r^j &= C_1T_{1,j} + \dots + C_\ell T_{\ell,j} \\ D_r^j &= D_1T_{1,j} + \dots + D_\ell T_{\ell,j} \end{aligned}$$

In the same way, we can obtain  $B_r(\theta_r), C_r(\theta_r), D_r(\theta_r)$  as defined in (3.9).

Note that the reduced system matrices  $\{A_r(\theta_r), B_r(\theta_r), C_r(\theta_r), D_r(\theta_r)\}$  have the same dimension as the matrices  $\{A(\theta), B(\theta), C(\theta), D(\theta)\}$  of the original system,

but their affine expansions as function of the parameter have been reduced from a length  $(\ell + 1)$  expansion to a length  $(r + 1)$  expansion in each and every state space matrix. To evaluate the performance of the reduced system, an error system is defined as the LPV system with input  $u$  and output  $y - y_r$  with  $y$  and  $y_r$  the outputs of (3.1) and (3.8). See Figure 3.1. This system is compactly denoted as  $\Sigma_e(\theta) = \Sigma(\theta) - \Sigma_r(\theta_r)$  where  $\theta_r$  is defined in (3.6). As such,  $\Sigma_e(\theta)$  is viewed as an error system in the parameter  $\theta$  only. Note that if the projection space  $\mathcal{T}$  is understood, the error system is affinely dependent on the parameters.



**Figure 3.1:** Interconnection of the reduced parameter error system, indicated by the blue area.

In the LTI setting, the error system can be evaluated by many well established and computable norms. Among these, the  $H_\infty$ -,  $H_2$ - and Hankel-norm are commonly used in model reduction. In this work, we introduce a composite error measure on the LPV system that consists of a norm over the parameter space and a system norm

$$\|\Sigma_e\|_{p_\infty, H} := \max_{\theta \in \Theta} \|\Sigma_e(\theta)\|_H, \quad (3.11)$$

which evaluates the maximal Hankel-norm of the system  $\Sigma(\theta)$  when ranging over the feasible parameter space, where the definition of Hankel-norm for parameter-dependent systems is given in the next section. Based on what we have discussed so far, we give the problem formulation of parameter reduction for LPV systems.

*Problem:* Given a stable and minimal system  $\Sigma(\theta)$  as defined in (3.1), find a  $\mathcal{T}$  of dimension  $r$  that is minimizing the righthand side of the (3.12)

$$\|\Sigma_e\|_{p_\infty, H} = \max_{\theta \in \Theta} \|\Sigma(\theta) - \Sigma(\Pi_{\mathcal{T}}\theta)\|_H \quad (3.12)$$

over all subspace  $\mathcal{T}$ . Here  $\Sigma(\Pi_{\mathcal{T}}\theta)$  denotes the parameterized reduced system that is defined in (3.7).

### 3.3 Hankel-norm reduction

It is well known that the Hankel norm of a stable LTI system can be expressed in terms of reachability and observability Gramians [13]. We establish a similar result for LPV systems first. The Hankel operator associated with a stable LTI system  $\Sigma$  is defined by

$$\begin{aligned} \mathcal{H} : \mathcal{L}_2(\mathbb{R}_-, \mathbb{R}^m) &\mapsto \mathcal{L}_2(\mathbb{R}_+, \mathbb{R}^q), u_- \mapsto y_+ \\ &\text{where} \\ y_+(t) := \mathcal{H}(u_-)(t) &= \int_{-\infty}^0 H(t - \tau)u(\tau)d\tau, \quad \text{with } t \in \mathbb{R}_+ \end{aligned} \quad (3.13)$$

it maps the past inputs  $u_-$  into future outputs  $y_+$ . The Hankel norm of the system  $\Sigma$ , denoted  $\|\Sigma\|_H$  is the  $\mathcal{L}_2$  induced norm of  $\mathcal{H}$  and is defined as

$$\|\mathcal{H}\|_{\mathcal{L}_2\text{-ind}} = \sup_{\|u_-\|_2} \frac{\|y_+\|_2}{\|u_-\|_2} = \|\Sigma\|_H. \quad (3.14)$$

The quantity  $\|\Sigma\|_H$  is the Hankel norm of the system  $\Sigma$  and equals the spectral norm  $\|\Sigma\|_H = \sigma_{\max}(\mathcal{H})$ . The Hankel singular values of the system  $\Sigma$  are defined as the singular values of the Hankel operator  $\mathcal{H}$  associated with  $\Sigma$ , and are given as follows

**Lemma 3.1** *Given a reachable, observable and stable LTI system  $\Sigma$  of dimension  $n$ , the Hankel singular values  $\sigma_i(\Sigma)$  are equal to the square roots of the eigenvalues of the product of  $\mathcal{P}\mathcal{Q}$*

$$\sigma_i(\Sigma) = \sqrt{\lambda_i(\mathcal{P}\mathcal{Q})}, i = 1, \dots, n \quad (3.15)$$

where  $\mathcal{P}$  and  $\mathcal{Q}$  are the controllability Gramian and observability Gramian of  $\Sigma$ .

For a time-invariant LPV system, the Hankel operator is defined as

$$\begin{aligned} \mathcal{H}_\theta : \Theta \times \mathcal{L}_2(\mathbb{R}_-, \mathbb{R}^m) &\mapsto \mathcal{L}_2(\mathbb{R}_+, \mathbb{R}^q), (\theta, u_-) \mapsto y_+ \\ &\text{where} \\ \mathcal{H}_\theta(\theta, u_-)(t) &= \int_{-\infty}^0 H(\theta, t - \tau)u(\tau)d\tau, \quad \text{with } t \in \mathbb{R}_+ \quad \text{and } \theta \in \Theta, \end{aligned} \quad (3.16)$$

it maps the past inputs and parameters into future outputs. It follows that the  $\mathcal{L}_2$ -induced norm of  $\mathcal{H}_\theta$  is now parameter dependent. Since  $\Sigma(\theta)$  is assumed to be stable  $\forall \theta \in \Theta$ , the Hankel norm of  $\Sigma(\theta)$  can be expressed as

$$\|\Sigma(\theta)\|_H = \sigma_{\max}(\mathcal{H}_\theta) = \sqrt{\lambda_{\max}(\mathcal{P}(\theta)\mathcal{Q}(\theta))} \quad \text{with } \theta \in \Theta \quad (3.17)$$

where  $\mathcal{P}(\theta)$ ,  $\mathcal{Q}(\theta)$  are the reachability and observability Gramians, defined as the

unique solutions of the parametrized Lyapunov equations

$$A(\theta)\mathcal{P}(\theta) + \mathcal{P}(\theta)A^T(\theta) + B(\theta)B^T(\theta) = 0 \quad (3.18)$$

$$A(\theta)^T \mathcal{Q}(\theta) + \mathcal{Q}(\theta)A(\theta) + C(\theta)^T C(\theta) = 0. \quad (3.19)$$

Finding an exact solution to the Lyapunov equation for the whole parameter space is not trivial and often computationally intractable. In the literature, a static Gramian [148] is proposed which often suffices, but leads to conservative solutions. For parameter dimension reductions, the parameter dependent Gramian is necessary as they express changes in the system due to parameter variations.

The following result shows that a relaxation of (3.18) to an inequality naturally leads to an upper bound on  $\mathcal{P}(\theta)$  and  $\mathcal{Q}(\theta)$  for all  $\theta$ . Subsequently, an upper bound of the Hankel norm (3.17) is derived.

**Theorem 3.2** *Consider an LPV system (3.1) that is stable, reachable and observable for all  $\theta \in \Theta$ . There exist unique solutions  $\mathcal{P}(\theta) = \mathcal{P}^T(\theta) \succeq 0$ ,  $\mathcal{Q}(\theta) = \mathcal{Q}^T(\theta) \succeq 0$  which satisfies (3.18) and (3.19). Furthermore, there exists  $P(\theta) = P^T(\theta) \succeq 0$  and  $Q(\theta) = Q^T(\theta) \succeq 0$  which satisfy*

$$A(\theta)P(\theta) + P(\theta)A^T(\theta) + B(\theta)B^T(\theta) \preceq 0. \quad (3.20)$$

$$A(\theta)^T Q(\theta) + Q(\theta)A(\theta) + C(\theta)^T C(\theta) \preceq 0. \quad (3.21)$$

for all  $\theta \in \Theta$ . All solutions  $\{P(\theta), Q(\theta)\}$  upper bound  $\{\mathcal{P}(\theta), \mathcal{Q}(\theta)\}$  in the sense that  $P(\theta) \succeq \mathcal{P}(\theta) \succeq 0$  and  $Q(\theta) \succeq \mathcal{Q}(\theta) \succeq 0$ . Moreover, the Hankel norm of  $\Sigma(\theta)$  is upper bounded by

$$\begin{aligned} \|\Sigma(\theta)\|_H &= \sqrt{\lambda_{\max}(\mathcal{P}(\theta)\mathcal{Q}(\theta))} \\ &\leq \sqrt{\lambda_{\max}(P(\theta)Q(\theta))}. \end{aligned} \quad (3.22)$$

The proof of the Lyapunov inequality in (3.20) and (3.21) is given in Appendix 3.A and the proof of Hankel Singular value in (3.22) is given in Appendix 3.B. Note that solutions of (3.20) and (3.21) are not unique.

For every sample point of  $\theta \in \Theta$ , there is one solution of (3.18) and (3.19). Therefore, it is of interest to find an upper bound of such solutions for  $P(\theta)$  and  $Q(\theta)$ ,  $\theta \in \Theta$ . Particularly, we are interested in an upper bound with affine structure.

**Theorem 3.3** *Consider an LPV system defined by (3.1) that is stable, reachable and observable for all  $\theta \in \Theta$ . Suppose that the parameter space is a compact set:  $\Theta = \text{Co}\{w_1, \dots, w_k\}$ . Then, for  $\mathcal{P}(\theta)$  which satisfies (3.18) there exists an affine function  $f : \Theta \rightarrow \mathbb{R}^{n \times n}$  such that*

$$f(\theta) \succeq \mathcal{P}(\theta) \text{ for all } \theta \in \Theta. \quad (3.23)$$

This theorem can be proved by making use of the convexity of affine functions  $f(\theta)$  with convex hull of  $\theta \in \Theta$ . The proof of the upper bound affine function for observability Gramian is given in Appendix 3.C. In the same manner, the upper bound affine function for controllability Gramian can be obtained.

### 3.3.a Parameter reduction using Hankel-norm approximation

The main ingredients of parameter reduction for LPV systems are to construct an upper bounded Hankel norm defined in *Theorem 3.2* which satisfies the affine form in *Theorem 3.3*, and then to find a reduced model with the least error for the given reduced parameter space.

We start with constructing a proper affine upper bounding controllability Gramian

$$P(\theta) = P_0 + P_1\theta_1 + \cdots + P_\ell\theta_\ell,$$

here each  $P_i$  is the solution of (3.20) for the  $\Sigma(\theta)$  in (3.2) where all parameters  $\theta_j$  except for  $\theta_i$  are zero  $\theta_j = 0, j \neq i$ . Notice that the  $P_0$  is the solution of (3.20) for the that all parameters are equal to zero  $\theta = 0$ . In such a way, the upper bounding Gramian with affine format is obtained. In the same manner, the affine upper bounding observability Gramian

$$Q(\theta) = Q_0 + Q_1\theta_1 + \cdots + Q_\ell\theta_\ell,$$

$Q(\theta)$  satisfies (3.21). With these affine forms, the Gramian can be rewritten as

$$\begin{aligned} P(\theta)Q(\theta) &= \bar{\theta}^T K \bar{\theta} \\ &= \underbrace{\begin{bmatrix} \theta_1 I_n \\ \vdots \\ \theta_\ell I_n \\ I_n \end{bmatrix}}_{\bar{\theta}^T} \underbrace{\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}}_K \underbrace{\begin{bmatrix} \theta_1 I_n \\ \vdots \\ \theta_\ell I_n \\ I_n \end{bmatrix}}_{\bar{\theta}} \end{aligned} \quad (3.24)$$

where  $I_n$  denotes identity matrix with dimension of  $n \times n$ . The  $K_{12} \in \mathbb{R}^{(\ell \cdot n) \times n}$ ,  $K_{21} \in \mathbb{R}^{n \times (\ell \cdot n)}$  and  $K_{22}$  a matrix with  $n \times n$ . Note that  $K_{22}$  is the a constant matrix.  $K_{12}$  and  $K_{21}$  are the matrices that denote the linear term of  $\theta$ . And  $K_{11} \in \mathbb{R}^{(\ell \cdot n) \times (\ell \cdot n)}$  represents all quadratic terms of  $\theta$ .

With the upper bounding Gramian proposed in (3.24), the system norm defined in (3.11) for the full-order LPV system  $\Sigma(\theta)$  (3.1) satisfies

$$\begin{aligned} \|\Sigma(\theta)\|_{p_\infty, H}^2 &= \max_{\theta \in \Theta} \|\Sigma(\theta)\|_H^2 \\ &= \max_{\theta \in \Theta} \lambda_{\max}(\mathcal{P}(\theta)Q(\theta)) \\ &\leq \max_{\theta \in \Theta} \lambda_{\max}(P(\theta)Q(\theta)) \\ &= \max_{\theta \in \Theta} \lambda_{\max}(\bar{\theta}^T K \bar{\theta}) \end{aligned} \quad (3.25)$$

In the same manner, the system norm for reduced-order LPV system  $\Sigma_r(\theta_r)$  also

satisfies the following inequalities:

$$\begin{aligned}
\|\Sigma_r(\theta_r)\|_{p_\infty, H}^2 &= \|\Sigma(\theta_T)\|_{p_\infty, H}^2 = \|\Sigma(T\theta_r)\|_{p_\infty, H}^2 \\
&= \max_{\theta_r \in \Pi_{\mathcal{T}}\Theta} \lambda_{\max}(\mathcal{P}(T\theta_r)\mathcal{Q}(T\theta_r)) \\
&\leq \max_{\theta_r \in \Pi_{\mathcal{T}}\Theta} \lambda_{\max}(P(T\theta_r)Q(T\theta_r)) \\
&= \max_{\theta_T \in \Theta} \lambda_{\max}(\bar{\theta}_T^T K \bar{\theta}_T) \tag{3.26}
\end{aligned}$$

here  $\bar{\theta}_T = \text{col}(\theta_T, 1)$  and  $\theta_T = \Pi_{\mathcal{T}}\theta = T\theta_r$  denotes the reduced parameter space as defined in (3.5).

As defined in (3.12), the  $\|\cdot\|_{p_\infty, H}$  system norm of the error system between the full-order LPV system  $\Sigma(\theta)$  and the reduced-order LPV system  $\Sigma_r(\theta_r)$  reads

$$\begin{aligned}
\|\Sigma_e(\theta)\|_{p_\infty, H}^2 &= \|\Sigma(\theta) - \Sigma_r(\theta_r)\|_{p_\infty, H}^2 \\
&= \max_{\theta \in \Theta} \|\Sigma(\theta) - \Sigma(\theta_T)\|_H^2. \tag{3.27}
\end{aligned}$$

Notice that the error system is not equal to the difference between their individual norms in (3.25) and (3.26). Instead of finding the exact solution of (3.27), we find it of particular interest to find the upper bound of the error system. With the affine upper bounding Gramian proposed in (3.24) and a given  $\mathcal{T}$ , we can obtain the following

$$\|\Sigma_e(\theta)\|_{p_\infty, H}^2 \leq \max_{\theta \in \Theta} \lambda_{\max}(\bar{\theta}^T K \bar{\theta} - \bar{\theta}_T^T K \bar{\theta}_T) \tag{3.28}$$

Once the upper bounded error system norm is given in (3.28), the problem stated in (3.12) can be reformulated as finding

$$\begin{aligned}
\min_{\mathcal{T}} \quad & \max_{\theta, \theta_r \in \Theta} \lambda_{\max}(\bar{\theta}^T K \bar{\theta} - \bar{\theta}_T^T K \bar{\theta}_T) \\
\text{s.t.} \quad & \bar{\theta}^T K \bar{\theta} = P(\theta)Q(\theta), \quad \bar{\theta}_T^T K \bar{\theta}_T = P(T\theta_r)Q(T\theta_r) \\
& P(\theta) = P_0 + P_1\theta_1 + \dots + P_\ell\theta_\ell, \quad P(T\theta_r) = P_0 + P_r^1\theta_r^1 + \dots + P_r^\ell\theta_r^\ell \\
& Q(\theta) = Q_0 + Q_1\theta_1 + \dots + Q_\ell\theta_\ell, \quad Q(T\theta_r) = Q_0 + Q_r^1\theta_r^1 + \dots + Q_r^\ell\theta_r^\ell \tag{3.29}
\end{aligned}$$

This formulation is used to express the loss in Hankel-norm in quadratic form. Minimising the loss over all parameters gives a min-max optimisation as

$$\min_{\mathcal{T}} \max_{\theta, \theta_r \in \Theta} \|\bar{\theta} - \bar{\theta}_T\|^2 \tag{3.30}$$

for a given reduced parameter space  $\mathcal{T}$  with  $\dim(\mathcal{T}) = r$ . Remark that the explicit expression of  $P(T\theta_r), Q(T\theta_r)$  is given and it is  $T$  dependent. For solving the above min – max optimization problem, the first step is to find an upper bound  $P(\theta), Q(\theta)$  that satisfies *Theorem 3.2* and *Theorem 3.3*. Finding a solution to the Lyapunov inequality in (3.20) and (3.21) is done using linear matrix inequalities (LMIs). The second step is to construct a  $T$  which minimizes the error between the original system and the reduced one. Specifically, with a given order  $r$ , we

build the projection matrix  $T = [\phi_1, \dots, \phi_r]$  in column-wise fashion. The above discussion is briefly summarized in Algorithm 1.

---

**Algorithm 1** : Parameter reduction for affine LPV systems
 

---

- 1: Construct parameter affine dependent upper bounds  $P(\theta), Q(\theta)$  that satisfy Theorem 3.2 and Theorem 3.3.
  - 2: Build the projection matrix  $v_r$  by solving (3.29) for a given  $r$ ,
  - 3: Orthonormalize  $v_r$  (e.g., Modified Gram-Schmidt method) such that  $v_r^T v_r = I_r$ .
  - 4: Assign  $T \leftarrow v_r$ .
  - 5: **Return**  $T$
- 

Notice that the orthonormalization procedure in Step 3 is to obtain a better numerically stable projection matrix, which is not required in the theorem developed in (3.29).

### 3.3.b Approximation error and sensitivity analysis

In LTI approximation the loss function is related to the energy of the error system since the Hankel singular values are a measure of energy in each state respectively. For the proposed method, the interpretation is more nuanced as the affine Gramian constitutes an upper bound on the actual Gramians. If this upper bound is tight, the loss function is a good indication of the error. The tightness of this upper bound is heavily dependent on the system, and determining the tightness is not trivial. Inspired by Adamjan-Arov-Krein (AAK) theorem, we provide a relative  $p_{\infty, \infty}$ -error which asses the approximation error

$$\frac{\|\Sigma(\theta) - \Sigma_r(\theta_r)\|_{p_{\infty, \infty}}}{\|\Sigma(\theta)\|_{p_{\infty, \infty}}}. \quad (3.31)$$

Here  $\|\cdot\|_{p_{\infty, \infty}}$  is defined as

$$\|\Sigma_e(\theta)\|_{p_{\infty, \infty}} := \max_{\theta \in \Theta} \|\Sigma_e(\theta)\|_{H_{\infty}} \quad (3.32)$$

which evaluates the  $H_{\infty}$ -norm of the system  $\Sigma(\theta)$  when varying over the feasible parameter space.

In literature, another commonly used method to evaluate the parametric importance of the LPV systems is sensitivity analysis [158]: the derivative towards the parameters is used to find parameter directions in which the output is most sensitive to changes. In the frequency domain it is expressed as the transfer-sensitivity covariance matrix (TSCM)  $\Pi$

$$\Pi = \begin{bmatrix} \Pi_{1,1} & \cdots & \Pi_{1,\ell} \\ \vdots & \ddots & \vdots \\ \Pi_{\ell,1} & \cdots & \Pi_{\ell,\ell} \end{bmatrix}, \quad (3.33)$$

of which the elements are defined as

$$\Pi_{ij} := \left\| \frac{d}{d\theta_i} H(\theta, s)^T \cdot \frac{d}{d\theta_j} H(\theta, s) \right\|_{p_{\infty, \infty}} \quad (3.34)$$

where  $H(\theta, s)$  is a parameter-dependent transfer function of the system. The resulting TSCM is a positive definitive matrix with a dimension of  $\ell \times \ell$ .

By taking the SVD of this matrix,  $\Pi = USU^*$ , the parameter transformation matrix

$$T = [u_1, \dots, u_r]$$

and  $\mathcal{T} = \text{im}(T)$  is found. This transformation orders the parameter directions in terms of transfer sensitivity covariance and is orthonormal, constituting a valid transformation as defined in (3.6).

## 3.4 Simulation results

To illustrate these methods of parameter reduction through sensitivity analysis and Hankel-norm approximation, two examples systems are used. The first system is an illustrated LPV system and the second a real thermal system.

### 3.4.a illustrative example

Consider an affine LPV system

$$\dot{x} = A_0 x + B_0 u + \sum_{i=1}^5 (A_i \theta_i x + B_i \theta_i u) \quad (3.35a)$$

$$y = C_0 x + D_0 u + \sum_{i=1}^5 (C_i \theta_i x + D_i \theta_i u), \quad (3.35b)$$

where  $\theta_i \in [0, 1]$  with  $i = 1, \dots, 5$  and  $u(t) \in \mathbb{R}^2$ ,  $y(t) \in \mathbb{R}^2$  and  $x(t) \in \mathbb{R}^{45}$ . To ensure stability all  $A_i \prec 0$ . For this system Hankel-norm and TSCM approximations are determined. As a comparison subsystem Hankel-norm approximation is used where reduction is based on singular values of the actual Hankel norm associated with the remaining parameter space.

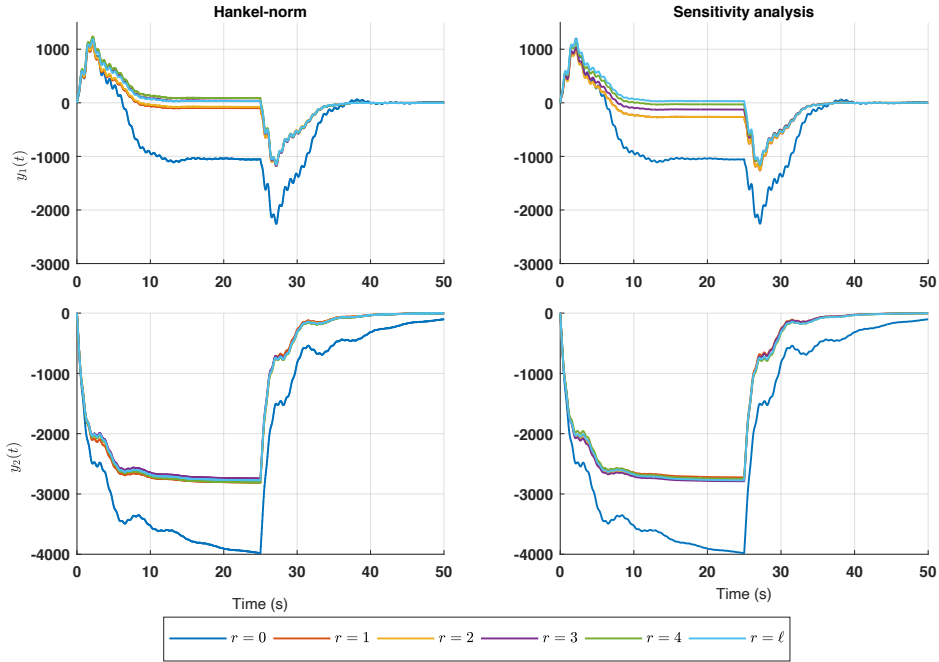
In Figure 3.2 the output evolution which consists of the transient response and the steady-state response is shown for an input  $u(t) = [u_1(t); u_2(t)]$ ,  $u_1(t) = 1$  and  $u_2(t) = 2$  for  $t = [0, 25]$  and  $u(t) = 0$  otherwise. The parameter space is defined as  $\theta \in \Theta$ ,  $\Theta = [0, 1]$  and a random parameter realization for  $\theta_i, i = 1, \dots, 5$  is chosen for both reduction methods.

To illustrate the accuracy of the reduced models, the errors are shown in Figure 3.3. For  $t = [0, 25]$ , both figures show a tendency that the more parameters are



preserved, the smaller the error is between the reduced and the original model. For  $t > 25[s]$ , all reduced models with  $r = 1, \dots, 4$  converge to zero and therefore show a stable system (which was to be expected).

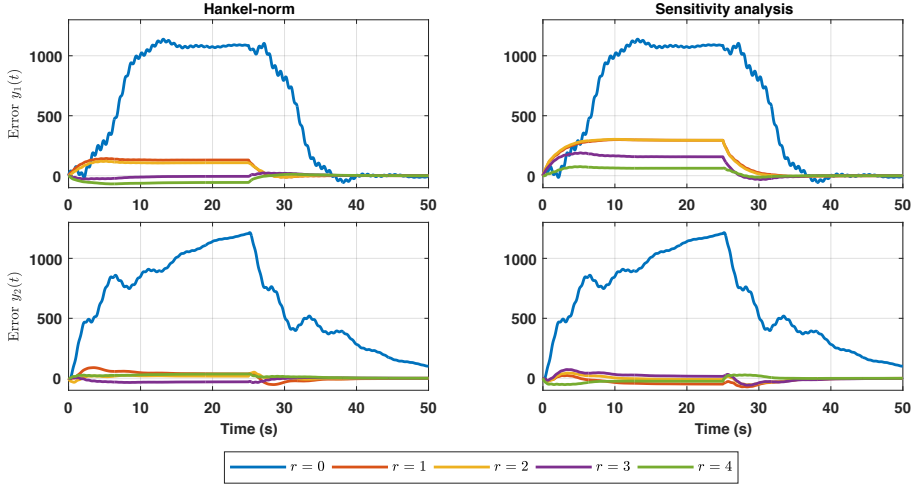
The output of the system changes with a different parameter realization, and therefore Figure 3.4 shows the relative  $p_{\infty, \infty}$  error defined in (3.31) of the reductions to be non-increasing with growing parameter order for both methods. It also shows that for this example, the sensitivity analysis outperforms the Gramian based method.



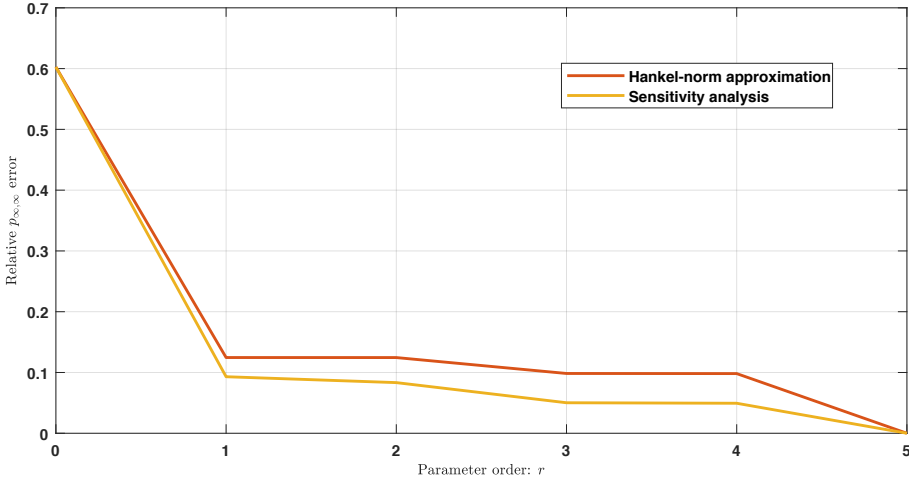
**Figure 3.2:** System evolution of the illustrated LPV system defined in (3.35) with a random parameter realization. Both plots show the outputs  $y_1, y_2$  of the original model and the outputs of the reduced model generated from the proposed Hankel-norm method (left side) and the sensitivity analysis method (right side) with different reduced parameter space of a dimension of  $r = 1, \dots, 4$ .

### 3.4.b Thermal simulation

The second system is a thermal simulation consisting of five coupled metal blocks, all having parameter dependent heat capacity. Using COMSOL Multiphysics the



**Figure 3.3:** Output error  $[y_1(t) - y_{1,r}(t), y_2(t) - y_{2,r}(t)]$  of the randomized LPV system with different reduced parameter space of a dimension of  $r = 1, \dots, 4$  for the proposed Hankel-norm approximation method and the sensitivity analysis method.



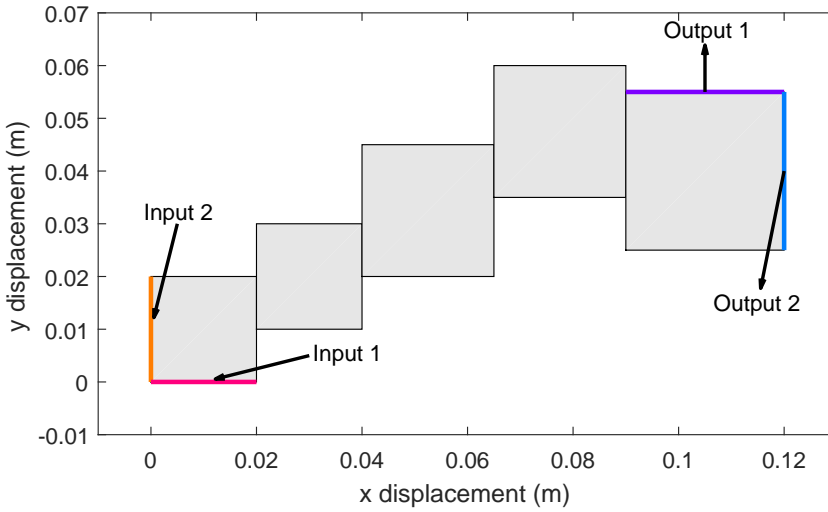
**Figure 3.4:** Relative  $p_{\infty, \infty}$  error defined in (3.31) of the randomized LPV system for the reduced model generated via the Hankel-norm approximation method, and the sensitivity analysis method with different reduced order  $r = 1, \dots, 4$ .

system is generated with the following structure.

$$\dot{x} = \left( A_0 x + \sum_{i=1}^5 A_i \theta_i x \right) + Bu \quad (3.36a)$$

$$y = C_0 x \quad (3.36b)$$

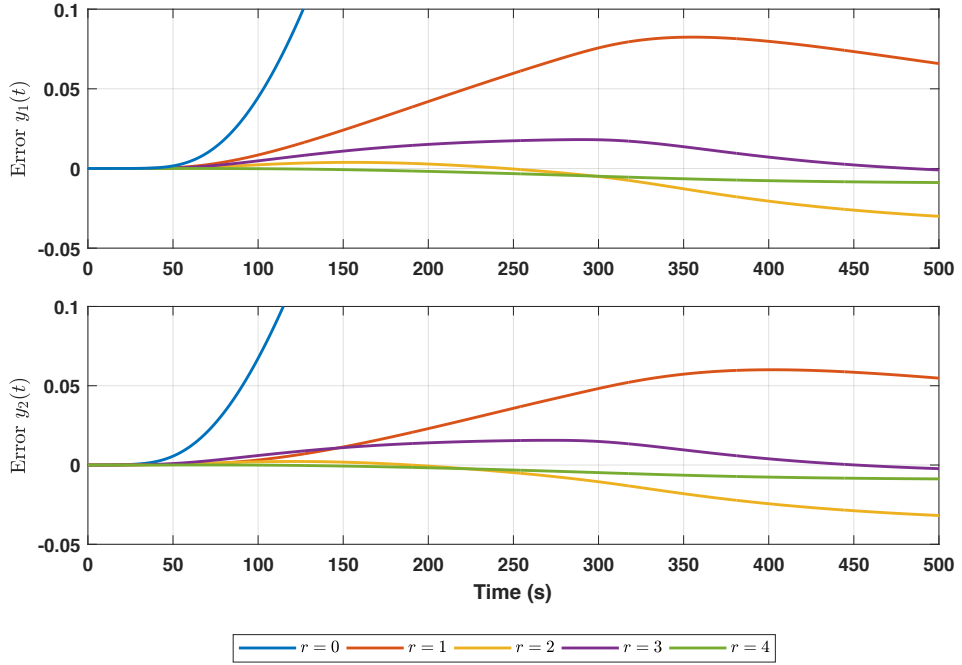
Where  $u \in \mathbb{R}^2$ ,  $y \in \mathbb{R}^2$  and  $x \in \mathbb{R}^{45}$ . The inputs and outputs represent heat power in [W] and temperature [K] respectively. Here the parameter  $\theta_i$ ,  $i = 1, \dots, 5$  represents the conductivity coefficient for 5 metal blocks. In Figure 3.5 an illustration of the system is shown.



**Figure 3.5:** Thermal model with five different material blocks. Orange and pink are heat inputs. Purple and blue are outputs.

To evaluate the error of the reduced model, a simulation is performed with constant power from  $u(t) = [u_1(t); u_2(t)]$ ,  $u_1(t) = 50[W]$  and  $u_2(t) = 45[W]$  for  $t \in [0 \ 250]$  and  $u(t) = [0; 0]$  afterwards. Here all parameters are normalized and the values are  $[\theta_1, \theta_2, \theta_3, \theta_4, \theta_5] = [0.11, 0.26, 0.23, 0.21, 0.24]$ . For different reduced parameter space, the error between the reduced model and full order model is shown in Figure 3.6. Clearly the non-parametric model  $r = 0$  performs the worst showing that the nominal model is not a good approximation of the system. The best performing model is the  $r = 4$  model as the error is almost zero. This result shows the performance with respect to a specific parameter space. Therefore, a global error bound is used to infer conclusions on performance. In Figure 3.7 the relative  $p_{\infty, \infty}$  error is plotted for different reduction orders. This figure illustrates that reducing the system using sensitivity analysis is comparable to reduction in subsystem using Hankel-norm. It is also clear that approximation in the  $p_{\infty, H}$  norm using transformation optimization yields improved results for  $r < 4$ . For approximation of  $r = 4$  the Hankel-norm optimization method performs worse in

$p_{\infty, \infty}$  error. This is due to a combination of issues, the non-convex optimization (3.29) and the error introduced in finding the upper bound of the affine Gramian (3.23).

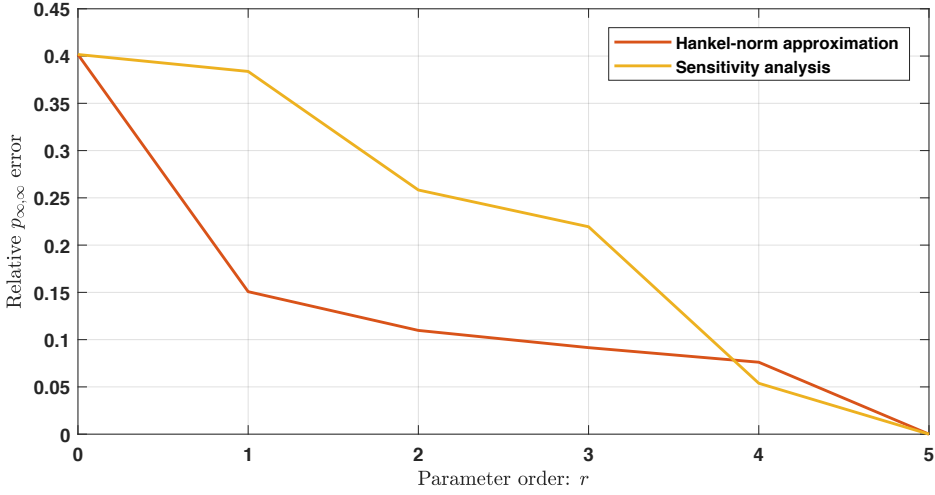


**Figure 3.6:** Simulation error of the thermal system for different parameter orders at a selected parameter in the parameter space.

For both systems computation of the Gramian takes in the order of approximately 1000 seconds (with a dual core PC and using YALMIP [111]). Because the optimization in (3.29) has not shown to be convex, it is using different initial condition to find a close approximation of the Hankel singular values. Due to this non-convexity, the resulting transformation matrix does not guarantee the global optimal approximation.

### 3.5 Conclusion

In this chapter, two methods for parameter reduction have been given for time-invariant LPV systems. The first method approximates the system using the Hankel-norm over the parameter space. The second method uses singular value decomposition on the covariance matrix. A system norm analysis of the performance of these methods has been presented in the chapter together with simulation results. It can be concluded that both methods, though substantially different, are computationally feasible and provide good approximations over the parameter space.



**Figure 3.7:** Relative  $p_{\infty, \infty}$  error of the thermal system for different parameter orders and different reduction methods.

The proposed methods are not limited to be only used for simplifying the complexity of parameters, but can also be integrated with state reduction problems, as joint state and parameter space reduction problems.

### 3.A Proof of upper bound of the Lyapunov function

Consider a stable LTI system with stable matrix  $A$ . Lyapunov equation theorem shows that there is a unique solution  $P \succ 0$  to

$$AP + PA^T + Q = 0 \quad (3.37)$$

where  $Q = BB^T \succ 0$ . The same system admits a solution space  $P \in \tilde{\mathcal{P}}$ , being all solutions that satisfy

$$AP + PA^T + Q \preceq 0 \quad (3.38)$$

The equality can be rewritten and substituted into the inequality to give

$$A(P - \mathcal{P}) + (P - \mathcal{P})A^T \preceq 0 \quad (3.39)$$

From Lyapunov equation theorem and stable  $A$ , we have  $P - \mathcal{P} \succeq 0$  as the solution of the above inequality. Extending to stable, observable and reachable time-invariant LPV systems, it suffices to show that for every  $\theta \in \Theta$  a solution can be found to the Lyapunov (in)equality concluding the proof.  $\square$

### 3.B Proof of the upper bound of Hankel-singular values

From [34] we take two properties of eigenvalue arithmetic of symmetric matrices  $A, B \in \mathbb{R}^{n_x \times n_x}$ ,

$$\begin{aligned}\lambda_i(A + B) &\geq 0 \text{ if } \lambda_i(A) \geq 0 \text{ and } \lambda_i(B) \geq 0, \\ \lambda_i(AB) &\geq 0 \text{ if } \lambda_i(A) \geq 0 \text{ and } \lambda_i(B) \geq 0,\end{aligned}\tag{3.40}$$

for  $i = 1, \dots, n$ . For clarity we drop  $\theta$ , and define a function  $F$  below:

$$F := PQ - \mathcal{P}\mathcal{Q} = (P - \mathcal{P})(Q - \mathcal{Q}) + \mathcal{P}(Q - \mathcal{Q}) + (P - \mathcal{P})\mathcal{Q}.\tag{3.41}$$

Given  $P(\theta) \succeq \mathcal{P}(\theta) \succ 0$  and  $Q(\theta) \succeq \mathcal{Q}(\theta) \succ 0$ , applying properties (3.40) and Theorem 1 to (3.41) yields  $\lambda_i(F) \geq 0$ ,  $i = 1, \dots, n$ . With the properties of Hankel matrix  $PQ = Q^{\frac{1}{2}}PQ^{\frac{1}{2}}$ , for the symmetric  $F = (Q^{1/2}PQ^{1/2} - \mathcal{Q}^{1/2}\mathcal{P}\mathcal{Q}^{1/2}) = F^T$  it is proven. Therefore,  $PQ \succeq \mathcal{P}\mathcal{Q}$  is concluded. Next consider the eigenvector  $x_1$  associated to the largest eigenvalue of  $\mathcal{P}\mathcal{Q}$ . Then the following holds

$$x_1^T \mathcal{P}\mathcal{Q} x_1 \leq x_1^T (PQ) x_1 \Rightarrow \lambda_1(\mathcal{P}\mathcal{Q}) \leq \frac{x_1^T (PQ) x_1}{\|x_1\|_2^2}.\tag{3.42}$$

By the definition of eigenvalue decomposition, we have

$$\frac{x_1^T (PQ) x_1}{\|x_1\|_2^2} \leq \sup_{y_1} \frac{y_1^T (PQ) y_1}{\|y_1\|_2^2} = \lambda_{\max}(PQ),\tag{3.43}$$

here  $y_1$  is the eigenvector associated to the largest eigenvalue of  $PQ$ . Thus, the  $\lambda_{\max}(\mathcal{P}\mathcal{Q}) \leq \lambda_{\max}(PQ)$  is proved.  $\square$

### 3.C Proof of upper bound with affine function $f(\theta)$

**Proof:** Since  $f$  is assumed to be affine in the parameter  $\theta$ , it takes the form

$$f(\theta) = P_0 + P_1\theta_1 + \dots + P_\ell\theta_\ell.$$

For suitable matrices  $P_0, \dots, P_\ell$ . Given  $\mathcal{P}$  satisfying (3.18), define

$$P_i := \sup_{\theta_i, \theta \in \Theta} (\mathcal{P}(\theta_i)).$$

Then since  $\Theta$  is compact there exists  $\theta_i$  assuming the supremum in the sense that  $P_i = \max(\mathcal{P}(\theta_i))$ ,  $i = 0, \dots, \ell$ . With the definition of  $f$ , we achieve that

$$f(\theta) = P_0 + P_1\theta_1 + \dots + P_\ell\theta_\ell \geq P_0 + \theta_1\mathcal{P}(\theta_1) + \dots + \theta_\ell\mathcal{P}(\theta_\ell) = \mathcal{P}(\theta)$$

which yields the result. Thus, Theorem 3.3 is proved.  $\square$

# Parameter calibration using adaptive parameterized model reduction

---

**T**his chapter extends the parameterized model order reduction technique to the parameter calibration problem using measurement data. We prove the equivalent Hermite interpolation condition between the calibration problem using the full-order parameter-dependent model and the parameterized reduced-order model. We propose an adaptive algorithm that iteratively updates parameter expansion points while minimizing the calibration error based on this equivalent condition. Numerical experiments of a real-world application show that the proposed algorithm well preserves the fidelity of the full-order parameter-dependent model for the parameter calibration problem.

---

## 4.1 Introduction

Many applications in engineering and applied sciences are fully dedicated to the question how to make adjustments to a system so as to improve its functionality or performance in a well defined sense. The development of efficient algorithms for automated performance enhancement is of evident interest in areas such as machine learning, design optimization, controller synthesis, statistical learning and system automation, and in questions related to calibration and parameter identification. Typically, these algorithms act on sample data and create or adjust mathematical models to make predictions or decisions on the system by considering the mismatch between model and data. Many of these techniques amount to

adjusting dedicated physical parameters in the model and/or the physical system on the basis of an optimization criterion that reflects this mismatch.

A few fundamental problems in this line of research can be identified. First, although the rigorous or governing models are numerically accessible, the physical parameters that are subject to adaptation may be fully or partially inaccessible. Second, the validated operating range of the model is generally limited and does not cover the full range of parameter values that is of potential interest. Third, the numerical analysis that is performed on a single parameter vector may be computationally expensive, in which case the gathering of insight in parametric changes is numerically demanding or computationally prohibitive.

Models that involve the discretization of partial differential equations are especially demanding in this context. Their complexity increases dramatically with the number of parameters, while increasing demands on accuracy translate into large scale computer models as a result of finer granularities of time and space. One possible solution for this problem amounts to using model order reduction (MOR) techniques in which the original large-scale model is substituted by a low-order simplified model with sufficient accuracy. Model order reduction techniques have led to many successful applications including VLSI simulation by [84], MEMS design by [130] and thermal designs [113]. A drawback of many MOR techniques for parameterized systems is the need to repeat the reduction procedure for each choice of parameters. The need for parametrized model order reduction (pMOR) was therefore introduced in [53]. In comparison to non-parameterised MOR, pMOR methods ideally preserve the character and physical structure of parameters over a given parameter range in the reduced model.

By reducing the complexity of a large-scale model via MOR/pMOR, one can alternatively find the best parameters of a large-scale model by minimizing an error between a reduced order parametric model and data. This requires the construction of fully parameterized reduced order models instead of reducing the models that are obtained by querying the parameter space. That is, the reduction of queried parametrized models is much less attractive than the querying of reduced order parametric models.

Various approaches have been reported in the recent past. In [120, 3], a component mode synthesis was applied for optimal design which only captures the stationary properties of the system. Non-parametrized MOR methods were introduced by [91]. Surprisingly, we also find this theme in biological problems [124] and economic problems suggested by [45]. The reduced order models show an efficient alternative to solve an estimation problem for a large-scale system. However, the proper orthogonal decomposition (POD) method used in both of these papers limits the validity of the reduced model when considering larger ranges of the input space. The authors of [176] proposed a general framework for design optimization using a simplified model generated by MOR method or pMOR method. The focus of this work is on maximizing the performance of the optimization step. However, in both frameworks the construction of the reduced order model is carried out independent of the optimization step. Consequently, this strategy does not impose guarantees on performance of the optimization problem by using the reduced order model. The existing literature to combine the model approxima-



tion method and the parameter optimization/estimation typically does not take the optimization criterion into account in the model reduction step.

The purpose of this chapter is to develop a model reduction strategy for parametrized models in which the cost function of the model misfit is defined in such a way that optimal parameters that minimize the difference between an observed model and the reduced model, explicitly preserves the parameters and a relevant range of frequency responses of the system. We propose a reduction method which guarantees a well-defined equivalence between the objective function using the reduced order model and the full order model. By evaluating the first-order derivatives of the optimization function with respect to the parameters, the proposed method is better geared towards finding optimal parameter values. The present chapter is an extension of our earlier work in [113].

The remainder of the chapter is structured as follows. In Section 4.2 we formalize the problem. In Section 4.3, the equivalence of the original problem and the simplified problem is derived. In Section 4.4, we present the proposed method to solve the optimization problem. Numerical examples are shown in Section 4.5. Conclusions are given in Section 4.6.

## 4.2 Problem formulation

In this chapter, we focus on models in the class of linear time-invariant systems that are parameterized in terms of parameter vectors  $\theta$  that are assumed to be time-independent and reside in a parameter set  $\Theta \subset \mathbb{R}^\ell$ . The transfer function  $\mathcal{H}$  of a large-scale system with  $m$  inputs and  $q$  outputs is assumed to be given and its frequency response at parameter  $\theta \in \Theta$  is given by the mapping  $s \mapsto \mathcal{H}(\theta, s)$  where  $s \in \mathbb{C}$  and  $\mathcal{H}(\theta, s) \in \mathbb{C}^{q \times m}$ . For simplicity, we start from a single-input single-output (SISO) system  $q = m = 1$ .

**Objective function.** The objective function of interest is a quadratic form of the error function, the difference between the frequency response of  $\mathcal{H}$  and the frequency measurement data  $\mathcal{H}^{\text{obs}}(s)$ . Note that the measurement data can be parameter independent  $\mathcal{H}^{\text{obs}}$  or parameter dependent  $\mathcal{H}^{\text{obs}}(\theta)$ <sup>1</sup>. In this chapter, we focus on the parameter independent measurement  $\mathcal{H}^{\text{obs}}$ . Then, the error function is defined as

$$e(\theta, s) = \mathcal{H}(\theta, s) - \mathcal{H}^{\text{obs}}(s). \quad (4.1)$$

For a frequency point  $s_i \in \mathbb{C}$  belonging to a given set  $\mathcal{S}$  of frequency points  $\mathcal{S} := \{s_1, \dots, s_k\} \subset \mathbb{C}$ , we define the matrices

$$\mathbb{H}(\theta) = \text{vec}(\mathcal{H}(\theta, s_i) \mid s_i \in \mathcal{S}) := \begin{bmatrix} \mathcal{H}(\theta, s_1) \\ \vdots \\ \mathcal{H}(\theta, s_k) \end{bmatrix},$$

---

<sup>1</sup>i.e., for the system whose boundary conditions are parameterized and the measurement data captures different boundary conditions, this leads to parameter dependent  $\mathcal{H}^{\text{obs}}(\theta)$ .

and

$$\mathbb{H}^{\text{obs}} = \text{vec}(\mathcal{H}^{\text{obs}}(s_i) \mid s_i \in \mathcal{S}) := \begin{bmatrix} \mathcal{H}^{\text{obs}}(s_1) \\ \vdots \\ \mathcal{H}^{\text{obs}}(s_k) \end{bmatrix}.$$

Here we vectorize the error expression (4.1) as:

$$\begin{aligned} \mathbb{E}(\theta) &= \text{vec}(e(\theta, s_i) \mid s_i \in \mathcal{S}) = \text{vec}(\mathcal{H}(\theta, s_i) - \mathcal{H}^{\text{obs}}(s_i) \mid s_i \in \mathcal{S}) \\ &= \mathbb{H}(\theta) - \mathbb{H}^{\text{obs}}. \end{aligned} \quad (4.2)$$

With  $\mathcal{S}$  a fixed and finite set of sampled frequencies, we define the objective function as

$$V(\theta) = \frac{1}{2} \mathbb{E}(\theta)^* P \mathbb{E}(\theta) \quad \text{with} \quad P \succeq 0 \quad (4.3)$$

where  $P$  of dimension  $k \times k$  is a weighting matrix that represents the covariance matrix of noise sequence that is supposed to act on the measured data, or is a matrix that allows to weight different frequency points differently. If necessary, a-priori information can be incorporated in the objective function (4.3) by introducing non-uniform weighting in the frequency points. Here  $(\cdot)^*$  denotes the conjugate transpose. This formulation of objective function in (4.3) also can be incorporated for multi-input multi-output (MIMO) systems via vectorizing both frequencies and entries of the MIMO transfer functions. The Jacobian of  $V(\theta)$  with respect to the parameters  $\theta$  is

$$\begin{aligned} \nabla_{\theta} V(\theta) &= \frac{1}{2} (\nabla_{\theta} \mathbb{E}(\theta))^* P \mathbb{E}(\theta) + \frac{1}{2} \mathbb{E}(\theta)^* P (\nabla_{\theta} \mathbb{E}(\theta))^* \\ &= \begin{bmatrix} (\nabla_{\theta_1} \mathbb{E}(\theta))^* P \mathbb{E}(\theta) \\ \vdots \\ (\nabla_{\theta_j} \mathbb{E}(\theta))^* P \mathbb{E}(\theta) \\ \vdots \\ (\nabla_{\theta_{\ell}} \mathbb{E}(\theta))^* P \mathbb{E}(\theta) \end{bmatrix} = (\nabla_{\theta} \mathbb{H}(\theta))^* P \mathbb{E}(\theta) \end{aligned} \quad (4.4)$$

Here,  $\nabla_{\theta_j}$  is differentiation  $\frac{d}{d\theta_j}$  w.r.t.  $\theta_j$ , the  $j$ th coordinate of  $\theta \in \Theta$ . Notice that the  $\nabla_{\theta} \mathbb{E}(\theta)$  is a vector and the Jacobian  $\nabla_{\theta} V(\theta)$  is a scalar for the single parameter derivative  $\ell = 1$ . The Hessian of  $V(\theta)$  with respect to the parameters  $\theta$  is

$$\begin{aligned} \nabla_{\theta}^2 V(\theta) &= (\nabla_{\theta} \mathbb{E}(\theta))^* P (\nabla_{\theta} \mathbb{E}(\theta))^* \\ &= (\nabla_{\theta} \mathbb{H}(\theta))^* P (\nabla_{\theta} \mathbb{H}(\theta))^* \end{aligned} \quad (4.5)$$

Parameter calibration now consists in finding the parameter estimate as a minim-

izing argument of the objective function  $V(\theta)$ :

$$\begin{aligned} & \underset{\theta \in \Theta}{\text{minimize}} && V(\theta) \\ & \text{where} && V(\theta) := \frac{1}{2} \mathbb{E}(\theta)^* P \mathbb{E}(\theta) \\ & && \mathbb{E}(\theta) = \mathbb{H}(\theta) - \mathbb{H}^{obs}. \end{aligned} \quad (4.6)$$

The assumption of parameter calibration using frequency data is that parameters are fixed but unknown or different from the prior-knowledge of those parameters. These deviations can be caused by the production process, installation error, different operating conditions, etc. While, with the prior-knowledge and the physical nature of these parameters, we can infer some numerical constraints for calibration. In addition, the parameters for calibration in this chapter are assumed to be *structural identifiable* in the sense that the changes of parameters in the model can be observed in the model transfer function such that the parameter calibration is meaningful. For more discussion on identifiability, we refer to the paper [159]. For a high-fidelity model with the state dimension of  $n \gg 10^4$ , the optimization problem (4.6) and the numerical evaluation of  $V(\theta)$  are of high computational demand since  $\mathbb{H}(\theta)$  has to be computed for multiple frequency points  $\mathcal{S}$  and parameter points  $\Theta$ , each requiring a large matrix inverse. Thus, replacing  $\mathbb{H}(\theta)$  by a low-order parametrized model  $\tilde{\mathbb{H}}(\theta)$  appears to be an attractive numerical alternative to solve the optimization problem (4.6). Specifically, in that case, the substitute problem is

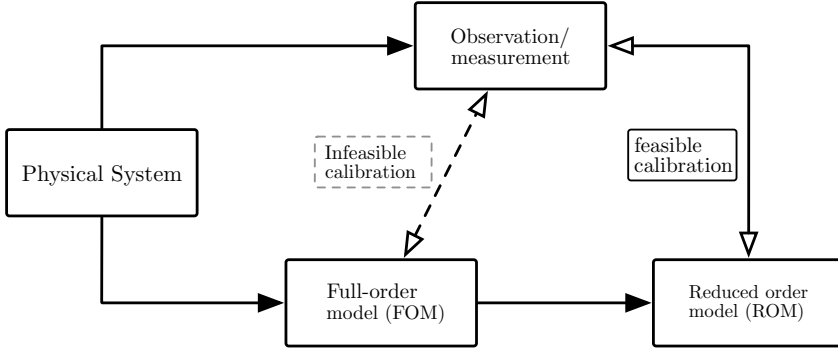
$$\begin{aligned} & \underset{\theta \in \Theta}{\text{minimize}} && \tilde{V}(\theta) \\ & \text{where} && \tilde{V}(\theta) = \frac{1}{2} \tilde{\mathbb{E}}(\theta)^* P \tilde{\mathbb{E}}(\theta) \\ & && \tilde{\mathbb{E}}(\theta) = \tilde{\mathbb{H}}(\theta) - \mathbb{H}^{obs}. \end{aligned} \quad (4.7)$$

By defining the reduced error function  $\tilde{e}(\theta, s) := \tilde{\mathcal{H}}(\theta, s) - \mathcal{H}^{obs}(s)$ , and adapting the notation above

$$\begin{aligned} \tilde{\mathbb{E}}(\theta) &= \text{vec}(\tilde{e}(\theta, s_i)) \mid s_i \in \mathcal{S} \\ &= \text{vec}(\tilde{\mathcal{H}}(\theta, s_i) - \mathcal{H}^{obs}(s_i)) \mid s_i \in \mathcal{S} \end{aligned}$$

Here  $\tilde{\mathcal{H}}(\theta, s)$  denotes the transfer function of the reduced model of the full-order  $\mathcal{H}(\theta, s)$ . The  $\tilde{\mathbb{E}}(\theta)$  is a vectorized representation of the error function of  $\tilde{e}(\theta, s)$  at the set of points  $\mathcal{S} := \{s_1, \dots, s_k\} \subset \mathbb{C}$ . There are a few methods to generate such a reduced model  $\tilde{\mathcal{H}}(\theta, s)$  that preserve the parameter-dependence of the original model [29]. The method for generating the reduced model in this work is given later. The above discussion is illustrated pictorially in Figure 4.1.

In general, the optimization problem (4.6) or its simplified version (4.7) described previously is a non-convex optimization problem, and often there is no global optimality guaranteed for the calibrated results. With a predefined parameter space  $\Theta$ , there may be one or several combinations of sampling points  $\theta \in \Theta$  such that the full-order model (FOM) matches or is close enough to the observation on



**Figure 4.1:** Parameter calibration using reduced order model for high-fidelity model. As indicated by the dashed line, the direct computation using the full-order model (FOM) is computational cumbersome and even infeasible. By taking the proper approximation of the FOM, the ROM is computational feasible for parameter calibration.

the cross product  $\Theta \times \mathcal{S}$  of these points with the frequency set  $\mathcal{S}$ . In other words, the parametric variation of the FOM is important for the parameter calibration. As indicated in Figure 4.1, the direct calibration using FOM is infeasible. The alternative is to use a ROM whose parametric variation of the FOM is preserved. However, there are two major difficulties that need to be addressed here:

- 1) The formulation of (4.7) depends on the approximation of  $\tilde{\mathcal{H}}(\theta, s)$  of  $\mathcal{H}(\theta, s)$ , which will therefore introduce the approximation errors w.r.t. the optimization (4.7). This induced error potentially could make the calibrated result distinctly different from the result obtained by performing the optimization in (4.6) on the FOM. Thus, the first question is whether we can show that the performance of the original objective function  $V(\theta)$  is close or equivalent to the one  $\tilde{V}(\theta)$  using a specific reduced model.
- 2) Among the state-of-art techniques reviewed above, the reduction procedure is set independently of the measurement data  $\mathcal{H}^{obs}$ . In other words, whether the choice of the reduced model can be refined by means of the attributes of  $\mathcal{H}^{obs}$  such that the results of  $\tilde{V}(\theta)$  is close to results of  $V(\theta)$ .

### 4.3 Error equivalent parameter calibration problem

In this section, we address the first difficulty mentioned above, and consider the equivalence of objective function (4.6) using the full-order model and the objective error function (4.7) with the reduced model is proved.

Various pMOR techniques have been proposed in the literature, among which the methods based on moment-matching are the most suitable for large-scale systems, in the sense that the implementation is computationally efficient and the method is input/output independent. In this work, we consider our early work [113] on

multi-parameter multi-frequency moment-matching method for generating the reduced model  $\tilde{\mathcal{H}}(\theta, s)$ . Before moving to the details, we briefly motivate this pMOR method. For the reasoning that the parameter  $\theta$  is separated from the variable  $\hat{s} = s \cdot \theta$  proposed in the previous work [53] and the following work. One of the benefits using this separation is that the physical interpretation is maintained in the reduced model. The parameter calibration using the derived reduced parameterized model in such a method becomes meaningful. The details are referred to in Chapter 2.

The dynamical system  $\Sigma(\theta)$  with transfer function  $\mathcal{H}(\theta, s)$  in this work is represented in state space form by

$$\Sigma(\theta) := \begin{cases} E(\theta)\dot{x}(t) = A(\theta)x(t) + B(\theta)u(t) \\ y(t) = C(\theta)x(t) \end{cases} \quad (4.8)$$

where  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$  and  $y(t) \in \mathbb{R}^q$  denote, respectively, the state vector, the input and the output. Here  $n$  is the dimension of the state variable. The state-space matrices are parameter-dependent functions  $A : \mathbb{R}^\ell \rightarrow \mathbb{R}^{n \times n}$ ,  $E : \mathbb{R}^\ell \rightarrow \mathbb{R}^{n \times n}$  etc. The  $\{E(\theta), A(\theta), B(\theta), C(\theta)\}$  are assumed to be affine

$$\begin{aligned} E(\theta) &= E_0 + E_1\theta_1 + \cdots + E_\ell\theta_\ell, \\ A(\theta) &= A_0 + A_1\theta_1 + \cdots + A_\ell\theta_\ell, \\ B(\theta) &= B_0 + B_1\theta_1 + \cdots + B_\ell\theta_\ell, \\ C(\theta) &= C_0 + C_1\theta_1 + \cdots + C_\ell\theta_\ell. \end{aligned}$$

There are several assumptions made for the given system:

#### Assumption 4.1

- (i) The matrix pair  $\{E(\theta), A(\theta)\}$  is regular for all  $\theta \in \Theta$ , i.e., the characteristic polynomial  $\mathcal{P}(\lambda) = \det(\lambda E(\theta) - A(\theta))$  is not identically zero for all  $\theta \in \Theta$ ;
- (ii) The system (4.8) considered in this chapter, we assume the parameter  $\theta \in \Theta$  with physically meaningful constraints  $\underline{\theta}_j \leq \theta_j \leq \bar{\theta}_j$  and  $-\infty < \underline{\theta}_j \leq \bar{\theta}_j < \infty$ .

Note that a class of nonlinear parameter dependent systems, i.e., parabolic systems represented by linear diffusion-convection-reaction equations, can be transformed into the linear affine form above. This enables the system representation  $\Sigma(\theta)$  to cover a rather broad scope of applications

With time-invariant parameters  $\theta$ , the transfer function of (4.8) is

$$\mathcal{H}(\theta, s) = C(\theta)[sE(\theta) - A(\theta)]^{-1}B(\theta). \quad (4.9)$$

**Multi-parameter multi-frequency moment-matching pMOR:** The work proposed by [113] provides a projection-based method for parameter-preserving interpolation parametric model order reduction. We start with a pair of expansion point

( $\theta = p_1, s = \sigma_1$ ), and the Taylor expansion of the transfer function is

$$\begin{aligned}
 \mathcal{H}(\theta, s) &= \frac{1}{0!} \mathcal{H}(p_1, \sigma_1) + \frac{1}{1!} \frac{\partial}{\partial s} \mathcal{H}(p_1, \sigma_1) (s - \sigma_1) + \frac{1}{1!} \frac{\partial}{\partial \theta} \mathcal{H}(p_1, \sigma_1) (\theta - p_1) \\
 &\quad + \frac{1}{2!} \frac{\partial^2}{\partial s^2} \mathcal{H}(p_1, \sigma_1) (s - \sigma_1)^2 + \frac{1}{2!} \frac{\partial^2}{\partial \theta^2} \mathcal{H}(p_1, \sigma_1) (\theta - p_1)^2 \\
 &\quad + \frac{1}{2!} \frac{\partial^2}{\partial s \partial \theta} \mathcal{H}(p_1, \sigma_1) (s - \sigma_1) (\theta - p_1) + \dots \\
 &= \sum_{\tau=0, \omega=0}^{\infty, \infty} \frac{1}{(\tau + \omega)!} \frac{\partial^{\tau+\omega}}{\partial \theta^\omega \partial s^\tau} \mathcal{H}(p_1, \sigma_1) (s - \sigma_1)^\tau (\theta - p_1)^\omega \\
 &= \sum_{\tau=0, \omega=0}^{\infty, \infty} m_{\{\tau, \omega\}}(p_1, \sigma_1) (s - \sigma_1)^\tau (\theta - p_1)^\omega \quad \text{with } \tau, \omega \in \mathbb{Z}_+
 \end{aligned}$$

where  $m_{\{\tau, \omega\}}(p_1, \sigma_1) = \frac{1}{(\tau + \omega)!} \frac{\partial^{\tau+\omega}}{\partial \theta^\omega \partial s^\tau} \mathcal{H}(p_1, \sigma_1)$ , the coefficient of the Taylor expansion, is defined as the moment of the system at  $(p_1, \sigma_1)$ . The 0th moment can be defined as

$$m_{\{0,0\}}(p_1, \sigma_1) := C(p_1)[\sigma_1 E(p_1) - A(p_1)]^{-1} B(p_1) = C(p_1) B_M.$$

Here we define  $B_M := [\sigma_1 E(p_1) - A(p_1)]^{-1} B(p_1)$ . The 1st moments are

$$\begin{aligned}
 m_{\{0,1\}}(p_1, \sigma_1) &:= C(p_1) \mathcal{B}_{\sigma_1} B_M, \\
 m_{\{1,0\}}(p_1, \sigma_1) &:= C(p_1) \mathcal{B}_{p_1} B_M
 \end{aligned}$$

here  $\mathcal{B}_{\sigma_1} = -[\sigma_1 E(p_1) - A(p_1)]^{-1} E(p_1)$  and  $\mathcal{B}_{p_1} = -[\sigma_1 E(p_1) - A(p_1)]^{-1} \sigma_1 E_1$ . The 2nd moments are

$$\begin{aligned}
 m_{\{0,2\}}(p_1, \sigma_1) &:= C(p_1) \mathcal{B}_{\sigma_1}^2 B_M, \\
 m_{\{2,0\}}(p_1, \sigma_1) &:= C(p_1) \mathcal{B}_{p_1}^2 B_M, \\
 m_{\{1,1\}}(p_1, \sigma_1) &:= C(p_1) \mathcal{B}_{\sigma_1} \mathcal{B}_{p_1} B_M.
 \end{aligned}$$

Defining

$$\mathcal{V} := \text{im}[B_M, \mathcal{B}_{\sigma_1} B_M, \mathcal{B}_{p_1} B_M, \mathcal{B}_{\sigma_1}^2 B_M, \mathcal{B}_{p_1}^2 B_M, \mathcal{B}_{\sigma_1} \mathcal{B}_{p_1} B_M] \quad (4.10)$$

and

$$\text{colspan}(V) = \mathcal{V},$$

we can compute the projection matrix  $V$  in such a way that columns consist of an orthonormal basis of the subspace spanned by (4.10). With the derived projection matrix  $V$ ,  $\dim(V) = n \times n_r$  with  $n_r \ll n$ , the parametric reduced order model of (4.8) is obtained via the Galerkin projection

$$\Sigma_r(\theta) := \begin{cases} E_r(\theta) \dot{x}_r(t) = A_r(\theta) x_r(t) + B_r(\theta) u(t) \\ y_r(t) = C_r(\theta) x_r(t) \end{cases} \quad (4.11)$$

where the reduced state  $\dim(x_r) = n_r$ . The reduced system matrices remain their affine form and can be computed by the following

$$\begin{aligned} E_r(\theta) &= W^T E_0 V + W^T E_1 V \theta_1 + \cdots + W^T E_\ell V \theta_\ell, \\ A_r(\theta) &= W^T A_0 V + W^T A_1 V \theta_1 + \cdots + W^T A_\ell V \theta_\ell, \\ B_r(\theta) &= W^T B_0 + W^T B_1 \theta_1 + \cdots + W^T B_\ell \theta_\ell, \\ C_r(\theta) &= C_0 V + C_1 V \theta_1 + \cdots + C_\ell V \theta_\ell, \end{aligned}$$

where, here, we set  $W = V$ . The transfer function of the reduced model (4.11) is

$$\tilde{\mathcal{H}}(\theta, s) := C_r(\theta)[sE_r(\theta) - A_r(\theta)]^{-1}B_r(\theta).$$

In the case of Petro-Galerkin projection  $W \neq V$ , the subspace of  $\mathcal{W}$  can be constructed in a similar way as  $\mathcal{V}$  described above. Defining

$$\mathcal{W} := \text{im}[C_M, \mathcal{C}_{\sigma_1} C_M, \mathcal{C}_{p_1} C_M, \mathcal{C}_{\sigma_1}^2 C_M, \mathcal{C}_{p_1}^2 C_M, \mathcal{C}_{\sigma_1} \mathcal{C}_{p_1} C_M] \quad (4.12)$$

where  $C_M = [\sigma_1 E(p_1) - A(p_1)]^{-T} C(p_1)^T$ ,  $\mathcal{C}_{\sigma_1} = -[\sigma_1 E(p_1) - A(p_1)]^{-T} E(p_1)^T$  and  $\mathcal{C}_{p_1} = -[\sigma_1 E(p_1) - A(p_1)]^{-T} \sigma_1 E_1$ . Then, the subspace is

$$\text{colspan}(W) = \mathcal{W} \quad \text{with} \quad \dim(W) = n \times n_r.$$

where  $W$  is a full rank matrix whose columns are mutually orthonormal. To obtain square reduced system matrices  $\{E_r, A_r\}$  for practical considerations, the projection matrices  $W, V$  should have the same dimension for a well-defined Petro-Galerkin projection. Often, this problem arises for MIMO systems where the number of inputs is different from the number of outputs, or rank deficiency in generating the projection matrices. As Theorem 2.4 states, the moment-matching method for the system  $\Sigma(\theta)$  determines a reduced model whose transfer function coincides in as many coefficients of its Taylor expansions about  $(s = \sigma, \theta = p)$  for a given order of the expanded moments. This yields that the transfer function of the full-order model and the transfer function of the reduced model satisfies the Hermite interpolation conditions.

**Lemma 4.1** *Given a transfer function  $\mathcal{H}(\theta, s)$  of a full-order parameter dependent system as in (4.8) and the reduced model transfer function  $\tilde{\mathcal{H}}(\theta, s)$ . Assume that  $\mathcal{H}(\theta, s)$  and  $\tilde{\mathcal{H}}(\theta, s)$  satisfy the Theorem 2.4 for the expansion point  $(\sigma, p)$  with up to the second order moment, then the following Hermite interpolation conditions satisfy*

$$\nabla_s^i \mathcal{H}(\theta, s) = \nabla_s^i \tilde{\mathcal{H}}(\theta, s), \quad (4.13)$$

$$\nabla_\theta^j \mathcal{H}(\theta, s) = \nabla_\theta^j \tilde{\mathcal{H}}(\theta, s), \quad (4.14)$$

up to an order  $i = 0, 1, 2$  and  $j = 0, 1, 2$  for the frequency expansion point  $s = \sigma \in \mathbb{C}$  not being a pole of  $\mathcal{H}$  and the parameter expansion point  $\theta = p \in \Theta$ , respectively.

There are several remarks here:

- In practice, the explicit computation of subspace  $\{\mathcal{V}, \mathcal{W}\}$  may lead to numerical instability where the matrices become linearly dependent after the first few moments. To overcome this problem, we deploy the implicit moment-matching method for constructing  $\{\mathcal{V}, \mathcal{W}\}$ .
- In the above, we only consider a single pair of expansion points for deriving the reduced model above. For multiple expansion  $\ell > 1$  points moment-matching, we refer to our previous work in Chapter 2. When more than one parameters are desired for calibration, or higher accuracy of the reduced model is required, multiple expansion points should be taken in the reduction. However, the number of columns in  $V$  (equivalent to the order of the reduced model) increases exponentially with the number of expansion points, and the computation of the matrix becomes much more expensive. With our experience, only a few number of parameter expansion points and moments suffice to achieve similar accuracy in comparison to the large number of frequency moments. As for the frequency expansion points, we have observed in many different systems, the reduced model with multiple expansion points  $\{\sigma_1, \sigma_2, \dots\}$  often outperforms the one with single expansion point  $\sigma_1$  for the same reduced order and comparable accuracy. The drawback is that it requires to solve the inverse problems  $[\sigma E(p) - A(p)]^{-1}$  for every new expansion point. This is one of the most computationally demanding steps in the implementation.
- The choice of the number and location of the expansion points  $\{p, \sigma\}$  plays an essential role in the efficiency and the accuracy of the multi-parameter multi-frequency moment-matching method. In particular, for the application of parameter calibration using the parameterized reduced model, accurate and robust calibrated results only can be achieved when the expansion points are appropriately chosen.

**Equivalent conditions between two objective functions:** In this part, we present the equivalent conditions of two objective functions with respect to parameter optimization.

A solution of the full-order objective function (4.6) can be found using the Lagrangian function

$$\begin{aligned} \mathcal{L}(\theta) &:= V(\theta) + \iota(\theta) \\ &= \frac{1}{2} \left( \mathbb{H}(\theta) - \mathbb{H}^{\text{obs}} \right)^* P \left( \mathbb{H}(\theta) - \mathbb{H}^{\text{obs}} \right) + \iota(\theta) \end{aligned} \quad (4.15)$$

where  $\iota(\theta)$  denotes the barrier function and it is assumed to be a smooth function of  $\theta$ . With the assumption that  $\mathcal{L}(\theta)$  is continuous and differentiable at  $\theta$ , then the first order derivative of the Lagrangian  $\mathcal{L}(\theta)$  with respect to the parameter  $\theta$  is

$$\begin{aligned} \nabla_{\theta} \mathcal{L}(\theta) &= \nabla_{\theta} V(\theta) + \nabla_{\theta} \iota(\theta) \\ &= \left( \nabla_{\theta} \mathbb{H}(\theta) \right)^* P \left( \mathbb{H}(\theta) - \mathbb{H}^{\text{obs}} \right) + \nabla_{\theta} \iota(\theta). \end{aligned} \quad (4.16)$$



In the same manner, we can derive a solution for the objective function  $\tilde{V}(\theta)$  in (4.7) using Lagrangian

$$\begin{aligned}\tilde{\mathcal{L}}(\theta) &:= \tilde{V}(\theta) + \iota(\theta) \\ &= \frac{1}{2} \left( \tilde{\mathbb{H}}(\theta) - \mathbb{H}^{\text{obs}} \right)^* P \left( \tilde{\mathbb{H}}(\theta) - \mathbb{H}^{\text{obs}} \right) + \iota(\theta).\end{aligned}\quad (4.17)$$

Notice that the indicator  $\iota(\theta)$  remains the same for both objective functions since both the parameter constraints  $\theta_i \in [\underline{\theta}_i, \bar{\theta}_i], i = 1, \dots, \ell$  are defined as the same one. The first order derivative about parameter  $\theta$  is

$$\begin{aligned}\nabla_{\theta} \tilde{\mathcal{L}}(\theta) &= \nabla_{\theta} \tilde{V}(\theta) + \nabla_{\theta} \iota(\theta) \\ &= \left( \nabla_{\theta} \tilde{\mathbb{H}}(\theta) \right)^* P \left( \tilde{\mathbb{H}}(\theta) - \mathbb{H}^{\text{obs}} \right) + \nabla_{\theta} \iota(\theta).\end{aligned}\quad (4.18)$$

**Theorem 4.2** Suppose the parameter expansion point  $\theta = p \in \Theta$  is such that the transfer function  $\mathcal{H}(p)$  of the system (4.8) and the transfer function  $\tilde{\mathcal{H}}(p)$  of the reduced model (4.11) are regular. If the reduced model transfer function  $\tilde{\mathcal{H}}(\theta, s)$  and the full-order transfer function  $\mathcal{H}(\theta, s)$  satisfy the Hermite interpolation conditions  $\nabla_{\theta}^j \mathcal{H}(\theta, s) = \nabla_{\theta}^j \tilde{\mathcal{H}}(\theta, s)$  for  $j = 0, 1$  at  $\theta = p$  for all  $s \in \mathcal{S}$ . Then

- a) the Lagrangian  $\mathcal{L}(\theta)$  of the full-order objective function (4.15) and the Lagrangian  $\tilde{\mathcal{L}}(\theta)$  of the reduced-order objective function (4.17) satisfy

$$\mathcal{L}(p) = \tilde{\mathcal{L}}(p),$$

- b) furthermore the first-order derivative with respect to  $\theta$  of the above two Lagrangians satisfies

$$\nabla_{\theta} \mathcal{L}(p) = \nabla_{\theta} \tilde{\mathcal{L}}(p)$$

where  $\theta$  is evaluated at  $p$ .

**Proof:** Define the approximation error

$$\delta(\theta, s) := \mathcal{H}(\theta, s) - \tilde{\mathcal{H}}(\theta, s),$$

with the Hermite interpolation conditions, we have

$$\nabla_{\theta}^j \delta(\theta, s) = 0 \mid_{(\theta=p, s \in \mathcal{S})}, \quad j = 0, 1.$$

The error function in (4.1) can be reformulated

$$\begin{aligned}e(\theta, s) &= \mathcal{H}(\theta, s) - \mathcal{H}^{\text{obs}}(s) \\ &= \mathcal{H}(\theta, s) - \tilde{\mathcal{H}}(\theta, s) + \tilde{\mathcal{H}}(\theta, s) - \mathcal{H}^{\text{obs}}(s) \\ &= \delta(\theta, s) + \tilde{e}(\theta, s)\end{aligned}$$

and, (4.3) leads to the following

$$\begin{aligned}
\nabla_{\theta}^j e(\theta, s) &= \nabla_{\theta}^j (\mathcal{H}(\theta, s) - \tilde{\mathcal{H}}(\theta, s) + \tilde{\mathcal{H}}(\theta, s) - \mathcal{H}^{\text{obs}}(s)) \\
&= \nabla_{\theta}^j \delta(\theta, s) + \nabla_{\theta}^j \tilde{e}(\theta, s) \\
&= 0 + \nabla_{\theta}^j \tilde{e}(\theta, s) \quad \text{at } \theta = p \quad \text{for } j = 0, 1
\end{aligned} \tag{4.19}$$

Followed by the above reformulation, the objective function with a set of frequency points  $\mathcal{S} := \{s_1, \dots, s_k\} \subset \mathbb{C}$ ,  $s_i \in \mathcal{S}$  in (4.6) reads

$$\begin{aligned}
V(\theta) &= \frac{1}{2} \mathbb{E}(\theta)^* P \mathbb{E}(\theta) \\
&= \frac{1}{2} (\mathbb{H}(\theta) - \tilde{\mathbb{H}}(\theta) + \tilde{\mathbb{H}}(\theta) - \mathbb{H}^{\text{obs}})^* P (\mathbb{H}(\theta) - \tilde{\mathbb{H}}(\theta) + \tilde{\mathbb{H}}(\theta) - \mathbb{H}^{\text{obs}}) \\
&= \frac{1}{2} \text{vec}(\delta(\theta, s_i) + \tilde{e}(\theta, s_i))^* P \text{vec}(\delta(\theta, s_i) + \tilde{e}(\theta, s_i)) \\
&= \frac{1}{2} \text{vec}(\delta(\theta, s_i))^* P \text{vec}(\delta(\theta, s_i)) + \frac{1}{2} \text{vec}(\delta(\theta, s_i))^* P \tilde{\mathbb{E}}(\theta) \\
&\quad + \frac{1}{2} \tilde{\mathbb{E}}(\theta)^* P \text{vec}(\delta(\theta, s_i)) + \underbrace{\frac{1}{2} \tilde{\mathbb{E}}(\theta)^* P \tilde{\mathbb{E}}(\theta)}_{\tilde{V}(\theta)}
\end{aligned} \tag{4.20}$$

Indeed, the objective function  $V(\theta)$  using the full-order model can be rewritten as a combination of the reduced model objective function  $\tilde{V}(\theta)$  and other terms. Using (4.19), term consists of  $\delta(\theta, s_i)$  or  $\nabla_{\theta} \delta(\theta, s_i)$  vanishes at  $\theta = p$ , and we have the following equivalence

$$\begin{aligned}
V(\theta) &= 0 + 0 + 0 + \tilde{V}(\theta) \\
\iff \mathcal{L}(\theta) &= \tilde{\mathcal{L}}(\theta) \quad \text{at } \theta = p
\end{aligned}$$

Thus, the Lagrangian equivalence is proved. In a similar way, the Jacobian of the objective function about  $\theta$  is

$$\begin{aligned}
\nabla_{\theta} V(\theta) &= \text{vec}(\nabla_{\theta} \delta(\theta, s_i))^* P \text{vec}(\delta(\theta, s_i)) \\
&\quad + \frac{1}{2} [\text{vec}(\nabla_{\theta} \delta(\theta, s_i))^* P \tilde{\mathbb{E}}(\theta) + \text{vec}(\delta(\theta, s_i))^* P \nabla_{\theta} \tilde{\mathbb{E}}(\theta)] \\
&\quad + \frac{1}{2} [\nabla_{\theta} \tilde{\mathbb{E}}(\theta)^* P \text{vec}(\delta(\theta, s_i)) + \tilde{\mathbb{E}}(\theta)^* P \text{vec}(\nabla_{\theta} \delta(\theta, s_i))] + \nabla_{\theta} \tilde{V}(\theta).
\end{aligned}$$

With (4.19), the term consists of  $\delta(\theta, s_i)$  or  $\nabla_{\theta} \delta(\theta, s_i)$  vanishes at  $\theta = p$ , and we have the equivalence

$$\begin{aligned}
\nabla_{\theta} V(\theta) &= \nabla_{\theta} \tilde{V}(\theta) \\
\iff \nabla_{\theta} \mathcal{L}(\theta) &= \nabla_{\theta} \tilde{\mathcal{L}}(\theta)
\end{aligned}$$

where  $\theta$  evaluated at  $p$ . Thus, it is proved.  $\square$

Theorem 4.2 gives an elementary condition for finding a solution of the optimization problems. Assume that the  $\theta^{\text{opt}} \in \Theta$  is a local minimizer of optimization problem in (4.6), and  $\mathcal{L}(\theta)$  is continuously differentiable in an open neighbourhood of  $\theta^{\text{opt}}$ , then the first order necessary condition of the full-order objective function (4.6) reads

$$\begin{aligned} \nabla_{\theta} \mathcal{L}(\theta^{\text{opt}}) \ni 0 \iff & \left[ \text{vec}(\nabla_{\theta} \delta(\theta^{\text{opt}}, s_i))^* P \text{vec}(\delta(\theta^{\text{opt}}, s_i)) \right. \\ & + \frac{1}{2} [\text{vec}(\nabla_{\theta} \delta(\theta^{\text{opt}}, s_i))^* P \tilde{\mathbb{E}}(\theta^{\text{opt}}) + \text{vec}(\delta(\theta^{\text{opt}}, s_i))^* P \nabla_{\theta} \tilde{\mathbb{E}}(\theta^{\text{opt}})] \\ & + \frac{1}{2} [\nabla_{\theta} \tilde{\mathbb{E}}(\theta^{\text{opt}})^* P \text{vec}(\delta(\theta^{\text{opt}}, s_i)) + \tilde{\mathbb{E}}(\theta^{\text{opt}})^* P \text{vec}(\nabla_{\theta} \delta(\theta^{\text{opt}}, s_i))] \\ & \left. + \nabla_{\theta} \tilde{V}(\theta^{\text{opt}}) + \underbrace{\sum_{i=1}^{\ell} \nabla_{\theta_i} \iota(\theta_i^{\text{opt}})}_{\nabla_{\theta} \tilde{\mathcal{L}}(\theta^{\text{opt}})} \right] \ni 0. \end{aligned} \quad (4.21)$$

If the expansion point is chosen as the local minimizer  $p = \theta^{\text{opt}}$ , we obtain the equivalent first-order derivative condition if and only if

$$\nabla_{\theta} \mathcal{L}(\theta^{\text{opt}}) \ni 0 \iff \nabla_{\theta} \tilde{\mathcal{L}}(\theta^{\text{opt}}) \ni 0.$$

Conclude that the necessary conditions for optimal solutions of the two problems (full order and reduced order) therefore match.

**Theorem 4.3** Suppose two objective functions  $V(\theta)$  in (4.15) and  $\tilde{V}(\theta)$  in (4.17) that satisfy Theorem 4.2:  $\nabla_{\theta}^i \mathcal{L}(p) = \nabla_{\theta}^i \tilde{\mathcal{L}}(p)$  at  $\theta = p$  for  $i = 0, 1$  and  $s \in \mathcal{S}$ . If a local minimizer  $\theta = \theta^{\text{opt}}$  for  $V(\theta)$  is chosen as the parameter expansion point for  $\tilde{\mathcal{H}}(\theta, s)$

$$p = \theta^{\text{opt}},$$

then we obtain the equivalent first-order condition for  $\nabla_{\theta} \mathcal{L}(\theta^{\text{opt}})$  and  $\nabla_{\theta} \tilde{\mathcal{L}}(\theta^{\text{opt}})$  if and only if both satisfy

$$\nabla_{\theta} \mathcal{L}(\theta^{\text{opt}}) \ni 0 \quad \text{and} \quad \nabla_{\theta} \tilde{\mathcal{L}}(\theta^{\text{opt}}) \ni 0. \quad (4.22)$$

In particular, if  $\theta^{\text{opt}}$  is chosen as a vector satisfies the first order derivative conditions

$$\nabla \tilde{\mathcal{L}}(\theta^{\text{opt}}) = 0$$

then it also satisfies the necessary optimality condition for the original problem  $\mathcal{L}(\theta)$ .

**Proof:** By considering the Hermite condition for the reduced model and the necessary first-order derivative condition stated in Theorem 4.2, the rest of the proof is straightforward.  $\square$

## 4.4 Moment-matching for parameter calibration

Once the observation data is obtained associated with the given sampling frequency point  $s_i$  and parameter  $\theta$ , the observed model  $\mathcal{H}^{\text{obs}}(\theta, s)$  can be defined. To solve the optimization problem using the reduced order model, we also need to choose the prior pair of expansion points  $(p, \sigma)$  for the reduced-order model  $\tilde{\mathcal{H}}(\theta, s)$ . The conventional method for parametric fitting of large-scale models is summarized in Algorithm 2.

---

**Algorithm 2** : Moment matching for parametric fitting
 

---

- 1: **Given**: the observed model  $\mathcal{H}^{\text{obs}}(s)$  with a set of frequency points  $s \in \mathcal{S} := \{s_1, \dots, s_k\}$ ,
  - 2: **Input**: initial points  $\{p_j^0, \sigma_i^0\}_{j=i=1}^{n_\ell, n_k}$  and the corresponding moments  $m_{(j,i)}$ ,
  - 3: Construct the subspace  $\mathcal{V}$  and  $\mathcal{W}$  such that they satisfy the (4.10) and (4.12),
  - 4: Build the reduced system matrices  $\{E_r(\theta), A_r(\theta), B_r(\theta), C_r(\theta)\}$  as in (4.11) using  $V$  and  $W$  and compute the frequency response of the reduced system  $\tilde{\mathcal{H}}(\theta, s)$ ,
  - 5: Solve the optimization problem (4.7) with a proper weighting matrix  $P$ ,
  - 6: **return**  $\theta \leftarrow \theta^{\text{opt}}$ .
- 

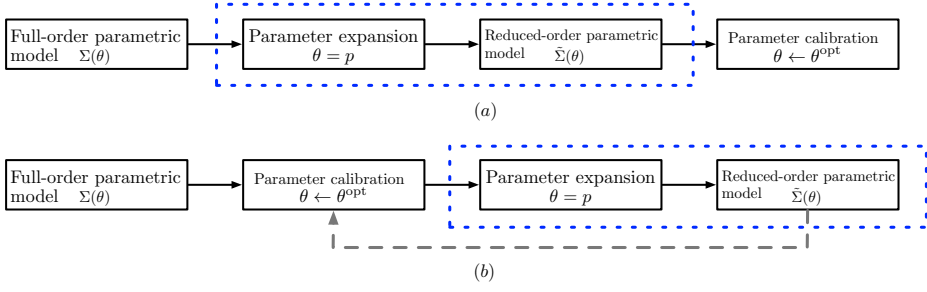
For the case of MIMO systems, we employ the tangential direction method [21], where the input tangential direction  $\mathbf{r} \in \mathbb{R}^n$  and output tangential direction  $\mathbf{l} \in \mathbb{R}^q$  can be pre-defined at *step 2*. The Algorithm 2 can be roughly illustrated as (a), depicted in Figure 4.2. The blue-dashed block denotes the procedure of parameterized model order reduction, which generates the  $\tilde{\Sigma}(\theta)$  for parameter calibration. Indeed, as indicated in Lemma 4.3, the equivalent parameter calibration result can be achieved if the parameter expansion point is chosen as the optimal  $\theta = p = \theta^{\text{opt}}$ . In other words, regardless of the computational difficulty, the ideal procedure to have equivalent calibration results between the FOM  $\Sigma(\theta)$  and the ROM  $\tilde{\Sigma}(\theta)$  is to obtain the  $\theta \leftarrow \theta^{\text{opt}}$  and choose the optimal  $\theta^{\text{opt}} = p$  as the expansion point. In this way, calibration using the reduced model  $\tilde{\Sigma}(\theta)$  leads to the same optimal results, depicted in (b) Figure 4.2.

Based on the discussion above, we take a closer look at (4.20)

$$\begin{aligned}
 V(\theta) - \tilde{V}(\theta) &= \underbrace{\frac{1}{2} \text{vec}(\delta(\theta, s_i))^* P \text{vec}(\delta(\theta, s_i))}_{\text{approximation error}} \\
 &\quad + \underbrace{\frac{1}{2} \text{vec}(\delta(\theta, s_i))^* P \text{vec}(\tilde{e}(\theta, s_i)) + \frac{1}{2} \text{vec}(\tilde{e}(\theta, s_i))^* P \text{vec}(\delta(\theta, s_i))}_{\text{mixed error}}.
 \end{aligned} \tag{4.23}$$

Here the approximated error is a quadratic function of the approximation error  $\delta(\theta, s_i)$ , and the mixed error includes the model mismatch  $\delta(\theta, s_i)$  and the reduced error  $\tilde{e}(\theta, s_i)$  between the reduced model and the measurement data.

**Remark 4.1** *The equation (4.23) has an interpretation that the error between these two objective functions  $V(\theta) - \tilde{V}(\theta)$  is due to the summation of the model mismatch  $\delta(\theta, s_i)$*



**Figure 4.2:** (a) The procedure of the conventional approach aiming at parameter calibrations using reduced-order parametric model  $\tilde{\Sigma}(\theta)$ . The blue-dashed block denotes the parameterized model reduction, including the selection of parameter expansion points; (b) Ideal procedure to achieve the equivalent performance of parameter calibration between  $\Sigma(\theta)$  and  $\tilde{\Sigma}(\theta)$ , where the grey-dashed line indicates that parameter calibration using  $\tilde{\Sigma}(\theta)$  results in the same optimization result  $\theta \leftarrow \theta^{opt}$  as the parameter calibration using  $\Sigma(\theta)$ .

induced by the model reduction procedure and the reduced error  $\tilde{e}(\theta)$  between the reduced order model and the measurement data. Although the optimization methods/techniques for obtaining the calibration results are not discussed in this work, we note that the error  $\tilde{e}(\theta)$  with one optimal  $\theta \leftarrow \theta^{opt}$  may not be equal to zero for all frequency points  $s \in \mathcal{S}$ . Under the assumption that the full-order parameterized model with the parameter space  $\theta \in \Theta$  matches or is close enough to the observation  $e(\theta, s_i) := \mathcal{H}(\theta, s_i) - \mathcal{H}^{obs}(s_i) \rightarrow 0$  at  $\theta \leftarrow \theta^{opt}$  for all  $s_i \in \mathcal{S}$ . The reduced error at the optimal parameter value can be expressed as

$$\begin{aligned} \tilde{e}(\theta^{opt}, s_i) &= \tilde{\mathcal{H}}(\theta^{opt}, s_i) - \mathcal{H}^{obs}(s_i) \\ &= \tilde{\mathcal{H}}(\theta^{opt}, s_i) - \mathcal{H}(\theta^{opt}, s_i) + \mathcal{H}(\theta^{opt}, s_i) - \mathcal{H}^{obs}(s_i) \\ &= -\delta(\theta^{opt}, s_i) + e(\theta^{opt}, s_i). \end{aligned}$$

Thus, the error between  $V(\theta)$  and  $\tilde{V}(\theta)$  due to the approximation error  $\delta(\theta, s_i)$  at those frequency points where the parameters at  $\theta^{opt}$  are most sensitive with respect to the input-output relation, is expected to dominate. Therefore, to minimize this error in (4.23), one can try to eliminate the approximation error  $\delta(\theta, s_i)$  induced by the model reduction.

Thus, based on the above considerations and Theorem 4.3, this suggests a set of novel frequency interpolation points as follows. Define, for any  $s \in \mathcal{S} := \{s_1, \dots, s_k\}$  the sensitivity of the  $j$ -th parameter in  $\tilde{\mathcal{H}}(\theta, s)$  by setting:

$$\mathcal{J}_j(\theta, s) := \nabla_{\theta_j} \tilde{\mathcal{H}}(\theta, s), \text{ for } j = 1, \dots, \ell \quad (4.24)$$

where  $\nabla_{\theta_j}$  denotes the derivative w.r.t to the  $j$ -th entry in  $\theta$ . For all points  $s \in \mathcal{S}$

we stack these sensitivities into a vector

$$\mathcal{J}_j(\theta, \mathcal{S}) = \begin{bmatrix} \nabla_{\theta_j} \tilde{\mathcal{H}}(\theta, s_1) \\ \nabla_{\theta_j} \tilde{\mathcal{H}}(\theta, s_2) \\ \vdots \\ \nabla_{\theta_j} \tilde{\mathcal{H}}(\theta, s_k) \end{bmatrix} \quad \text{at and } j = 1, \dots, \ell \quad (4.25)$$

Now define

$$\tilde{\sigma}_j = \arg \max_{s \in \mathcal{S}} |\mathcal{J}_j(\theta, s)| \quad (4.26)$$

as the adapted frequency interpolation point. We determine these frequency points for  $j = 1, \dots, \ell$  and with  $\theta \in \Theta$  in (4.26) chosen as the computed optimal  $\theta = \theta^{\text{opt}}, \theta \in \Theta$  from the optimization (4.7)

With the proposed frequency point candidate at hand, we can discuss the corresponding order/moment. Since the objective of the reduced model is for parameter calibration, in which often gradient-type optimization methods are considered, 2 moments for the refined points are sufficient. Here is a sketch of the proposed algorithm.

---

**Algorithm 3** : Adaptive moment matching for parametric fitting

---

- 1: **Given**: the observed model  $\mathcal{H}^{\text{obs}}(s)$  with a set of frequency points  $\mathcal{S} := \{s_1, \dots, s_k\}$ ,
  - 2: **Input**: initial expansion points  $\{p_j^0, \sigma_i^0\}_{j=1}^{n_\ell, n_k}$ , the corresponding moments  $m_{(j,i)}$ , the tolerance  $tol > 0$  and the weight matrix  $P \succ 0$ ,
  - 3: Construct the subspace  $\mathcal{V} = \text{colspan}(V)$  and  $\mathcal{W} = \text{colspan}(W)$  such that they satisfy the equation (4.10) and the equation (4.12),
  - 4: Build  $\{E_r(\theta), A_r(\theta), B_r(\theta), C_r(\theta)\}$  as in (4.11) using  $V$  and  $W$  and compute the frequency response of the reduced system  $\tilde{\mathcal{H}}(\theta, s)$ ,
  - 5: Solve the optimization problem (4.7)  $\theta \leftarrow \theta^{\text{opt}}$  and obtain the Jacobian matrix  $\nabla_{\theta} \tilde{\mathcal{H}}(\theta^{\text{opt}}, s)$ ,
  - 6: **While** (relative change in  $\{\tilde{\sigma}_j\} > tol$ )  
 do for  $j = 1, \dots, \ell$   
 i)  $\tilde{\sigma}_j = \arg \max_{s \in \mathcal{S}} |\mathcal{J}_j(\theta^{\text{opt}}, s)|$ ,  
 ii) Assign  $\{\tilde{p}_j, \tilde{\sigma}_j\}, \tilde{p}_j = \theta_j^{\text{opt}}$  with suitable moments  $\tilde{m}_j$ ,  
 iii) Combine  $\{\tilde{p}_j, \tilde{\sigma}_j\} \cup \{p_j^0, \sigma_i^0\}_{j=1}^{n_\ell, n_k}$ , and refine the  $\mathcal{S} := \{\mathcal{S}, \tilde{\sigma}_j\}$   
 enddo  
 iv) Update  $\mathcal{V} = \text{colspan}(V)$ ,  $\mathcal{W} = \text{colspan}(W)$  with the expansion points in (iii), and construct  $\{E_r(\theta), A_r(\theta), B_r(\theta), C_r(\theta)\}$ ,  
 v) Solve the optimization problem (4.7) and obtain  $\nabla_{\theta} \tilde{\mathcal{H}}(\theta^{\text{opt}}, s)$ ,
  - 7: **return**  $\theta \leftarrow \theta^{\text{opt}}$ .
- 

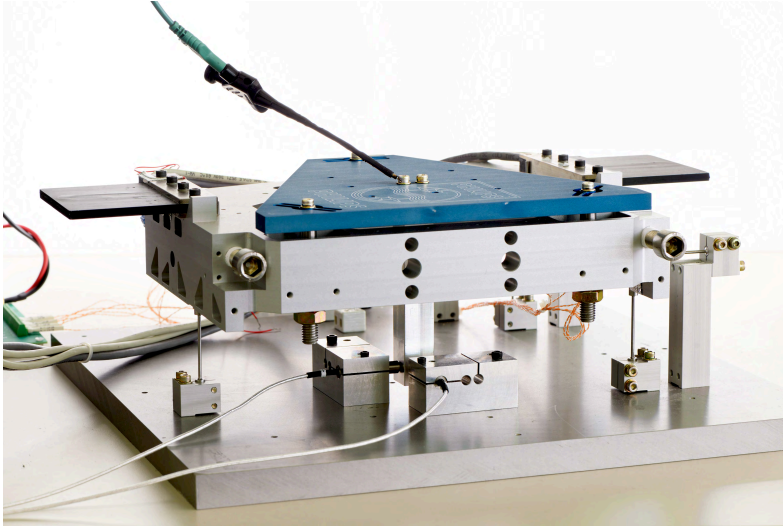
Some remarks on Algorithm 3 are in order.

**Remark 4.2**    *a) The choice of the tolerance  $tol$  in Step 2 is dependent on the accuracy requirement of the reduced model and the requirement for the parameter calibration.*

- b) The construction of the subspace  $\mathcal{V}, \mathcal{W}$  in Step 3 and Step 6, item iv) usually involves computing a (sparse) matrix inverse that is often ill-conditioned:  $(\sigma_j^0 E(p_j^0) - A(p_j^0))^{-1}$ . For every new pair of expansion points  $\{p_j^0, \sigma_j^0\}$ , the algorithm requires only one matrix inverse. The columns/subspace computed in Step 3) can be reused in Step 6, item iv) if the number of the expansion points remains the same.
- c) In Step 6, item iii), we add the updated expansion points  $\{\tilde{p}_j, \tilde{\sigma}_j\}$  to the initial expansion points  $\{p_j^0, \sigma_i^0\}$  due to which the reduced-order model dimension grows. The alternative is to replace the initial expansion points with the updated ones, and in such a way, the reduced-order model remains invariant. Both methods are feasible, and we consider the first one in this chapter.
- d) In Step 6, item iii), we add the  $\theta^{opt}$  to the parameter expansion points, which leads to an increase of the reduced-order model. One alternative is to deploy the Algorithm 4 we proposed in Chapter 5 where the parameter-dependent system transforms into a linear fractional representation, in such a way that the number of the reduced model is less in comparison to other pMOR methods.

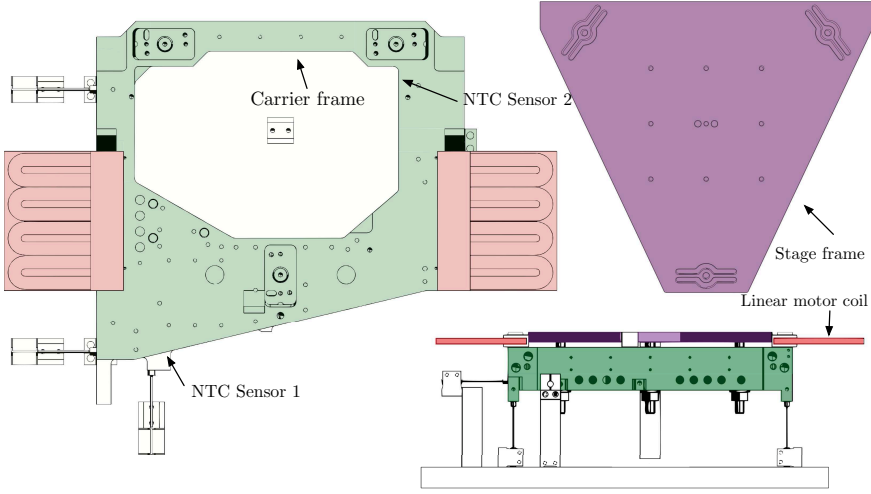
## 4.5 Numerical example: the PSA setup

We demonstrate our algorithm for a high-accuracy thermal setup: the Precision Stage Application (PSA), depicted in Figure 4.3. This PSA setup consists of three



**Figure 4.3:** Overview of the PSA setup

main parts: a carrier frame (in green), a stage frame (in purple), and two linear motors (in red). The schematic plots of these three parts are given in Figure 4.4. During the operation, the linear motors are identified as the main heat source: the coils generate heat flux which diffuses to the carrier frame and the stage frame.



**Figure 4.4:** Schematic plot of the PSA setup which includes three main parts: the carrier frame (left top), the stage frame (right top) and two linear motors (the components in pink). The right bottom denotes the side view of the PSA.

The governing equation for describing the temporal-spatial thermal behaviour is given by

$$\rho c_p \frac{\partial T}{\partial t} + \nabla \cdot J = u \quad \text{in } \mathcal{G}$$

$$\text{Boundary condition: } \nabla T(r, t) \cdot \bar{n} = \gamma \quad \text{on } \partial \mathcal{G}$$

$$\text{Initial condition: } T(r, 0) = T_0(r)$$

here  $T$  denotes the temperature of the setup,  $J = -\kappa \nabla T(r, t)$ ,  $r \in \mathcal{G}$  is the heat flux.  $\{\rho, c_p, \kappa\} \in \mathbb{R}_{>0}$  represents the density, heat capacity and thermal conductivity of materials in the PSA setup, respectively. It is worth mentioning that the green part of the carrier frame is built in aluminium, and the linear motor is covered with steel. Due to the design requirement, there are 2 points of interest (POI) that are equipped with negative temperature coefficient (NTC) thermistors. In this chapter, the temperatures of these POI are considered as the outputs, and the power to the linear motor (we assume the lossless conversion of electricity to thermal energy) which generates the heat is chosen as the input of the thermal model, respectively. With a proper spatial discretization method, e.g., finite element method (FEM), we obtain the approximated ordinary differential equation, given in the state space form:

$$\Sigma(\theta) := \begin{cases} E(\theta) \dot{x}(t) = A(\theta)x(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \quad (4.27)$$

where the state  $x \in \mathbb{R}^{38476}$  represents the temperature of different locations of the setup,  $u \in \mathbb{R}^2$  and  $y \in \mathbb{R}^2$  denote, respectively, the input and the output. The system matrices are functions  $E: \mathbb{R}^2 \rightarrow \mathbb{R}^{38476 \times 38476}$ ,  $A: \mathbb{R}^2 \rightarrow \mathbb{R}^{38476 \times 38476}$ ,  $B \in$



$\mathbb{R}^{38476 \times 2}$  and  $C : \mathbb{R}^{2 \times 38476}$ . The parameter dependent matrices  $E(\theta), A(\theta)$  are assumed to be affine and of the form:

$$E(\theta) = E_0 + \theta_1 E_1 + \theta_2 E_2$$

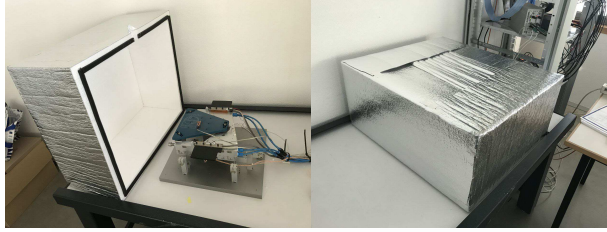
$$A(\theta) = A_0 + A_1 \theta_3 + A_4 \theta_4$$

where  $\theta_1, \theta_2$  denote the heat capacity for three materials: steel and aluminum in the PSA:  $C_{p_{ste}}, C_{p_{alu}}$ , and the  $\kappa_{ste}, \kappa_{alu}$  represent the thermal conductivity for the steel and the aluminum, the variation range of each  $\theta_i, i = 1, \dots, 4$  is summarized in the following table

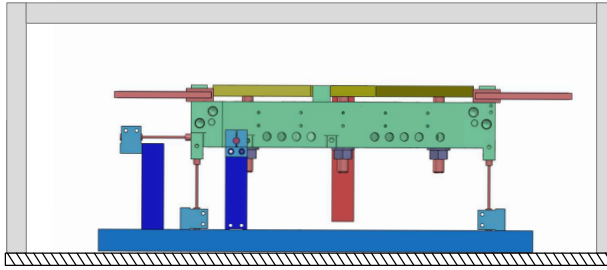
Name	Nominal Value (unit)	Lower bound	Upper bound
$\theta_1: C_{p_{ste}}$	475 (J/kg · K)	450	500
$\theta_2: C_{p_{alu}}$	900 (J/kg · K)	825	975
$\theta_3: \kappa_{ste}$	44.5 (W/m · K)	39.5	49.5
$\theta_4: \kappa_{alu}$	27 (W/m · K)	23	31

**Table 4.1:** 4 parameters  $\theta$ : nominal value, upper bound and lower bound.

To facilitate this demonstration, we perform a temperature measurement of the PSA in Figure 4.3 and construct a MIMO system. The system is excited via a random-phase multisine limited to 0.1 [Hz] with a peak of 4 [W], which is centred at offset 4 [W]. The data acquisition runs at a sampling frequency of  $F_s = 5$  [Hz]. We excite the system with a period length of 1 [hour] and repeated with 36 times.



(a) Photograph of the PSA setup with isolation shield

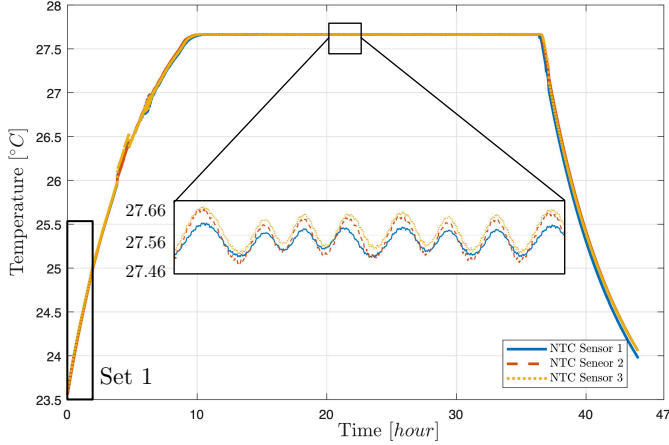


(b) Schematic plot of the PSA setup with isolation shield

**Figure 4.5:** Photograph of the PSA setup with the isolation shield (a) and the schematic plot of the PSA setup including the isolation shield (b).

As pointed in [66], the ambient air temperature has significant impact on the low

frequency dynamics. To overcome the disturbance induced by the ambient environment, we have installed an isolation shield which is made from polystyrene, depicted in Figure 4.5. The temperature profile over 45 hours is given in Figure 4.6. We use the Local Polynomial Method (LPM) [135, 137] to identify the frequency



**Figure 4.6:** Temperature response of the PSA setup over 45 hour period. After 36 repetitions, the system starts to cool down. Set 1 covers the first 2 hours which contains 2 full period of multisine excitation. In this period, there is the initial response and a strong transient.

response function  $\mathcal{H}^{\text{obs}}(s)$  of the PSA setup with Set 1 indicated in Figure 4.6, since the LPM exploits the local smoothness, otherwise it would lead to a bias.

Then, we choose the initial expansion points: 1) the nominal value of each parameter is chosen as the initial parameter expansion points:  $p_1^0 = 475, p_2^0 = 900, p_3^0 = 44.5, p_4^0 = 27$  with 2 moments for each of these, 2) two initial frequency expansion points are at  $\{0, 1\}$  with the order of  $\{100, 10\}$ . In this chapter, the tolerance sets as  $\text{tol} = 0.05$  which is the resolution of frequency range  $\mathcal{S}$ , and an identity weighting matrix is used  $P = I$ .

With the above configuration, we can calibrate the 4 parameters using the Algorithm 2 and the Algorithm 3. The calibrated results are summarized in Table 4.2

	Order of $\Sigma_r(\theta)$	$\theta_1^{\text{opt}}$	$\theta_2^{\text{opt}}$	$\theta_3^{\text{opt}}$	$\theta_4^{\text{opt}}$
Alg. 2	440	457.31	944.24	40.13	25.23
Alg. 3 (Step 5)	440	457.31	944.24	40.13	25.23
Alg. 3 (1st loop in Step 6)	488	459.40	944.78	45.90	23.01
Alg. 3 (2nd loop in Step 6)	536	458.07	945.60	47.57	26.74
Alg. 3 (3rd loop in Step 6)	584	458.22	946.34	46.61	27.94

**Table 4.2:** Optimization results of two algorithms. For Algorithm 3.

Some remarks for Table 4.2:

- With the same expansion points and the associated moments, the calibrated results and the order of the reduced model from Algorithm 2 are the same as the calibration results and the reduced order after Step 5 in Algorithm 3.
- Before reaching the tolerance in Step 6, Algorithm 3, it has run three loops in total. The new calibrated results are obtained for each loop since the objective function changes due to a refined reduced-order model.
- 2 moments for every parameter expansion are added  $\tilde{p}_j = \theta_j^{\text{opt}}$  for  $j = 1, \dots, 4$  in each loop.
- After Step 6, item i) in each loop, the adapted frequency point  $\tilde{\sigma}_j, j = 1, \dots, 4$  which represents the most sensitive frequency point with respect to the calibration results from the previous loop, can be automatically determined by solving (4.26).
- For every adapted frequency point  $\tilde{\sigma}_j, j = 1, \dots, 4$ , we choose 5 moments for generating the reduced order model.

The adapted frequency interpolation points are listed in Table 4.3 The frequency

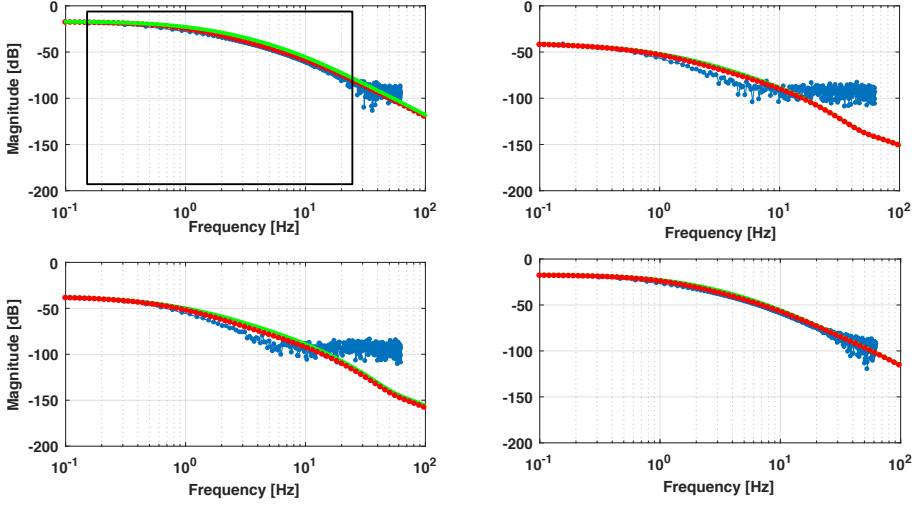
	$\tilde{\sigma}_1$	$\tilde{\sigma}_2$	$\tilde{\sigma}_3$	$\tilde{\sigma}_4$
Alg. 3 (1st loop in Step 6)	0.47	1.89	11.30	15.08
Alg. 3 (2nd loop in Step 6)	0.78	2.35	14.29	13.98
Alg. 3 (3rd loop in Step 6)	0.83	3.14	16.022	16.50

**Table 4.3:** The adapted frequency interpolation points for each loop in Algorithm 3.

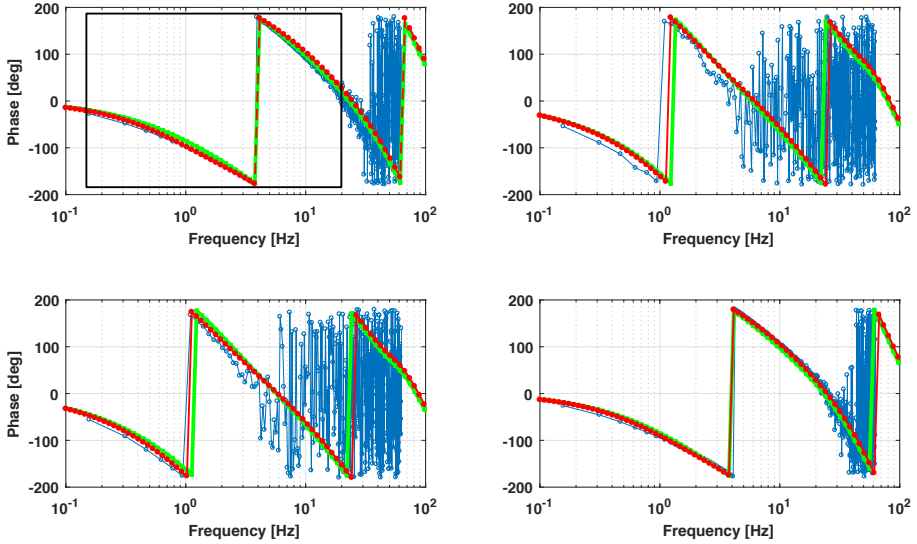
response of two parameterized reduced order models  $\Sigma_{r,2}(\theta^{\text{opt}}, s)$ ,  $\Sigma_{r,3}(\theta^{\text{opt}}, s)$  which are generated from the Algorithm 2 and the Algorithm 3 are given in Figure 4.7 and Figure 4.8, as well as the frequency response of the  $\mathcal{H}^{\text{obs}}(s)$ . Both the magnitude plot and the phase plot show that the  $\Sigma_{r,3}(\theta^{\text{opt}}, s)$  using the adaptive scheme introduces a less calibration error in comparison to the  $\Sigma_{r,2}(\theta^{\text{opt}}, s)$  obtained from Algorithm 3. Note that the order of  $\Sigma_{r,3}(\theta^{\text{opt}}, s)$  is 584 is higher than the order of  $\Sigma_{r,2}(\theta^{\text{opt}}, s)$  which is 440. Additionally, the Algorithm 3 takes longer computation time due to the adaptive scheme.

## 4.6 Conclusion

In this chapter, we have presented the calibration problems that involve large-scale models. An adapted optimization scheme that minimizes the calibration error has been combined with a parameterized model order reduction procedure. The new approach avoids re-sampling the parameter space for the parameter calibration and guarantees the equivalent Hermite interpolation condition for calibration problem between the full-order model and the reduced-order model. This method has been applied to a real-world application with measurement data. The calibration results show that the adaptive scheme introduced less calibration error in comparison to the current existing one.



**Figure 4.7:** Magnitude plot of  $2 \times 2$  systems: blue dot line denotes the  $\mathcal{H}^{\text{obs}}(s)$ , the red solid line represents the  $\Sigma_{r,3}(\theta^{\text{opt}}, s)$  using Algorithm 3, with a dimension of 584, and the green line is  $\Sigma_{r,2}(\theta^{\text{opt}}, s)$  using Algorithm 2, with a dimension of 440. The black box indicates the frequency range for calibration  $[0.15, 22]$  [Hz].



**Figure 4.8:** Phase plot of  $2 \times 2$  systems: blue dot line denotes the  $\mathcal{H}^{\text{obs}}(s)$ , the red solid line represents the  $\Sigma_{r,3}(\theta^{\text{opt}}, s)$  using Algorithm 3, with a dimension of 584, and the green line is  $\Sigma_{r,2}(\theta^{\text{opt}}, s)$  using Algorithm 2, with a dimension of 440. The black box indicates the frequency range for calibration  $[0.15, 22]$  [Hz].

# Model reduction of LFT systems using moment-matching method

---

In this chapter, model reduction methods are presented for a large-scale system with uncertainties, and a linear fractional transformation (LFT) represents is used to represent these model uncertainties. The reduction method involves a generalized definition of moments, parameterized moments and moment-matching based model reduction for LFT systems. The resulting methods apply to state order reduction for both uncertain linear time-invariant (LTI) systems and affine linear parameter-varying (LPV) systems with parametric uncertainties, represented by LFT's. By applying the proposed methods, we show that the reduced models preserve the moments/parameterized moments and the LFT structure. Based on the proposed methods, we present a parameterized model order reduction (pMOR) algorithm for LPV systems with bounded parameters. This algorithm preserves the parameters in the reduced model with fewer expansion points, and it leads to a smaller order reduced model. Numerical results based on a thermal system are carried out for both LTI and LPV models.

---

## 5.1 Introduction

The process of modelling and optimization for high-precision systems often results in models that have high state order and complicated uncertainty descriptions. These models are computationally expensive for controller/observer design

or analysis. For large-scale models without uncertainty, there exist many well-known state order reduction methods, examples of which include the interpolation based approaches for LTI systems [24, 130, 50, 15], balanced truncation [17, 30, 14, 83, 82], Krylov based method [85, 81, 75, 18, 23, 130, 84, 33, 21] and parametric models [29, 53, 170, 113]; and the reduced-based methods [143, 89, 97, 97]. Some Krylov based methods use the linear fractional transformation (LFT) structure to derive an iterative-type reduced order model algorithm for LTI systems. In [74, 131], the authors regarded the reduced order model (ROM) as the nominal model and the full order model (FOM) as the uncertain model. In contrast, there is little of work about reducing uncertain systems [168, 11, 10, 26]. In [26], the authors gave a systematic approach for the reduction of LTI systems with uncertainties, which is based on balancing the state variable. In addition [26] derived an error bound on the mismatch. This idea was extended to LPV systems in [25]. However, for models with high state dimension, MOR methods based on Krylov subspaces and moment matching are more attractive since it is simpler to implement and requires less computational effort than balanced reduction.

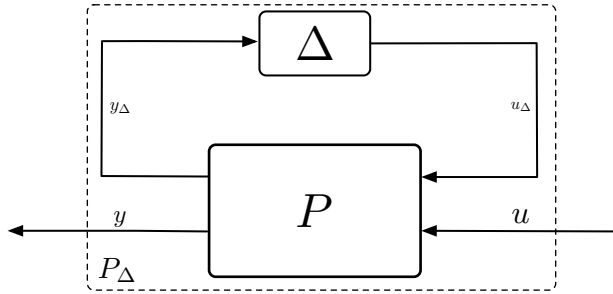
The idea of moment-matching for uncertain systems was first introduced in [167] where the authors analyzed and interpolated the uncertain systems at some points of the frequency response using the concept of *moment*. In such a way, the parametric variation of the model was approximated with the preservation of the first two moments. Besides this implementation of a moment matched reduction of an uncertain system, little work has been reported in the literature on the reduction of uncertain systems. There are still some important problems remaining open for the reduction of large-scale uncertain models. Specific problems include the question: i) whether the LFT structure for uncertainty descriptions is adequate for reduction purposes; ii) whether the uncertainty range of a model can be (or needs to be) preserved in the reduction process; iii) whether and how error bounds can be derived between an uncertain model and its simplification.

In this chapter, we will consider two types of uncertain systems: LTI systems and LPV systems. For the first type, we consider the setting of an uncertain large-scale LTI system that is described by LFTs on a repeated scalar uncertainty structure. To study the impact of the uncertainty set in the reduced model, we need the additional assumption that the uncertainty has a well defined upper and lower bound and has a diagonal structure. For the second one, we focus on a large-scale LPV model with parametric uncertainty that is related to the variation of physical parameters. More specifically, we consider a model with inaccurate density, heat capacitance, thermal conductivity coefficient due to process variation.

In Section 5.2, we give the background material on two types of uncertain systems. The model reduction for uncertain systems is considered in Section 5.3. Afterwards, Section 5.4 presents the results of MOR and pMOR for uncertain systems. We close this chapter with some conclusions and considerations on further work in Section 5.5.

## 5.2 Preliminaries

One way to model physical systems is through first principle modelling. Since in any modeling strategy there is always a mismatch between the model and the physical system. Such a model mismatch (or model discrepancy) can be viewed as a type of uncertainty in the underlying model. The field of system theory has developed various techniques to model the discrepancy between system and model and aim to quantify the mismatch, named uncertainty in this chapter. This chapter focuses on a system in which the performance is dominated by uncertainty, while the controller design or prediction is based on an idealized mathematical model that performs non-optimal or even poorly in the real physical system. The source of such uncertainty might be unknown nonlinearities, unmodelled dynamics, parameter variation, or the combination of any of these. To have a mathematical description of a system with uncertainty, we employ the LFT paradigm, which is depicted in Figure 5.1.



**Figure 5.1:** Example of a uncertain system  $P_\Delta$  consists of the nominal plant  $P$  and the uncertainty block  $\Delta$ .

In Figure 5.1 the system with block  $P$  represents the nominal model of the system, which is the input-output relation between input  $u$  and output  $y$  of the plant at the nominal values. The external signals of the uncertain system are the input  $u$  and the output  $y$  while the internal signals are denoted as  $y_\Delta$  and  $u_\Delta$ . The block indicated by  $\Delta$  denotes an uncertain system that is not fixed and is only known to be an element of a class of systems, say  $\mathbf{\Delta}$ , which we refer to as the uncertainty class. This class may be dynamic or static, linear or nonlinear, bounded with respect to various gains, structured or non-structured in its relation between  $u_\Delta$  and  $y_\Delta$ . A specific example of a structured static uncertainty set

$$\mathbf{\Delta} = \{\Delta(\lambda) \mid \lambda \in \mathbb{C} \text{ and } I - A\lambda \text{ is regular}\} \quad (5.1)$$

In particular, if the system realization matrix  $P$  is partitioned as

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix},$$

the input-output mapping of the interconnection in Figure 5.1 is defined by map-

ping  $\text{col}(u_\Delta, u)$  to  $\text{col}(y_\Delta, y)$  as a transfer function of an LTI system. Then let  $\Delta : y_\Delta \rightarrow u_\Delta$  and assume, for the time being, that  $\Delta$  is a linear mapping belonging to a class  $\mathbf{\Delta}$  of linear mappings. Then the configuration of Figure 5.1 represents an uncertain transfer function  $P_\Delta$  from input  $u$  to output  $y$  that is given by

$$y = (\Delta \star P)u = (P_{22} + P_{21}\Delta(I - P_{11}\Delta)^{-1}P_{12})u.$$

Here the Redheffer star notation  $\star$  is only used where it is understood that the inverse of  $(I - P_{11}\Delta)$  exists. Note that we require this nonsingular condition for all  $\Delta \in \mathbf{\Delta}$  to guarantee that  $\Delta \star P$  is a well defined class of uncertain models for all  $\Delta \in \mathbf{\Delta}$ . We mention here that, throughout this chapter, we call  $P_{22} + P_{21}\Delta(I - P_{11}\Delta)^{-1}P_{12}$  the *transfer operator* which is associated with  $\Delta \in \mathbf{\Delta}$ .

### 5.2.a LFR representations of LTI systems

First, we consider an LTI system of the form

$$\Sigma := \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \quad (5.2)$$

where  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$  and  $y(t) \in \mathbb{R}^q$  denote, respectively, the state vector, the input and the output. The  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ ,  $C \in \mathbb{R}^{q \times n}$  and  $D \in \mathbb{R}^{q \times m}$  are given system matrices. In most practical cases, we can assume that the state dimension is much higher than the input dimension and output dimension,  $n \gg q$  and  $n \gg m$ . The matrix pencil  $A - \sigma I$  is assumed to be *regular*. Let  $P$  be the static matrix

$$P = \begin{bmatrix} A & B \\ C & D \end{bmatrix}. \quad (5.3)$$

If we let  $\Delta$  denote the repeated copies of the shift operator, e.g.,  $\Delta = \lambda I$ ,  $\Delta \in \mathbf{\Delta}$  in (5.1) and  $I$  with dimension of  $n \times n$  in (5.1), then the input-output mapping of this uncertain system  $(\Delta, P)$  can be represented as

$$y = (\Delta \star P)u = (D + C\lambda(I - A\lambda)^{-1}B)u. \quad (5.4)$$

Note that with  $u_\Delta = x$ ,  $y_\Delta = \dot{x}$  in Figure 5.1 and  $\lambda = \frac{1}{s}$ , the equation in (5.4) is equal to the transfer function  $H(s) = C(sI - A)^{-1}B + D$ . Specifically,  $\Delta \star P = G(\lambda)$ .

### 5.2.b LFR representations of uncertain LPV systems

Second, we consider a linear parameter-varying model with parameters  $\theta \in \Theta$ ,  $\Theta \subset \mathbb{R}^\ell$  described as

$$\Sigma(\theta) := \begin{cases} \dot{x}(t) = A(\theta)x(t) + B(\theta)u(t) \\ y(t) = C(\theta)x(t) + D(\theta)u(t) \end{cases} \quad (5.5)$$

where  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$  and  $y(t) \in \mathbb{R}^q$  denote, respectively, the state vector, the input and the output. The state-space matrices are functions  $A : \mathbb{R}^\ell \rightarrow \mathbb{R}^{n \times n}$ ,



$B : \mathbb{R}^\ell \rightarrow \mathbb{R}^{n \times m}$ ,  $C : \mathbb{R}^\ell \rightarrow \mathbb{R}^{q \times n}$  and  $D : \mathbb{R}^\ell \rightarrow \mathbb{R}^{q \times m}$ . The system matrices are assumed to have affine dependence on the parameter vector, which means that we can write

$$\begin{aligned} A(\theta) &= A_0 + \sum_{i=1}^{\ell} \theta_i A_i & B(\theta) &= B_0 + \sum_{i=1}^{\ell} \theta_i B_i \\ C(\theta) &= C_0 + \sum_{i=1}^{\ell} \theta_i C_i & D(\theta) &= D_0 + \sum_{i=1}^{\ell} \theta_i D_i. \end{aligned} \quad (5.6)$$

In this work, we assume that  $A(\theta)$  is Hurwitz for all  $\theta \in \Theta$ . Furthermore, the parameter  $\theta \in \Theta$  is not varying with time, i.e.  $\frac{d}{dt}\theta = 0$  but with physically meaningful constraints  $\underline{\theta}_i \leq \theta_i \leq \bar{\theta}_i$  and  $-\infty < \underline{\theta}_i \leq \bar{\theta}_i < \infty$ . This means that  $\Theta = \prod_{i=1}^{\ell} [\underline{\theta}_i, \bar{\theta}_i]$ , i.e.  $\Theta$  has Cartesian structure. With such upper and lower bounds, we can define the nominal value of  $i$ th parameter according to

$$\theta_i^0 = \frac{\bar{\theta}_i + \underline{\theta}_i}{2} \quad \text{with} \quad \underline{\theta}_i \leq \theta_i \leq \bar{\theta}_i, \quad \underline{\theta}_i, \bar{\theta}_i \in \mathbb{R}$$

and introduce the scaled error  $\delta_i : \mathbb{R} \rightarrow \mathbb{R}$  defined by the expression

$$\delta_i(\theta_i) = \frac{\theta_i - \theta_i^0}{\bar{\theta}_i - \theta_i^0}$$

This defines then  $\Delta(\delta) : \Theta \rightarrow \mathbb{R}^{n \times n}$  and we write this as  $\Delta(\delta) = \{\text{diag}(\delta_i(\theta_i)I) \mid i = 1, \dots, \ell\}$ . Thus, we have  $\theta_i = \theta_i^0 + w_i \delta_i(\theta_i)$  where the weighting function  $w_i$  is defined as  $w_i = \bar{\theta}_i - \theta_i^0$  and  $|\delta_i(\theta_i)| \leq 1$ . By reformulating the system (5.5) in such a way, we shift and normalize the uncertain system to the nominal system, which is the centre of the uncertainty set with respect to the parametric uncertainty. Therefore, system matrices of (5.5) with  $\theta \in \Theta$  can be represented as

$$\begin{bmatrix} A(\theta) & B(\theta) \\ C(\theta) & D(\theta) \end{bmatrix} = \begin{bmatrix} A(\theta^0) & B(\theta^0) \\ C(\theta^0) & D(\theta^0) \end{bmatrix} + \sum_{i=1}^{\ell} \delta_i \begin{bmatrix} w_i A_i & w_i B_i \\ w_i C_i & w_i D_i \end{bmatrix}. \quad (5.7)$$

Now factorize

$$\begin{bmatrix} w_i A_i & w_i B_i \\ w_i C_i & w_i D_i \end{bmatrix} = \begin{bmatrix} L_i^1 \\ L_i^2 \end{bmatrix} \begin{bmatrix} R_i^1 & R_i^2 \end{bmatrix} = \begin{bmatrix} L_i^1 R_i^1 & L_i^1 R_i^2 \\ L_i^2 R_i^1 & L_i^2 R_i^2 \end{bmatrix} \quad (5.8)$$

where  $\begin{bmatrix} L_i^1 \\ L_i^2 \end{bmatrix}$  and  $[R_i^1, R_i^2]$  have full column and row rank, respectively. In this way, the system (5.5) can be described as

$$\begin{bmatrix} y_{1,\Delta} \\ \vdots \\ y_{\ell,\Delta} \\ \eta \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & \cdots & 0 & R_1^1 & R_1^2 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & R_\ell^1 & R_\ell^2 \\ \hline L_\ell^1 & \cdots & L_\ell^1 & A(\theta^0) & B(\theta^0) \\ L_\ell^2 & \cdots & L_\ell^2 & C(\theta^0) & D(\theta^0) \end{bmatrix}}_{P^\theta} \begin{bmatrix} u_{1,\Delta} \\ \vdots \\ u_{\ell,\Delta} \\ x \\ u \end{bmatrix} \quad (5.9)$$

with

$$\begin{bmatrix} u_{1,\Delta} \\ \vdots \\ u_{\ell,\Delta} \end{bmatrix} = \begin{bmatrix} \delta_1(\theta_1)I & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \delta_\ell(\theta_\ell)I \end{bmatrix} \begin{bmatrix} y_{1,\Delta} \\ \vdots \\ y_{\ell,\Delta} \end{bmatrix} = \Delta(\delta) \begin{bmatrix} y_{1,\Delta} \\ \vdots \\ y_{\ell,\Delta} \end{bmatrix}.$$

In order to avoid confusion regarding the notation of the matrix  $P$  in (5.3), the matrix in the equation (5.9) is denoted by  $P^\theta$ , which is algebraic and the parameter variations are captured in the uncertainty  $\Delta(\delta)$ .

**Theorem 5.1** The set of matrices  $\begin{bmatrix} A(\theta), B(\theta) \\ C(\theta), D(\theta) \end{bmatrix}$  defined in (5.6) with  $\theta \in \Theta$ ,  $\Theta := \Pi_{i=1}^\ell [\underline{\theta}_i, \bar{\theta}_i]$  can be rewritten as

$$\begin{bmatrix} A(\theta^0) + \sum_{i=1}^\ell \delta_i L_i^1 R_i^1 & B(\theta^0) + \sum_{i=1}^\ell \delta_i L_i^1 R_i^2 \\ C(\theta^0) + \sum_{i=1}^\ell \delta_i L_i^2 R_i^1 & D(\theta^0) + \sum_{i=1}^\ell \delta_i L_i^2 R_i^2 \end{bmatrix} = \begin{bmatrix} A(\delta) & B(\delta) \\ C(\delta) & D(\delta) \end{bmatrix} \\ = (\Delta(\delta) \star P^\theta)$$

with  $\Delta(\delta) = \{\text{diag}(\delta_i I) \mid i = 1, \dots, \ell\}$  where  $|\delta_i| \leq 1$  for all  $i = 1, \dots, \ell$ , and  $P^\theta$  is given in (5.9).

**Proof:** See the derivation in (5.7) and (5.8) proceeding the theorem.  $\square$

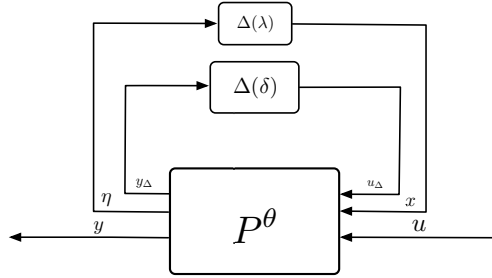
**Remark 5.1** We assume that all weighting factors  $w$  are accurate to describe the parametric uncertainty range and real-valued functions. In the above derivation, the weighting functions are absorbed in the nominal model. As mentioned in [147], the alternative approach is to incorporate the weighting functions in  $\Delta(\delta) = \text{diag}([w_1 \delta_1 I, \dots, w_\ell \delta_\ell I])$ . In this chapter, we choose the former approach.

**Remark 5.2** As mentioned in [147], the factorization in (5.8) is chosen such that the number of columns/rows of the factors is minimal. This minimality results in minimal blocks of the uncertainty  $\Delta(\delta)$ . If an arbitrary factorization is applied, it could lead to a larger factorization and inefficient system realization.

To describe the transfer function of the uncertain parametric system (5.5), let us take a second uncertainty function  $\Delta(\lambda) = \lambda I$ . Then the input-output mapping or transfer operator of the system (5.5) can be represented by Redheffer star product, that is

$$\begin{aligned} y &= (\Delta(\lambda) \star (\Delta(\delta) \star P^\theta))u \\ &= (\lambda I \star \begin{bmatrix} A(\delta) & B(\delta) \\ C(\delta) & D(\delta) \end{bmatrix})u \\ &= [D(\delta) + C(\delta)\lambda I(I - A(\delta)\lambda I)^{-1}(B(\delta))]u. \end{aligned} \quad (5.10)$$

The system model  $(\Delta(\lambda), (\Delta(\delta) \star P^\theta))$  is depicted in Figure 5.2.



**Figure 5.2:** System  $(\Delta(\lambda), (\Delta(\delta) \star P^\theta))$  where  $\Delta(\delta(\theta))$  represents the parametric uncertainty and the general uncertainty  $\Delta(\lambda)$  is denoted by shift operator  $\Delta(\lambda) = \lambda I$

**Remark 5.3** Indeed, by combining the  $\Delta(\lambda)$  and  $\Delta(\delta)$  as one uncertainty block, such as in [167]

$$\tilde{\Delta}(\lambda, \delta) = \left[ \begin{array}{c|c} \Delta(\lambda) & \Delta(\delta) \end{array} \right] = \left[ \begin{array}{c|ccc} \lambda I & \delta_1(\theta_1)I & \cdots & 0 \\ \hline & \vdots & & \vdots \\ & 0 & & \delta_\ell(\theta_\ell)I \end{array} \right] \quad (5.11)$$

the expression of the transfer operator  $(\lambda I \star (\Delta(\delta) \star P^\theta))$  can be simplified to  $(\tilde{\Delta}(\lambda, \delta) \star P^\theta)$ . Nevertheless, we intend to distinguish the shift operator  $\lambda I$  from the parametric uncertainty  $\Delta(\delta)$  such that the impact of the parameter variation to the plant is differentiable.

**Theorem 5.2** The transfer function  $H(s, \theta)$  for the uncertain system in (5.5) belongs to the class of uncertain systems  $G := \{G(s, \theta) \mid \theta \in \Theta\}$  that satisfies

$$G = \{\Delta(\tilde{\lambda}, \delta) \star P^\theta \mid \tilde{\Delta}\}$$

assuming the form  $\tilde{\Delta}(\lambda, \delta)$  in (5.11) with  $\Delta \in \mathbf{\Delta}$  and  $\lambda = \frac{1}{s}$ .

### 5.2.c Projection-based MOR

Suppose (5.2) is a stable system with a high dimensional state vector typically  $n \geq 10^4$ . It is desired to have a low-order system realization  $r \ll n$ . We employ the projection-based method to find the reduced order model. Define the projection spaces

$$\mathcal{V} := \text{im}(V); \quad \mathcal{W} := \text{im}(W). \quad (5.12)$$

$W, V \in \mathbb{R}^{n \times r}$  are full rank matrices with  $r \ll n$ . Let  $\Pi_{\mathcal{V}}$  and  $\Pi_{\mathcal{W}}$  denote the (canonical) projections of  $\mathbb{R}^n$  onto  $\mathcal{V}$  and  $\mathcal{W}$ , respectively. That is  $\Pi_{\mathcal{V}} = V(V^T V)^{-1}V^T$  and  $\Pi_{\mathcal{W}} = W(W^T W)^{-1}W^T$ . Then, the reduced state is defined

$$x_r := (V^T V)^{-1}V^T x,$$

and satisfies  $Vx_r = \Pi_{\mathcal{V}}x$ . Since  $Vx_r$  is the orthogonal projection of  $x$  in  $\mathcal{V}$  we view  $Vx_r$  as the best approximation of  $x$  in  $\mathcal{V}$  and view the  $r$  dimensional vector  $x_r$  of coefficients of  $x$  in the projected space  $\mathcal{V}$  as the reduced order state.  $x_r$  is an approximation of the state  $x$  in (5.2) and has a dimension of  $r$ . Replacing  $x$  by its approximation  $Vx_r$  incurs a residual error

$$\rho := V\dot{x}_r - AVx_r - Bu$$

in the state evolution equation of (5.2). The projection  $\Pi_{\mathcal{W}}$  is meant to require that  $\Pi_{\mathcal{W}}\rho = 0$  for all time instants  $t$ . Equivalently,

$$W^T \rho = W^T V\dot{x}_r - W^T AVx_r - W^T Bu = 0.$$

Combing the state and the residual projection, we have the reduced system realization

$$\Sigma := \begin{cases} \dot{x}_r(t) &= A_r x_r(t) + B_r u(t) \\ y(t) &= C_r x_r(t) + D_r u(t) \end{cases} \quad (5.13)$$

where the system matrices are

$$\begin{aligned} A_r &= W^T AV, & B_r &= W^T B \\ C_r &= CV & D_r &= D. \end{aligned}$$

This method can also be extended to LPV systems. Consider the system in (5.5), assume the system is stable for all  $\theta \in \Theta$ . Using the projection-based MOR as described above, the reduced model is

$$\Sigma_r(\theta) := \begin{cases} \dot{x}_r(t) &= A_r(\theta)x_r(t) + B_r(\theta)u(t) \\ y(t) &= C_r(\theta)x_r(t) + D_r(\theta)u(t), \end{cases} \quad (5.14)$$

the reduced system matrices are defined as  $A_r(\theta) = W^T A(\theta)V$ ,  $B_r(\theta) = W^T B(\theta)$ ,  $C_r(\theta) = C(\theta)V$  and  $D_r(\theta) = D(\theta)$ . This reads

$$A_r(\theta) = W^T \left( A_0 + \sum_{i=1}^{\ell} \theta_i A_i \right) V = A_{r,0} + \sum_{i=1}^{\ell} \theta_i A_{r,i}$$

$$\begin{aligned}
B_r(\theta) &= W^T(B_0 + \sum_{i=1}^{\ell} \theta_i B_i) = B_{r,0} + \sum_{i=1}^{\ell} \theta_i B_{r,i} \\
C_r(\theta) &= (C_0 + \sum_{i=1}^{\ell} \theta_i C_i)V = C_{r,0} + \sum_{i=1}^{\ell} \theta_i C_{r,i} \\
D_r(\theta) &= D(\theta).
\end{aligned}$$

Hence, if (5.5) affinely depends on the parameters  $\theta$ , so does (5.14). If  $W = V$ , this is called an (ordinary) Galerkin projection. If  $W \neq V$ , this is a Petrov-Galerkin projection. For a high-order state model, the construction of  $V, W$  using the SVD-based method often is computationally expensive or infeasible. In those cases, a moment-matching method is much easier and more efficient to generate the projection matrices.

### 5.3 Model reduction of LFT systems

Before considering the construction of the projection matrices  $V, W$  for uncertain systems, we first present the moments for uncertain systems. Then, the reduced order model for uncertain systems generated via moment-matching is given.

#### 5.3.a Moment and model reduction for LTI systems in LFT representation

Consider the uncertain LTI system model  $(\Delta, P)$  where the uncertainty set  $\Delta \in \mathbf{\Delta}$  is defined as in (5.1) and  $P$  has the realization in (5.3). Suppose  $A\Delta$  is nonsingular for all  $\Delta \in \mathbf{\Delta}$ . Using this condition for conceptual verification of the regularity of the matrix pencil  $(I - A\Delta)$ , we can derive an efficient and easily implemented regularity-based condition for uncertain systems.

With the condition of the regularity of the matrix pencil  $(I - A\Delta)$ , we proceed to the definition of *moments*. The transfer operator of the model  $(\Delta, P)$  with  $\Delta = \lambda I$  is defined as

$$G(\lambda) := \Delta \star P = D + C\lambda(I - A\lambda)^{-1}B.$$

Consider analytic expansion of  $G(\lambda)$  around the point  $\lambda_0 \in \mathbb{C}$ . Thus

$$G(\lambda) = \sum_{k=0}^{\infty} m_k(\lambda - \lambda_0)^k = m_0 + m_1(\lambda - \lambda_0) + m_2(\lambda - \lambda_0)^2 + \dots$$

with moments  $m_k = m_k(\lambda_0)$ ,  $k \geq 0$ . Then

$$m_k(\lambda_0) = \frac{1}{k!} \frac{d^k}{d\lambda^k} G(\lambda) \big|_{\lambda=\lambda_0} = \frac{1}{k!} G^{(k)}(\lambda_0), k \geq 0$$

and moments  $m_k$  satisfy Cauchy integral expression

$$m_k(\lambda_0) = \frac{1}{2\pi i} \oint \frac{G(\lambda)}{(\lambda - \lambda_0)^{k+1}} d\lambda$$

counterclockwise integration over any closed curve in  $\mathcal{C}$  enclosing  $\lambda_0$  and belonging to domain of analyticity of  $G$ .

**Theorem 5.3** *Let  $\lambda_0 \in \mathbb{C}$ . Then the  $k$ -th moment  $m_k = m_k(\lambda_0)$  of  $G(\lambda) = (\Delta \star P)(\lambda)$  with  $P = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ ,  $\Delta = \Delta(\lambda) := \lambda I_n \in \mathbf{\Delta}$  is given by*

$$m_k(\lambda_0) = \begin{cases} D + C\lambda_0(I - \lambda_0 A)^{-1}B & \text{if } k = 0 \\ \mathcal{C}(\lambda_0)[\mathcal{A}(\lambda_0)]^{k-1}\mathcal{B}(\lambda) & \text{if } k \geq 1. \end{cases} \quad (5.15)$$

where  $\mathcal{C}(\lambda) := C(I - \lambda A)^{-1}$ ;  $\mathcal{B}(\lambda) := (I - \lambda A)^{-1}B$ , and  $\mathcal{A}(\lambda) := A(I - \lambda A)^{-1}$ .

Notice that for a special case  $\lambda_0 = 0$ , (5.15) can be simplified to

$$m_k(0) = \begin{cases} D & \text{if } k = 0 \\ CA^{k-1}B & \text{if } k \geq 1. \end{cases}$$

Next suppose that  $\lambda_0 \in \mathbb{C}$  is given and we wish to match  $2N$  moments  $m_0(\lambda_0), \dots, m_{2N-1}(\lambda_0)$  in a reduced order  $r < n$  model. Define a full rank matrix  $V$  such that  $\mathcal{V} = \text{im}V$  is given by the Krylov subspace

$$\mathcal{V} = \text{colspan}\{f, Mf, \dots, M^{N-1}f\} \quad (5.16)$$

where

$$\begin{aligned} f &:= \mathcal{B}(\lambda_0) = (I - A\lambda_0)^{-1}B \in \mathbb{R}^{n \times m} \\ M &:= \mathcal{A}(\lambda_0) = A(I - A\lambda_0)^{-1} \in \mathbb{R}^{n \times n} \end{aligned}$$

Similarly,

$$\mathcal{W} = \text{colspan}\{l, M^T, \dots, (M^T)^{N-1}l\} \quad (5.17)$$

where

$$\begin{aligned} l &:= \mathcal{C}(\lambda_0)^T = (I - A\lambda_0)^{-T}C^T \in \mathbb{R}^{n \times q} \\ M^T &:= \mathcal{A}(\lambda_0)^T = (I - A\lambda_0)^{-T}A^T \in \mathbb{R}^{n \times n} \end{aligned}$$

and let  $W$  be any full rank matrix such that  $\mathcal{W} = \text{im}(W)$ , and  $r = N \times m$ .

Based on the discussion above, the reduced LTI system in LFR representation is

defined as  $(\Delta_r, P_r)$  with  $r \ll n$ , by setting:

$$P_r = \begin{bmatrix} A_r & B_r \\ C_r & D_r \end{bmatrix}$$

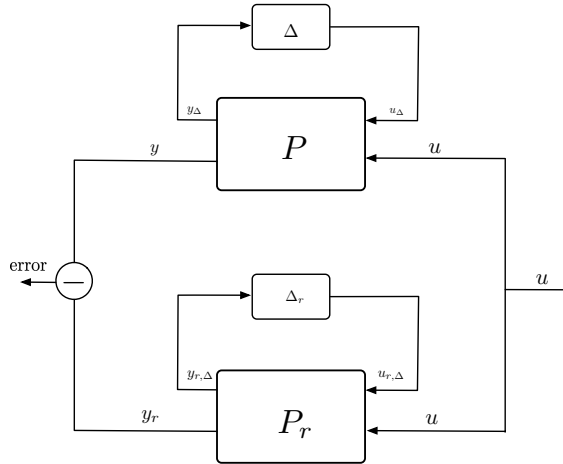
with

$$\begin{aligned} A_r &= W^T A V \in \mathbb{R}^{N \times N} & B_r &= W^T B \in \mathbb{R}^{N \times m} \\ C_r &= C V \in \mathbb{R}^{q \times N} & D_r &= D \in \mathbb{R}^{q \times m}, \end{aligned}$$

and the uncertainty set is

$$\Delta_r \in \Delta_r, \quad \Delta_r = \Delta$$

where  $\Delta_r$  represents a lower order uncertainty set of  $\Delta \in \Delta$  defined in (5.1). We again assume that the  $\Delta_r$  represents the shift operator  $\Delta_r(\lambda) = \lambda I_r$ .



**Figure 5.3:** Error system  $(\Delta_e, P_e)$  between two LFTs models remains LFT.

The difference between the full order model  $(\Delta(\lambda), P)$  and reduced order model  $(\Delta_r(\lambda), P_r)$  can be formed as the difference realization of these two models, that is

$$(\Delta_e \star P_e) = (\Delta(\lambda) \star P) - (\Delta_r(\lambda) \star P_r) \quad (5.18)$$

with

$$P_e = \begin{bmatrix} A & 0 & B \\ 0 & A_r & B_r \\ C & -C_r & D - D_r \end{bmatrix} \quad \text{and} \quad \Delta_e = \begin{bmatrix} \Delta & 0 \\ 0 & \Delta_r \end{bmatrix}$$

The interconnection of the  $(\Delta_e, P_e)$  is shown in Figure 5.3.

**Remark 5.4** Notice that the realization of the difference between the reduced and the original uncertain systems represented by the LFT, is an LFT again with block uncertainty  $\Delta_e$  satisfying  $\Delta_e(\lambda) = \lambda I_{n+r}$

Different from SVD-based MOR techniques, moment-matching MOR method is preserving the local information around the expansion point, which does not have global error bound. In what follows, we present the lemma using the result for LTI systems in Theorem 2.3.

**Theorem 5.4** *Let  $\lambda_0 \in \Delta$  be chosen such that both the matrix pencil  $(I - A\lambda_0), (I_r - A_r\lambda)$  are regular, and both the transfer operators  $(\lambda I_n \star P), (\lambda I_r \star P_r)$  are analytic in the neighbourhood of  $\lambda_0$ , then the Hermite interpolation condition is satisfied as*

$$\frac{d^i}{d\lambda^i}(\lambda I_n \star P) \big|_{\lambda=\lambda_0} = \frac{d^i}{d\lambda^i}(\lambda I_r \star P_r) \big|_{\lambda=\lambda_0}, \quad \text{for } i = 0, \dots, N-1.$$

*This yields a reduced model whose transfer operator  $(\lambda I_r \star P_r)$  coincides in as many moments as possible for a given order.*

**Proof:** We refer the proof to Section 2.3. □

Based on Theorem 5.4, the error system defined in (5.18) also satisfies the Hermite condition, that is

$$\frac{d^i}{d\lambda^i}(\Delta_e \star P_e)(\lambda_0) = 0 \quad \text{for } i = 0, \dots, N-1$$

$$\text{where } \Delta_e = \begin{bmatrix} \lambda I_n & 0 \\ 0 & \lambda I_r \end{bmatrix} = \lambda I_{n+r}.$$

A generic implementation of the moment-matching for uncertain LTI systems as described in (5.16) and (5.17) can be carried out by using Arnoldi algorithm, and the details we refer to [Section 3.3, [17]].

### 5.3.b Moments and model reduction for uncertain LPV systems

Similarly, we present the moment for the LPV systems with parametric uncertainty. Consider the LFT paradigm shown in Figure 5.2, where  $P^\theta$  denotes the nominal plant as in (5.9),  $\Delta(\delta)$  specifies the parametric uncertainty which is represented by variables  $\delta$ , and  $\Delta$  represents the general uncertainty that can be represented as copies of the shift operator  $\Delta = \lambda I$ .

Assume the model  $(\Delta(\delta), P^\theta)$  is stable for every  $\Delta(\delta) \in \Delta$ , that is, all the generalized eigenvalues of the matrix pencil  $(I - A(\delta)\Delta)$  lie in the left half open plane. Suppose the transfer operator  $G(\lambda, \delta) = (\lambda I \star (\Delta(\delta) \star P^\theta))$  is analytic in the neighbourhood of the expansion point  $(\lambda_0, \delta_0)$ . Then, the multivariable Taylor series expansion of  $G(\lambda, \delta)$  is given by the expression derived in (2.35) and reads

$$\begin{aligned} G(\lambda, \delta) &= G(\lambda_0, \delta_0) + [D^1 G] \begin{pmatrix} \lambda - \lambda_0 \\ \delta - \delta_0 \end{pmatrix} \\ &+ \dots + [D^{|k|} G] \left( \begin{pmatrix} \lambda - \lambda_0 \\ \delta - \delta_0 \end{pmatrix}, \dots, \begin{pmatrix} \lambda - \lambda_0 \\ \delta - \delta_0 \end{pmatrix} \right) + \dots \end{aligned}$$



where  $[D]^{|k|}G$  is defined in (2.32). This can be written as

$$G(\lambda, \delta) = \sum_{I=0}^{\infty} \sum_{|J|=0}^{\infty} m_{I,J}(\lambda_0, \delta_0)(\lambda - \lambda_0)^I(\delta - \delta_0)^J$$

where  $I$  is a integer and  $J$  is a multi-index  $J = (j_1, \dots, j_\ell)$  with length  $|J| = \sum_{k=1}^{\ell} j_k$  and  $(\delta - \delta_0)^J$  is to be interpreted as

$$(\delta - \delta_0)^J = (\delta_1 - \delta_{0,1})^{j_1} \dots (\delta_\ell - \delta_{0,\ell})^{j_\ell}.$$

Here  $m_{I,J}(\lambda_0, \delta_0)$  denotes the  $(I, J)$ -th moment of the uncertain system  $G(\lambda, \delta) = (\lambda I_n, (\Delta(\delta) \star P^\theta))$  and is given

$$\begin{aligned} G(\lambda, \delta) &= G(\lambda_0, \delta_0) + \frac{\partial}{\partial \lambda} G(\lambda_0, \delta_0) \left( \frac{1}{\lambda} - \frac{1}{\lambda_0} \right) + \frac{\partial}{\partial \delta} G(\lambda_0, \delta_0) (\delta - \delta_0) + \dots \\ &= \sum_{I=0, J=0} m_{\{I, J\}}(\lambda_0, \delta_0) \left( \frac{1}{\lambda} - \frac{1}{\lambda_0} \right)^I (\delta - \delta_0)^J \end{aligned}$$

and where

$$m_{I,J}(\lambda_0, \delta_0) = \frac{1}{(I+J)!} \frac{\partial^{I+J}}{\partial \lambda^I \partial \delta^J} G(\lambda_0, \delta_0)$$

for  $I \geq 0$  and  $J \geq 0$ . Here  $m_{I,J}(\lambda_0, \delta_0)$  denotes the  $\{I+J\}$ th moment of the uncertain system  $(\lambda I, (\Delta(\delta) \star P^\theta))$  is defined below.

**Definition 5.5** Given the system in (5.5) with parameters  $\theta \in \Theta, \Theta = \Pi_{i=1}^{\ell} [\theta_i, \bar{\theta}_i]$  and the uncertainty set  $\Delta = \lambda I, \Delta \in \mathbf{\Delta}$  defined in (5.1). Assume the matrix pencil  $(I - A(\delta)\Delta)$  with  $\delta = \text{vec}(\delta_i), |\delta_i| \leq 1, i = 1, \dots, \ell$  is regular and the transfer function  $G(\lambda, \delta)$  is analytic in the neighbourhood of  $(\lambda_0, \delta_0)$ , then the  $\{I+|J|\}$ th moment around  $(\lambda_0, \delta_0)$  is defined as

$$m_{I,J}(\lambda_0, \delta_0) = \frac{1}{(I+|J|)!} \frac{\partial^{I+J}}{\partial \lambda^I \partial \delta^J} G(\lambda_0, \delta_0), \quad \text{with } I \geq 0, |J| \geq 0.$$

Note that if  $\ell = 1$ , then  $J = j_1$  and  $|J| = J = j_1$  in which case

$$\begin{aligned} G(\lambda, \delta) &= G(\lambda_0, \delta_0) + \frac{\partial}{\partial \lambda} G(\lambda_0, \delta_0) + \frac{\partial}{\partial \delta} G(\lambda_0, \delta_0) (\delta - \delta_0) + \dots \\ &= \sum_{I=0}^{\infty} \sum_{J=0}^{\infty} m_{I,J}(\lambda_0, \delta_0) (\lambda - \lambda_0)^I (\delta - \delta_0)^J. \end{aligned}$$

Based on the above definition, the first three moments of (5.5) are:

$$\begin{aligned}
 m_{0,0}(\lambda_0, \delta_0) &= D(\delta_0) + C(\delta_0)\lambda_0 I(I - A(\delta_0)\lambda_0 I)^{-1}B(\delta_0) \\
 m_{1,0}(\lambda_0, \delta_0) &= C(\delta_0)(I - A(\delta_0)\lambda_0 I)^{-1}I(I - A(\delta_0)\lambda_0 I)^{-1}B(\delta_0) \\
 m_{0,1}(\lambda_0, \delta_0) &= C_{\delta_0}\lambda_0 I(I - A(\delta_0)\lambda_0 I)^{-1}B(\delta_0) + \\
 &\quad C(\delta_0)\lambda_0 I(I - A(\delta_0)\lambda_0 I)^{-1}B_{\delta_0} + \\
 &\quad C(\delta_0)\lambda_0 I(I - A(\delta_0)\lambda_0 I)^{-1}A_{\delta_0}(I - A(\delta_0)\lambda_0 I)^{-1}B(\delta_0)
 \end{aligned}$$

here  $C_{\delta_0} = \frac{d}{d\delta}C(\delta)$ ,  $B_{\delta_0} = \frac{d}{d\delta}B(\delta)$  and  $A_{\delta_0} = \frac{d}{d\delta}A(\delta)$  at  $\delta = \delta_0$ . In the same manner, the rest of moments for the uncertain LPV system can be derived.

Let us define two sets  $\{\lambda_1, \dots, \lambda_k\}$  and  $\{\delta_1, \dots, \delta_\ell\}$ , we can extend the above derivation to multiple shift operator and multiple parameter expansion points. Here we abuse the notation of  $\ell$  which indicates both the dimension of a parameter space and the dimension of parameter sampling point. The projection matrices  $V, W$  are defined as any full rank matrices that satisfy (5.12) with follows

$$\mathcal{V} = \text{im}[v(\lambda_1, \delta_1), \dots, v(\lambda_k, \delta_\ell)] \quad (5.19)$$

$$\mathcal{W} = \text{im}[w(\lambda_1, \delta_1), \dots, w(\lambda_k, \delta_\ell)] \quad (5.20)$$

where

$$\begin{aligned}
 v(\lambda_i, \delta_j) &= \mathcal{K}_{\sigma_{i,j}^v}(M_{i,j}^v, F_{i,j}) \\
 w(\lambda_i, \delta_j) &= \mathcal{K}_{\sigma_{i,j}^w}(M_{i,j}^w, L_{i,j})
 \end{aligned} \quad (5.21)$$

with

$$\begin{aligned}
 F_{i,j} &= (I - A(\delta_j)\lambda_i)^{-1}B_{\delta_j} \\
 M_{i,j}^v &= A_{\delta_j}(I - A(\delta_j)\lambda_i)^{-1} \\
 L_{i,j} &= (I - A(\delta_j)\lambda_i)^{-T}C_{\delta_j}^T \\
 M_{i,j}^{wT} &= (I - A(\delta_j)\lambda_i)^{-T}A_{\delta_j}^T.
 \end{aligned}$$

Here  $B_{\delta_j} = \frac{d}{d\delta_j}(B(\theta))$ ,  $A_{\delta_j} = \frac{d}{d\delta_j}(A(\theta))$ ,  $C_{\delta_j} = \frac{d}{d\delta_j}(C(\theta))$  and  $\sigma_{i,j}^v, \sigma_{i,j}^w$  represent the number of preserved moments for each pair  $(\lambda_i, \delta_j)$  in  $v(\lambda_i, \delta_j)$  and  $w(\lambda_i, \delta_j)$ , respectively.

As a result, the transfer operator of the reduced model coincides in as many moments of the original model as possible for a given order. That is, we have the following theorem.

**Theorem 5.6** Consider the transfer operator  $G(\lambda, \delta) = (\lambda I \star (\Delta(\delta) \star P^\theta))$  is analytic in the neighbourhood of the expansion point  $(\lambda_0, \delta_0)$ . If the matrices  $W, V$  are the basis of the  $\mathcal{W}, \mathcal{V}$  as defined in (5.19) and (5.20), this leads to the transfer operator of the reduced system. Then, the Hermite interpolation condition is satisfied as

$$m_{\{I_i, J_j\}}(\lambda_i, \delta_j) = m_{r, \{I_i, J_j\}}(\lambda_i, \delta_j) \quad (5.22)$$

where, for every  $(i, j)$

$$\begin{aligned} I_i &= 0, \dots, \sigma_{i,j}^v + \sigma_{i,j}^w - 1 \\ J_j &= 0, \dots, \sigma_{i,j}^v + \sigma_{i,j}^w - 1. \end{aligned} \quad (5.23)$$

Here  $m_{r,\{I_i, J_j\}}(\lambda_i, \delta_j)$  denotes the  $\{I_i, J_j\}$ th moment of the reduced model. With the construction of the projection matrices as described above, the reduced system matrices of (5.7) are

$$\begin{bmatrix} A_r(\theta) & B_r(\theta) \\ C_r(\theta) & D_r(\theta) \end{bmatrix} = \begin{bmatrix} A_r(\theta^0) & B_r(\theta^0) \\ C_r(\theta^0) & D_r(\theta^0) \end{bmatrix} + \sum_{i=1}^{\ell} \delta_i \begin{bmatrix} w_i A_{r,i} & w_i B_{r,i} \\ w_i C_{r,i} & w_i D_{r,i} \end{bmatrix} \quad (5.24)$$

with

$$\begin{aligned} A_r(\theta^0) &= W^T A(\theta^0) V & A_{r,i} &= W^T A_i V \\ B_r(\theta^0) &= W^T B(\theta^0) & B_{r,i} &= W^T B_i \\ C_r(\theta^0) &= C(\theta^0) V & C_{r,i} &= C_i V \\ D_r(\theta^0) &= D(\theta^0) & D_{r,i} &= D_i. \end{aligned}$$

By constructing the moments for the reduced order model in (5.24) using the Definition 5.5, the matching property of the Hermit interpolation condition for the full-order model and the reduced-order model can be derived.

In particular, the factorization proposed in (5.8) also can be preserved in the reduced model, that is

$$\begin{aligned} \begin{bmatrix} w_i A_{r,i} & w_i B_{r,i} \\ w_i C_{r,i} & w_i D_{r,i} \end{bmatrix} &= \begin{bmatrix} w_i W^T A_i V & w_i W^T B_i \\ w_i C_i V & w_i D_i \end{bmatrix} \\ &= \begin{bmatrix} W^T L_i^1 R_i^1 V & W^T L_i^1 R_i^2 \\ L_i^2 R_i^1 V & L_i^2 R_i^2 \end{bmatrix} = \begin{bmatrix} W^T L_i^1 \\ L_i^2 \end{bmatrix} \begin{bmatrix} R_i^1 V & R_i^2 \end{bmatrix} \end{aligned} \quad (5.25)$$

Since  $W, V$  are constructed to be full row matrices, and it is readily seen that the  $\begin{bmatrix} W^T L_i^1 \\ L_i^2 \end{bmatrix}$  and  $\begin{bmatrix} R_i^1 V & R_i^2 \end{bmatrix}$  are full column rank and full row rank matrices respectively, if the  $\begin{bmatrix} L_i^1 \\ L_i^2 \end{bmatrix}$  are full column rank and the  $(R_i^1, R_i^2)$  are full row rank. Therefore, the reduced system realization can be represented as

$$\begin{bmatrix} y_{1,\Delta} \\ \vdots \\ y_{\ell,\Delta} \\ \eta_r \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & \cdots & 0 & R_1^1 V & R_1^2 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & R_\ell^1 V & R_\ell^2 \\ \hline W^T L_\ell^1 & \cdots & W^T L_\ell^1 & A_r(\theta^0) & B_r(\theta^0) \\ L_\ell^2 & \cdots & L_\ell^2 & C_r(\theta^0) & D_r(\theta^0) \end{bmatrix}}_{P_r^\theta} \begin{bmatrix} u_{1,\Delta} \\ \vdots \\ u_{\ell,\Delta} \\ x_r \\ u \end{bmatrix} \quad (5.26)$$

with

$$\begin{bmatrix} u_{1,\Delta} \\ \vdots \\ u_{\ell,\Delta} \end{bmatrix} = \begin{bmatrix} \delta_1(\theta_1)I & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \delta_\ell(\theta_\ell)I \end{bmatrix} \begin{bmatrix} y_{1,\Delta} \\ \vdots \\ y_{\ell,\Delta} \end{bmatrix} = \Delta(\delta) \begin{bmatrix} y_{1,\Delta} \\ \vdots \\ y_{\ell,\Delta} \end{bmatrix}.$$

Thus, the error system between the original system  $(\Delta(\lambda), (\Delta(\delta) \star P^\theta))$  and the reduced system  $(\Delta_r(\lambda), (\Delta(\delta) \star P_r^\theta))$  can be formed as

$$(\Delta_e \star (\Delta_e(\delta) \star P_e^\theta)) = [\Delta(\lambda) \star (\Delta(\delta) \star P^\theta) - \Delta_r(\lambda) \star (\Delta(\delta) \star P_r^\theta)] \quad (5.27)$$

where

$$P_e^\theta = \begin{bmatrix} A(\theta) & 0 & B(\theta) \\ 0 & A_r(\theta) & B_r(\theta) \\ C(\theta) & -C(\theta) & D(\theta) - D_r(\theta) \end{bmatrix} \quad \Delta_e(\delta) = \begin{bmatrix} \Delta(\delta) & 0 \\ 0 & \Delta(\delta) \end{bmatrix}$$

and

$$\Delta_e = \begin{bmatrix} \Delta & 0 \\ 0 & \Delta_r \end{bmatrix}$$

At the expansion point of  $\lambda_i, \theta_i$ , we can obtain an error free system, that is

$$(\Delta_e \star (\Delta_e(\delta) \star P_e(\theta)))(\lambda_i, \delta_i) = 0,$$

and the higher order derivatives of the error system at the expansion points are also error free:

$$\begin{aligned} \frac{\partial^i}{\partial \lambda^i} (\Delta_e \star (\Delta_e(\delta) \star P_e(\theta)))(\lambda_i) &= 0 \quad \text{for } i = 1, \dots, I_i - 1 \\ \frac{\partial^j}{\partial \delta^j} (\Delta_e \star (\Delta_e(\delta) \star P_e(\theta)))(\delta_j) &= 0 \quad \text{for } j = 1, \dots, I_j - 1 \end{aligned}$$

where  $I_i, I_j$  are defined in (5.23)

The implementation of the multiple shift operators and multiple parameters moment matching described above in (5.21) is summarized in Algorithm 4.

Different from the method proposed in [29], instead of interpolating the parameters or the neighbour of the parameters, we choose the parametric uncertain variables as the parameter expansion points. In this algorithm, we don't discuss the extension to tangential method that can be easily adapted to the proposed one. Certainly, the performance of this procedure strongly depends on the choices of the interpolation points. The possible refinement of this algorithm is to iteratively select the expansion points as IRKA in [84].

## 5.4 Study cases

In this chapter, we will derive two systems (an LTI in LFT representation and an uncertain LPV) from the Precision Stage Application(PSA), depicted in Figure 5.4.

**Algorithm 4** : Robust moment matching for uncertain LPV systems

- 1: Given:  $\begin{bmatrix} A(\theta) & B(\theta) \\ C(\theta) & D(\theta) \end{bmatrix}$  with  $\theta_i \in [\theta_1, \dots, \theta_\ell]$  and the bound of each parameter  $\theta_i \in [\underline{\theta}_i, \bar{\theta}_i]$
- 2: System realization  $\begin{bmatrix} A(\theta^0) & B(\theta^0) \\ C(\theta^0) & D(\theta^0) \end{bmatrix} + \sum_{i=1}^{\ell} \delta_i \begin{bmatrix} w_i A_i & w_i B_i \\ w_i C_i & w_i D_i \end{bmatrix}$
- 3: Choose the expansion point as a pair  $(\lambda_i, \delta_j)$  where  $\lambda_i \in [\lambda_1, \dots, \lambda_k]$ ,  $\delta_j \in [\delta_1, \dots, \delta_\ell]$  and the corresponding moments  $\begin{bmatrix} \sigma_{1,1}^v, \sigma_{1,2}^v, \dots, \sigma_{k,\ell}^v \\ \sigma_{1,1}^w, \sigma_{1,2}^w, \dots, \sigma_{k,\ell}^w \end{bmatrix}$
- 4: Compute a basis of  $\{v_1, \dots, v_{k \cdot \ell}\}$  for

$$v_1 = \mathcal{K}_{\sigma_{i,j}^v} \{(I - A(\delta_j)\lambda_i)^{-1}, (I - A(\delta_j)\lambda_i)^{-1}B(\theta_j)\}$$

- 5: Compute a basis of  $\{w_1, \dots, w_{k \cdot \ell}\}$  for

$$w_1 = \mathcal{K}_{\sigma_{i,j}^v} \{(I - A(\delta_j)\lambda_i)^{-T}, (I - A(\delta_j)\lambda_i)^{-T}C^T(\theta_j)\}$$

- 6: Set  $V = \text{im}[v_1, \dots, v_{k \cdot \ell}]$  and  $W = \text{im}[w_1, \dots, w_{k \cdot \ell}]$

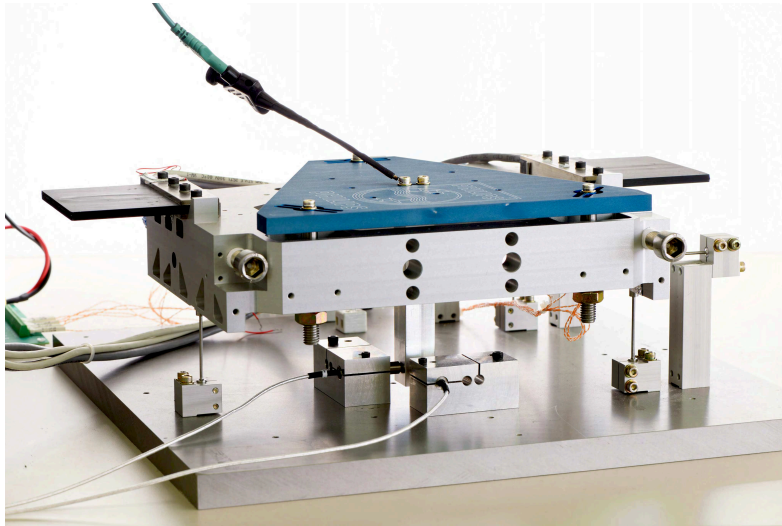
- 7: Construct the reduced model  $\begin{bmatrix} A_r(\theta) & B_r(\theta) \\ C_r(\theta) & D_r(\theta) \end{bmatrix}$  as in (5.24).

Let  $\mathcal{G}$  denote the spatial geometry on which the system is considered. For the purpose of this chapter  $\mathcal{G}$  is a bounded subset of  $\mathbb{R}^3$ . That is, we consider a 3-dimensional configuration space. Let  $\partial\mathcal{G}$  denote its boundary. The governing equation for describing the temporal-spatial thermal behaviour is given by

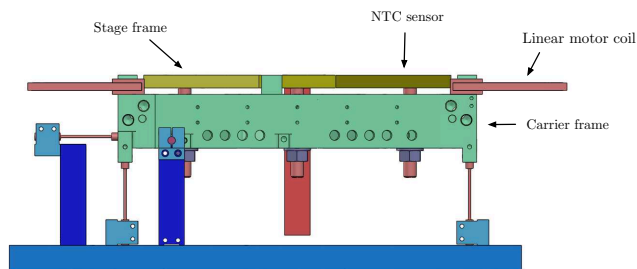
$$\begin{aligned} \rho c_p \frac{\partial T}{\partial t} + \nabla \cdot J &= u \quad \text{in } \mathcal{G} \\ \text{Boundary condition: } \nabla T(r, t) \cdot \bar{n} &= \gamma \quad \text{on } \partial\mathcal{G} \\ \text{Initial condition: } T(r, 0) &= T_0(r) \end{aligned} \tag{5.28}$$

here  $T(r, t)$  is the temperature at location  $r \in \mathcal{G}$  and at time  $t \geq 0$ .  $J = J(r, t) = -\kappa \nabla T(r, t)$ ,  $r \in \mathcal{G}$  is the heat flux.  $\{\rho, c_p, \kappa\} \in \mathbb{R}_{>0}$  represents the density, heat capacity and thermal conductivity of materials in the PSA setup. The  $T_0(r)$  and the  $\gamma$  denotes the initial condition and boundary condition, respectively. In this chapter, the input  $u$  denotes the input power to the linear motor coil, and the temperature measured from the NTC sensor is considered as the output. Notice that we consider the single-input and single-output for the following two cases.

For two different reduction techniques, the numerical results of two types reduced uncertain systems are given. For a detailed explanation of the PSA, we refer to Section 4.5.



(a) Photograph of the PSA setup



(b) Schematic plot of the PSA setup

**Figure 5.4:** Photograph (a) and the schematic plot (b) of the PSA setup. In (b), the stage frame, the carrier frame, the linear motor coil and location of the NTC sensor are indicated.

### 5.4.a Case one: LFR representation of an LTI system

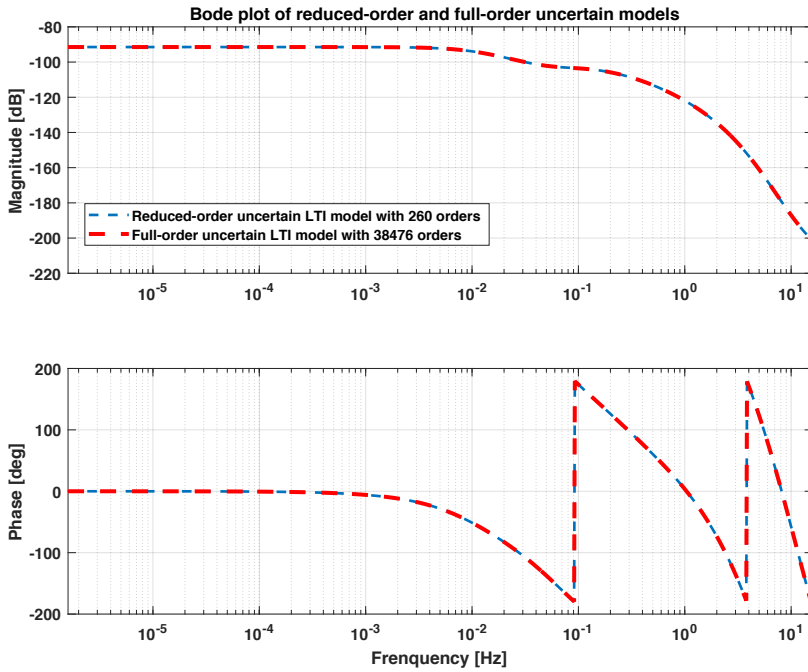
With a proper spatial discretization method, we derive a LTI system based on the governing equation (5.28)

$$\Sigma := \begin{cases} \dot{x}(t) = Ax(t) + bu(t) \\ y(t) = c^T x(t) \end{cases} \quad (5.29)$$

with  $\dim(x) = 38476$ . Here  $u$  denotes one input power to the system and the temperature of the NTC sensor is chosen as the output  $y$ . The reduced order system has been derived from the technique to match expansion points in a set  $\lambda$  of the LFR representation of the system up to appropriate orders. Here, we took

$$\lambda = \{\lambda_1, \lambda_2\} = \{10^{10}, 10\}$$

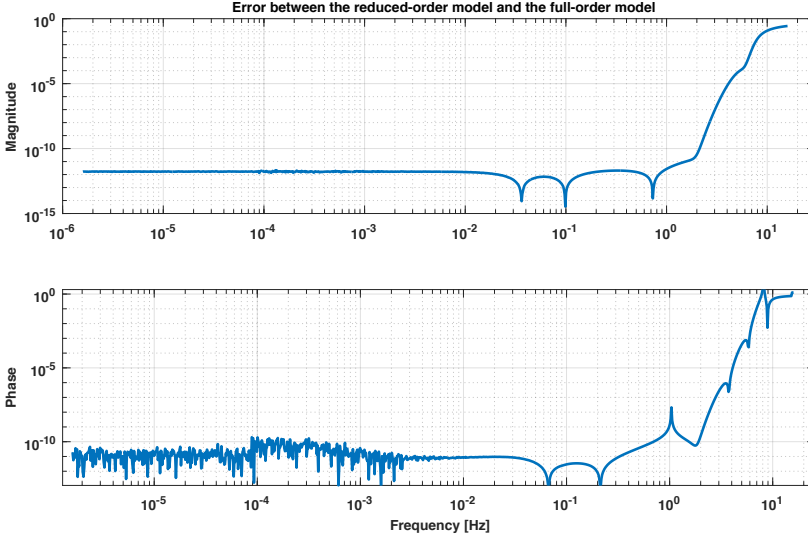
(corresponding to  $s_1 = 10^{-10} \approx 0, s_2 = 0.1$ ) with moment orders 200 and 60, respectively. That would make a reduced order system with a dimension of  $r \geq 260$  order emphasizing matching of the low frequent behavior of the PSA system. In



**Figure 5.5:** Bode plot of the reduced-order uncertain LTI system and the full-order LTI system

Figure 5.5, the frequency response of the reduced-order model with  $r = 260$  orders is close to the full-order model with  $n = 38476$  orders, and the error is plotted in Figure 5.6. Up to frequency 0.8[Hz], the error induced by reduction is approximately below  $10^{-10}$ , and the error is increasing slowly for higher frequencies.

Figure 5.5 shows an accurate fit of the low frequent frequency behavior of the PSA system both in gain as well as in phase. This is due to the choice of expansion points chosen:  $\lambda_0 = [10^{10}, 10]$  which is equal to frequency expansion points at  $s_0 = [10^{-10}, 0.1]$ . We can suppress those high-frequency errors by adding more points at higher frequency points.



**Figure 5.6:** Error between the reduced-order uncertain LTI model and the full-order LTI model:  $(\Delta_e \star P_e) = (\Delta \star P - \Delta_r \star P_r)$ .

### 5.4.b Case two: an LPV system

In this case, the parametric uncertainty of the system is due to the thermal conductivity coefficient of the two materials in the stage frame and the carrier frame  $\theta = [\kappa_1, \kappa_2]$ , and the nominal values are  $\kappa_1^0 = 27, \kappa_2^0 = 44.5$  with variations of  $24.4 \leq \kappa_1 \leq 29.7$  and  $40.05 \leq \kappa_2 \leq 48.95$ . The parameterized state space realization of the system reads

$$\Sigma(\theta) := \begin{cases} \dot{x}(t) = A(\theta)x(t) + bu(t) \\ y(t) = c^T x(t) \end{cases} \quad \text{with } A(\theta) = A_0 + \kappa_1 A_1 + \kappa_2 A_2. \quad (5.30)$$

Here  $\dim(x) = 38476$  and  $\Theta = [24.4, 29.7] \times [40.05, 48.95]$ . Similar to the previous LTI study case,  $u$  denotes the input power and  $y$  represents the temperature of NTC sensor. With the parametric upper bound and the lower bound, the LFT system realization is

$$\Sigma(\delta) := \begin{cases} \dot{x}(t) = A(\delta)x(t) + bu(t) \\ y(t) = c^T x(t) \end{cases} \quad (5.31)$$



here  $A(\delta) = A(\theta^0) + \delta_1 A_1 + \delta_2 A_2$  and  $A(\theta^0) = A_0 + \kappa_1^0 A_1 + \kappa_2^0 A_2$  and  $\delta \in [\delta_1, \delta_2]$ ,  $-1 \leq \delta_i \leq 1$ . This transformation from (5.30) to (5.31) can be obtained via Theorem 5.1.

For generating the reduced-order uncertain LPV systems, the expansion points consist of

$$\begin{aligned} \lambda_0 &\in \{\lambda_0^1, \lambda_0^2, \lambda_0^3\} = \{10^{10}, 10^{10}, 10\} \\ \delta_1 &\in \{\delta_1^1, \delta_1^2, \delta_1^3\} = \{0, -1, 1\} \end{aligned} \quad (5.32)$$

$$\delta_2 \in \{\delta_2^1, \delta_2^2, \delta_2^3\} = \{0, -1, 1\} \quad (5.33)$$

and first set is  $(\lambda_0^1, \delta_1^1, \delta_2^1) = (10^{10}, 0, 0)$  with associated moments  $(200, 50, 50)$ , the second set at  $(10^{10}, 1, 1)$  with order  $(200, 50, 50)$  and the third set  $(10, -1, -1)$  with order  $(100, 20, 20)$ .

In addition, since the parameter expansion points  $\delta_1$  in (5.32) and  $\delta_2$  in (5.33) have the same expansion point  $(0, 1, -1)$ , we can further rewrite the  $A(\delta) = A(\theta^0) + \delta_1 A_1 + \delta_2 A_2$  to  $A(\delta) = A(\theta^0) + \tilde{\delta} \tilde{A}_1$  where  $\tilde{A}_1 = A_1 + A_2$  with  $-1 \leq \tilde{\delta} \leq 1$ , and the expansion points are

$$\begin{aligned} \{\lambda_0^1, \tilde{\delta}^1\} &= \{10^{10}, 0\} \text{ with order of } \{200, 50\} \\ \{\lambda_0^2, \tilde{\delta}^2\} &= \{10^{10}, 1\} \text{ with order of } \{200, 50\} \\ \{\lambda_0^3, \tilde{\delta}^3\} &= \{10, -1\} \text{ with order of } \{100, 20\} \end{aligned}$$

As we see, the  $A(\delta) = A(\theta^0) + \tilde{\delta} \tilde{A}_1$  with  $\tilde{\delta}_0 = [0; 1; -1]$  is equivalent to  $A(\delta) = A(\theta^0) + \delta_1 A_1 + \delta_2 A_2$  with  $[\delta_1, \delta_2]$  at  $[0, 0; 1, 1; -1, -1]$ . And the order of the reduced order model is further reduced from 740 to 620.

In Figure 5.7, Figure 5.9 and Figure 5.11, the frequency response of the full-order uncertain LPV model and two reduced-order uncertain LPV models at the expansion points are given. Both reduced models closely match with the full-order model, respectively. The corresponding error between the full-order and the reduced-order model are shown in Figure 5.8, Figure 5.10 and Figure 5.12. Based on the results, the approximation errors are all below  $10^{-4}$  for both magnitude and phase plots. Notice that we only generate one set of projection matrices  $\{V, W\}$  and reconstruct the two reduced-order models at 3 different uncertain values.

## 5.5 Conclusion

In this chapter, we have presented parametric moment-matching based techniques for reducing the state order while preserving the uncertainty of large-scale uncertain systems. Two type of systems are considered in this chapter: uncertain LTI systems and uncertain LPV systems. Both systems are described by an LFT with a constant nominal plant on a structured uncertainty set, and systems are assumed to be stable for the uncertain set. The fidelity of the reduced uncertain systems is well demonstrated by two real-world models.

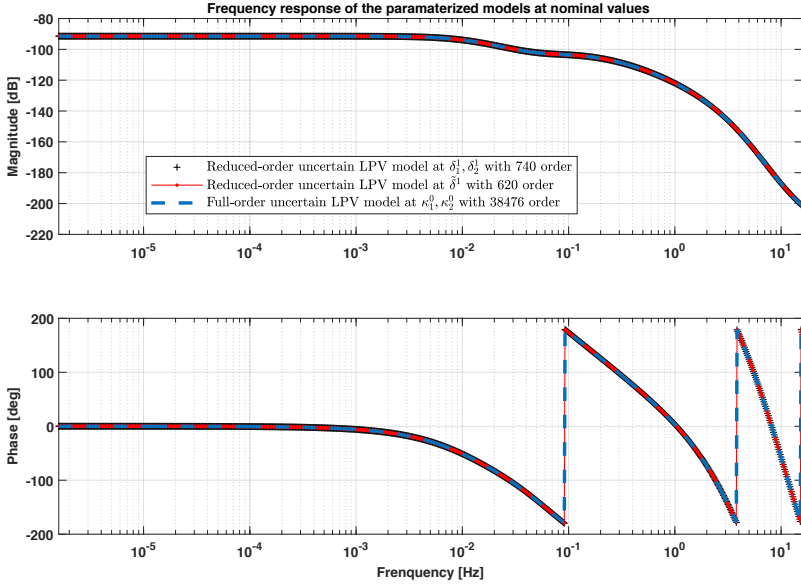


Figure 5.7: Frequency response of three models at nominal values.

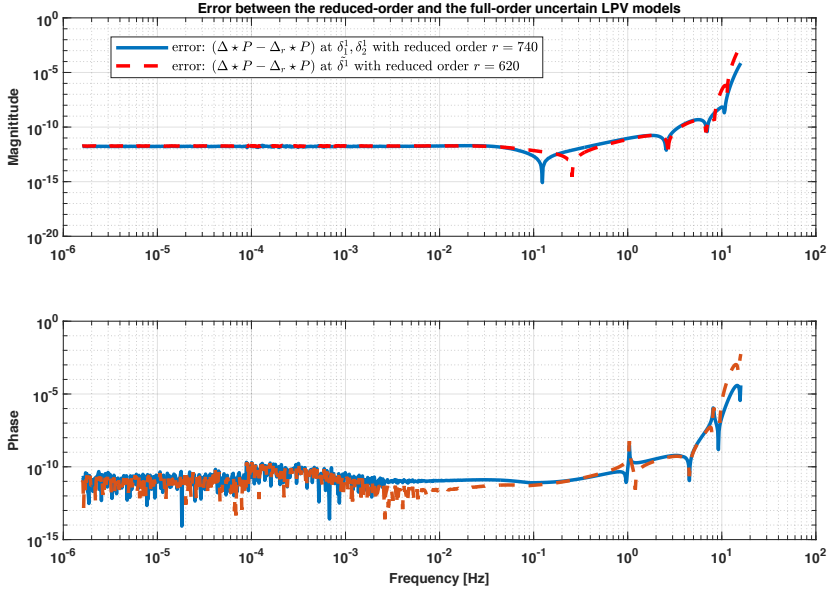


Figure 5.8: Error between the full-order uncertain LPV model and two different reduced-order uncertain LPV models:  $(\Delta_e \star P_e) = (\Delta \star P - \Delta_r \star P_r)$  at  $\{\delta_1^1, \delta_2^1\}$  and  $\tilde{\delta}^1$ , respectively.

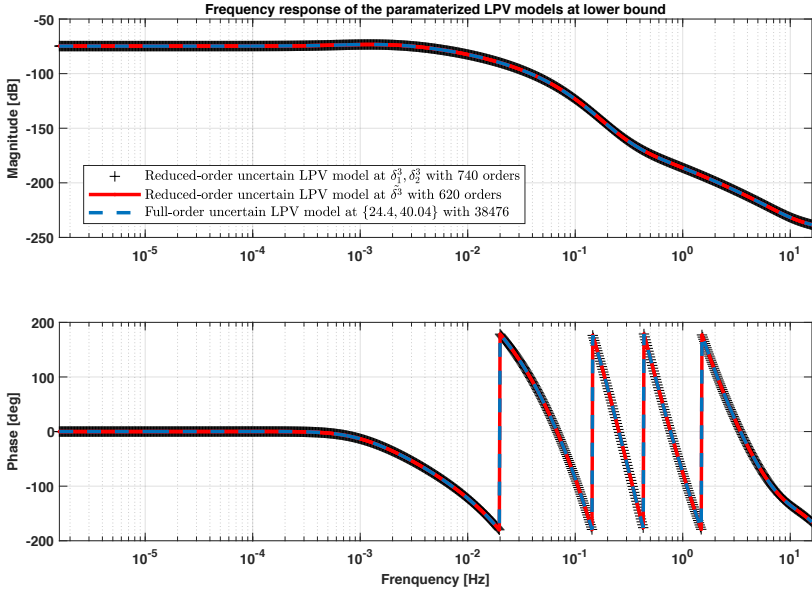


Figure 5.9: Frequency response of three models at lower bound.

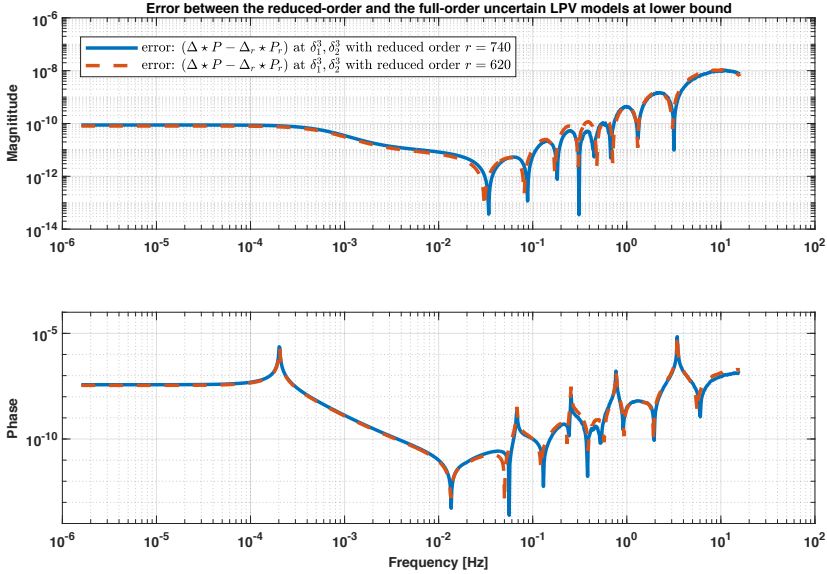


Figure 5.10: Error between the full-order uncertain LPV model and two different reduced-order uncertain LPV models:  $(\Delta_e \star P_e) = (\Delta \star P - \Delta_r \star P_r)$  at  $\{\delta_1^3, \delta_2^3\}$  and  $\hat{\delta}^3$ , respectively.

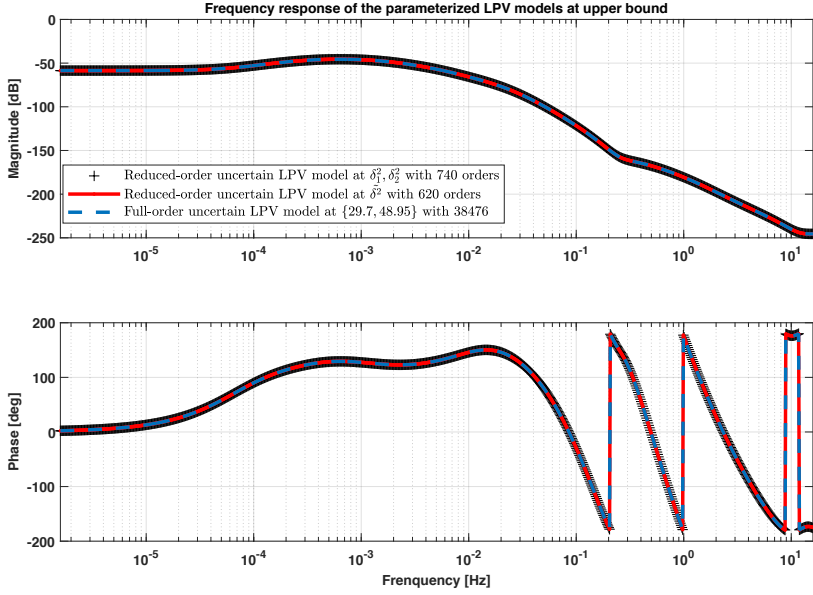


Figure 5.11: Frequency response of three models at upper bound.

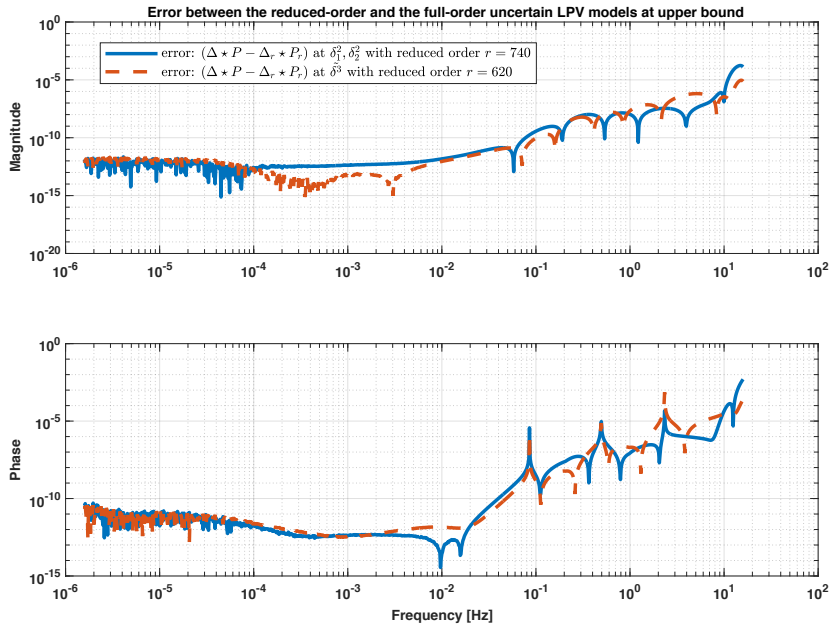


Figure 5.12: Error between the full-order uncertain LPV model and two different reduced-order uncertain LPV models:  $(\Delta_e \star P_e) = (\Delta \star P - \Delta_r \star P_r)$  at  $\{\delta_1^2, \delta_2^2\}$  and  $\tilde{\delta}^2$ , respectively.

## **Part II**

# **Stability and dissipation analysis for thermal systems**



*Thermodynamics, in contrast, is concerned with the macroscopic consequences of the myriads of atomic coordinates that, by virtue of the coarseness of macroscopic observations, do not appear explicitly in a macroscopic description of a system.*

Herbert B. Callen (1919-1993)

# Thermodynamic equilibrium and stability analysis of thermal systems

---

**A**n entropy related function is introduced to serve as a Lyapunov function candidate. Stability of a thermodynamic equilibrium is defined as a point-to-set property and it is shown how this property is verifiable through a suitable Lyapunov function. We show that the proposed Lyapunov function naturally extends to assess the stability and passivity of interconnected thermal systems. Two examples are given to demonstrate the time evolution of the Lyapunov function.

---

The understanding of thermal behaviour plays an increasingly vital role to enhance the performance of high-precision systems. These systems often consist of a large number of components and multi-physics interactions. To characterize such interactions, the coupled balance equations, encoding geometric structures, are used to describe the evolution of different physical quantities in space and time. Although many energy-based methods and tools, e.g., [163, 5], have been developed for modelling and control synthesis of such multi-physics systems, a rigorous integration of the second law of thermodynamics is often lacking. The entropy-based methods [54, 63, 142] are developed using the entropy-balance equation, as these methods permit to include the irreversible processes.

In general, from energy perspective, conservation laws and energy functions are often used to prove stability of equilibria of systems, to analyze their behavior, and for purposes of control and optimization. The role of Lyapunov functions in autonomous systems, together with their generalizations in the direction of open dynamical systems, passive and dissipative dynamical systems, bond-graph theory and port-Hamiltonian systems are fundamental in modern systems theory

[101, 171, 172]. The essence of these methodologies lies in the observation that the efficiency of many physical processes is limited by their energy storage and the amount of energy that has been supplied by its environment. Indeed, Carnot's principle claims the limited efficiency of heat engines; in Hamiltonian and Lagrangian mechanics, the recoverable energy is always bounded by the Hamiltonian or Lagrangian function; in an ideal reversible thermodynamical process, all heat can be converted into work.

In the last decades, non-equilibrium thermodynamics has been developed as a research field, aimed at describing physics of thermal processes beyond and away from thermodynamic equilibrium. In particular, it aims at incorporating the time-course of intensive variables such as temperature and pressure and to generalize the concept of entropy to thermal states that are not in equilibrium. Starting with Onsager's reciprocal relations [126, 127] in 1931, important contributions on non-equilibrium thermodynamics have been made by Denbigh [61] on steady state principles, De Groot's work on linear irreversible processes [55, 54] and Prigogine's minimum entropy production principle [139, 104] that has been extended to the theory of entropy generation minimization by Bejan [27]. This theory claims that the maximum efficiency of a thermodynamic system is achieved while the dissipated energy is minimal.

For many applications, quadratic Lyapunov functions have been widely used to provide insight in the qualitative behavior of systems. Although in [59], a quadratic function of temperature is proposed to satisfy Lyapunov stability criteria, these functions do not naturally inherit the energy attributes of thermodynamic systems, and only point-wise stability can be derived using such Lyapunov functions. In this chapter, we propose an alternative function, directly related to entropy increase, to prove stability properties of thermal systems that are away from their thermodynamic equilibrium.

The main contributions of this work can be summarized as follows: (1) We formally distinguish between the concept of an equilibrium (steady state) of a dynamical system described by differential equations and a thermal equilibrium of a thermodynamic system. We show that not every equilibrium is a thermodynamic equilibrium. In order to avoid confusion in the sequel, we refer the steady state in the sense of the equilibrium in dynamical system theory. (2) We propose a physically relevant Lyapunov function, which characterizes thermal dissipation by entropy generation. It is shown that this function characterizes the Lyapunov stability of any thermodynamic equilibrium for the system we consider. (3) We show the usage of this function for both distributed and lumped thermodynamic models describing thermal conduction.

This chapter is organized as follows. Section 6.1 introduces a number of fundamental concepts of quasi-static and time-dependent thermodynamics. These include entropy balance (second law of thermodynamics), and the definition of thermal equilibria. In Section 6.2, we give the derivation of the physically relevant Lyapunov function and the corresponding stability analysis for both infinite-dimensional and finite-dimensional systems. In Section 6.3, some simulation results are given to validate the equilibrium analysis and the stability analysis based on the proposed Lyapunov function. Conclusions are given in Section 6.4.



We first give some notational conventions that will be used in the chapter. The  $\text{col}$  operator stacks its arguments in a column vector as in  $y = \text{col}(y_1, \dots, y_n)$ .  $\nabla$  is the gradient operator with  $\nabla f := \text{col}(\frac{\partial f}{\partial y_i} \mid i = 1, \dots, n)$  for a differentiable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Its divergence  $\nabla \cdot f = \sum_{i=1}^n \frac{\partial f}{\partial y_i}$  and its Laplacian  $\nabla^2 f = \nabla \cdot (\nabla f) = \sum_{i=1}^n \frac{\partial^2 f}{\partial y_i^2}$ . The directional derivative  $\frac{\partial f}{\partial \bar{n}} = \nabla f \cdot \bar{n}$ .

A real-valued continuously differentiable function  $V : \mathcal{G} \rightarrow \mathbb{R}$  is *positive definite* on a neighborhood  $\mathcal{G}$  of  $x^*$  if  $V(x^*) = 0$  and  $V(x) > 0$  for any  $x \neq x^*$  in  $\mathcal{G}$ . If it satisfies  $V(x) \geq 0$  for any  $x \neq x^*$  in  $\mathcal{G}$ , then  $V$  is said to be *positive semidefinite*. A function  $\alpha : [0, a) \rightarrow \mathbb{R}_+ = [0, \infty)$  is of class  $\mathcal{K}$  if it is continuous, strictly increasing, and  $\alpha(0) = 0$ .

## 6.1 Thermodynamics fundamentals

To clarify the terminology used in thermodynamics, we review some essential definitions that are subjected to analysis. The *system* is referred to the study subject, which includes the collection of matter, the region in space. The *surroundings* is the portion of the region in space outside the system selected for analysis. The systems and the surroundings are separated by the surface called *boundary*. Following the authoritative work of [46], a (macroscopic homogeneous) thermodynamic system is described by *extensive* and *intensive* variables whose behavior is a subset

$$\mathcal{T} \subset \mathcal{X}^{\text{ext}} \times \mathcal{X}^{\text{int}}$$

where the extensive variables

$$x^{\text{ext}} = \text{col}(U, S, V, M, N_1, \dots, N_r) \in \mathcal{X}^{\text{ext}}$$

consist of internal energy  $U$ , entropy  $S$ , volume  $V$ , mass  $M$  and the mole numbers  $N_i$  of the  $r$  constituent chemical components. The intensive variables

$$x^{\text{int}} = \text{col}(T, P, \mu_1, \dots, \mu_r) \in \mathcal{X}^{\text{int}}$$

consist of temperature  $T$ , pressure  $P$  and the electro-chemical potentials  $\mu_i$  of each component. The extensive variables  $x^{\text{ext}}$  are related either by an *energetic fundamental relation*

$$U = U(S, V, N_1, \dots, N_r)$$

or an *entropic fundamental relation*

$$S = S(U, V, N_1, \dots, N_r). \quad (6.1)$$

Either of these relations define the intensive variables  $x^{\text{int}}$  through the partial derivatives

$$\begin{aligned} T &:= \left( \frac{\partial U}{\partial S} \right)_{V, N_1, \dots, N_r}, \quad P := - \left( \frac{\partial U}{\partial V} \right)_{S, N_1, \dots, N_r}, \quad \mu_i := \left( \frac{\partial U}{\partial N_i} \right)_{S, V, \dots, N_k, \dots} \\ \frac{1}{T} &:= \left( \frac{\partial S}{\partial U} \right)_{V, N_1, \dots, N_r}, \quad -\frac{P}{T} := - \left( \frac{\partial S}{\partial V} \right)_{S, N_1, \dots, N_r}, \quad \frac{\mu_i}{T} := \left( \frac{\partial S}{\partial N_i} \right)_{S, V, \dots, N_k, \dots}. \end{aligned} \quad (6.2)$$

See [46, 145]. The first law of thermodynamics claims the preservation of internal energy expressed by Gibbs' equation

$$dU = TdS - PdV + \sum_{i=1}^r \mu_i dN_i. \quad (6.3)$$

where  $TdS$  is the quasi-static flux of heat delivered to the system,  $-PdV$  is mechanical work done on the system and  $\sum_{i=1}^r \mu_i dN_i$  represents quasi-static chemical work done on the system. A similar expansion applies to the entropic representation and reads

$$dS = \frac{1}{T}dU + \frac{P}{T}dV - \sum_{i=1}^r \frac{\mu_i}{T}dN_i. \quad (6.4)$$

### 6.1.a Irreversible thermodynamics

We characterize the rate of the state as the flux, which is defined by

$$J_u \equiv \frac{dU}{dt}, \quad J_v \equiv \frac{dV}{dt}, \quad J_{n_i} \equiv \frac{dN_i}{dt}. \quad (6.5)$$

Here  $J_u, J_v, J_{n_i}$  denote the flux of the internal energy, the volume and the mole numbers, respectively. Under the assumption that  $x^{\text{ext}}$  and  $x^{\text{int}}$  represent a *quasi-static process* consisting of an ordered succession of equilibrium states [46, 107], we may derive the local balance equation for the time evolution of state variables according to [46, p. 309]

$$\frac{dS}{dt} = \frac{\partial S}{\partial U} \frac{dU}{dt} + \frac{\partial S}{\partial V} \frac{dV}{dt} - \sum_{i=1}^r \frac{\partial S}{\partial N_i} \frac{dN_i}{dt}. \quad (6.6)$$

By substituting (6.2)

$$\frac{dS}{dt} = \frac{1}{T} \frac{dU}{dt} + \frac{P}{T} \frac{dV}{dt} - \sum_{i=1}^r \frac{\mu_i}{T} \frac{dN_i}{dt}. \quad (6.7)$$

Define the rate of mechanical work  $\dot{W}_m := -P \frac{dV}{dt}$ , the rate of chemical work  $\dot{W}_c := \sum_{i=1}^r \mu_i \frac{dN_i}{dt}$  and the rate of heat flow  $\dot{Q} := T \frac{dS}{dt}$  (all in Joule per second).

The time evolution of (6.6) then leads to the balance equation

$$\frac{dU}{dt} = \dot{Q} + \dot{W}_c + \dot{W}_m$$

which may be viewed as an extension of the first law of thermodynamics to non-equilibrium states [60]. Notice that the  $\dot{W}_m$ ,  $\dot{W}_c$  and  $\dot{Q}$  are meant symbolically here to denote the rate of quantities, respectively. A wide range of macroscopic systems has been studied in this context. See, e.g., [54, 106, 73, 61, 139, 100].

### 6.1.b Entropy generation

According to the fundamental postulate of thermodynamics [54], the entropy change of a system can be decomposed according to

$$\frac{dS_{sys}}{dt} = \frac{dS_e}{dt} + \frac{dS_i}{dt}, \quad (6.8)$$

where  $S_e$  is the total entropy supplied by its surrounding across the external boundary of the system and  $S_i$  denotes the total entropy production due to processes in the interior of the system. The change in the entropy of a system is associated with the heat flow and the mass flow but not the work flow  $P \frac{dV}{dt}$  [145]. Irreversibility is the difference between reversible work and actual work and leads to a net increase in entropy. In this chapter, the  $S_e$  we consider can be contributed by two parts: the mass flow  $S_{e,m}$  and heat flow  $S_{e,q}$ :

$$S_e := S_{e,m} + S_{e,q} \quad (6.9)$$

Specifically, we have the following entropy balance equation

$$\frac{dS_{sys}}{dt} = \sum_{i=1}^N \dot{m}_i \hat{s}_i + \frac{\dot{Q}}{T} + S_{gen}, \quad (6.10)$$

and the here  $\dot{m}_i$  denotes the net rate of the  $i$ th mass flow  $m_i$  into or out of the system and  $\hat{s}_i = S_{e,m}/m_i$  is the entropy per unit mass.  $\dot{Q}/T$  represents the rate of entropy flow  $S_{e,q}$  due to the heat flow across the boundary and  $S_{gen}$  characterizes the internal entropy generation within the system (6.8). The entropy generation rate is

$$S_{gen} = \frac{dS_i}{dt}. \quad (6.11)$$

It is important to remark that  $S_{gen} = 0$  for reversible processes, while  $S_{gen} > 0$  for irreversible processes.

### 6.1.c Thermodynamic equilibria

In this chapter we distinguish between thermodynamic equilibria and the steady state of autonomous differential equations. In order to avoid confusion in the

sequent,  $(\cdot)^*$  denotes the thermodynamic equilibrium and we refer the  $(\cdot)^*$  to the steady state. A thermodynamic equilibrium is defined as a state  $x = \text{col}(x^{\text{ext}}, x^{\text{int}})$  that (1) does not vary with time; (2) is spatially uniformly distributed, e.g., without temperature or pressure gradients; (3) causes no flow of heat, mass, or work between the system and its surroundings; (4) causes the net rate of all chemical reactions to be zero.

The first requirement demands a state variable at equilibrium  $x^*$  in (6.6) to be time-invariant:

$$x^* \in \mathcal{X}^{\text{ext}} \times \mathcal{X}^{\text{int}} \implies \frac{dx^*}{dt} = 0. \quad (6.12)$$

The second requirement imposes a *consensus constraint* addressed by [125] and [46] which states that the equilibrium is uniformly established at all compartments of a composite system. This phenomenon is known as the *equipartition of energy*, that was elaborated by [32, 115, 133] and requires a uniform spatial distribution of state variables in thermodynamic equilibrium:

$$x^* \in \mathcal{X}^{\text{ext}} \times \mathcal{X}^{\text{int}} \implies \nabla x^* = 0. \quad (6.13)$$

The third condition states that

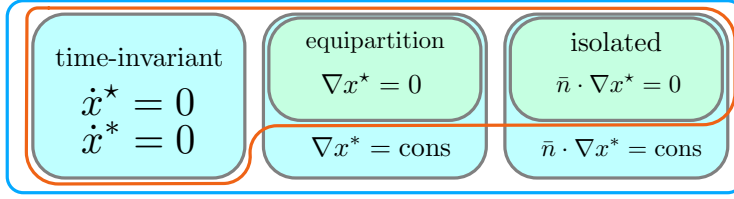
$$x^* \in \mathcal{X}^{\text{ext}} \times \mathcal{X}^{\text{int}} \implies \nabla x^* \cdot \bar{n} = 0 \text{ on } \partial\mathcal{G} \quad (6.14)$$

where  $\partial\mathcal{G}$  denotes the boundary of the geometry of the system and  $\bar{n}$  is an outward pointing unit vector at the boundary. The fourth condition establishes the chemical equilibrium [153]  $\sum_{i=1}^r \mu_i^* v_i^* = 0$  where  $v_i$  denote the stoichiometric coefficients.

**Definition 6.1** For a dynamic system  $\dot{x} = f(x, t)$  defined on a geometric domain  $\mathcal{G}$ , let  $\nabla x \cdot \bar{n} = \gamma$  on its boundary  $\partial\mathcal{G}$ . A point  $x^*$  is in thermodynamic equilibrium (in the absence of chemical reactions) if

1. it is time-invariant:  $\frac{dx^*}{dt} = 0$ ,
2. it satisfies the equipartition property  $\nabla x^* = 0$ ,
3. it is independent of interaction across the boundary  $\nabla x^* \cdot \bar{n} = 0$  on  $\partial\mathcal{G}$ .

**Remark 6.1** For (autonomous) dynamical systems described by ordinary differential equations, the steady state is a constant solution to the differential equation. The steady state in dynamical systems only meets the first requirement in Definition 6.1. In particular, a dynamical system can still be at steady state while the inflow rate is equal to the outflow rate. Hence, a dynamical system can be at steady state without being at thermodynamic equilibrium if (2) or (3) are not satisfied. See Figure 6.1 for an illustration of the various requirements.



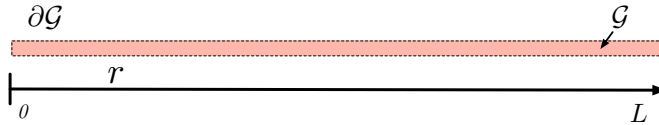
**Figure 6.1:** A set of steady state  $x^*$  (blue frame) and its subset of thermodynamic equilibrium  $x^*$  (red frame).

## 6.2 Stability of thermodynamic systems

In this section, we analyse the stability of thermodynamic equilibria. For an example of thermal conduction in a system described by partial differential equations (PDE), we present a Lyapunov function that warrants stability of a thermodynamic equilibrium. A similar example is given in the second subsection for a system described by ordinary differential equations.

### 6.2.a Equilibria in distributed systems

Consider a model of 1-dimensional thermal conduction in a solid, as depicted in Figure 6.2, whose governing equation, boundary conditions and initial condition are given by



**Figure 6.2:** Heat conduction in a 1-dimensional solid with geometry  $\mathcal{G}$  and boundary  $\partial\mathcal{G}$  (dashed line).

$$\rho c_p \frac{\partial T(r, t)}{\partial t} + \nabla \cdot J(r, t) = 0 \quad \text{in } \mathcal{G}, \quad (6.15a)$$

$$\nabla T(r, t) \cdot \bar{n} = \gamma \quad \text{on } \partial\mathcal{G}, \quad (6.15b)$$

$$T(r, 0) = T_0(r) \quad \text{at initial time } t = 0. \quad (6.15c)$$

Here,  $T : \mathcal{G} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{> 0}$  is an analytic function of the 1-dimensional domain  $\mathcal{G}$  that represents temperature. The heat flux  $J(r, t) := -\kappa \nabla T(r, t)$ ,  $r \in \mathcal{G}$  where  $\kappa > 0$  assumed constant throughout the medium represents the heat transfer coefficient. Note that the notation  $J(r)$  is also used so as to represent  $J(r) = -\kappa \nabla T(r)$  where, in this notation,  $T$  is interpreted as a distribution of temperature with  $T(r)$  the temperature at location  $r$ . Density and heat capacity of the solid are defined by  $\rho$  and  $c_p$ , respectively. The case where  $\gamma = 0$  corresponds to an isolated system.

We assume a time-independent non-positive constant  $\gamma$ . For every time instance  $t \geq 0$ , the temperature  $T(\cdot, t)$  is assumed to be a mapping from  $\mathcal{G}$  to  $\mathbb{R}_{>0}$ . The set of all such mappings is denoted by  $\mathcal{D}$ . With a given initial condition  $T_0 \in \mathcal{D}$  and  $\gamma$  defined on  $\partial\mathcal{G}$ , we will assume that  $T(r, t)$ , with  $r \in \mathcal{G}$ ,  $t \geq 0$  is uniquely defined by (6.15).

Let  $\mathcal{D}$  be equipped with an inner product

$$\langle T_1, T_2 \rangle := \int_{\mathcal{G}} T_1(r) T_2(r) \, dr \quad (6.16)$$

where  $T_1, T_2 \in \mathcal{D}$ . Then  $(\mathcal{D}, \langle \cdot, \cdot \rangle)$  becomes an inner product space with induced norm

$$\|T\| := \sqrt{\langle T, T \rangle}. \quad (6.17)$$

Given the general form (6.8) and the heat conduction equation mentioned above, we can derive the entropy balance equation for (6.15) with  $\partial r = \bar{r}$

$$\frac{d}{dt} S_{sys}(T(r, t)) = - \int_{\partial\mathcal{G}} \frac{J(r)}{T(r)} \cdot \bar{n} \, d\bar{r} + \int_{\mathcal{G}} J(r) \cdot \nabla \frac{1}{T(r)} \, dr. \quad (6.18)$$

The detailed derivation is given in Appendix 6.A. Since  $\gamma$  is assumed constant in (6.15b), the entropy flux across the boundary  $\partial\mathcal{G}$  is constant. The sign of this term is determined by the outward pointing direction of the normal  $\bar{r}$ . The second term denotes the entropy generation within the system, that we will propose as a candidate Lyapunov function. Let  $S_{gen} : \mathcal{D} \rightarrow \mathbb{R}$  be defined as

$$\begin{aligned} S_{gen}(T) &:= \kappa \|\nabla(\ln T)\|^2 \\ &= \int_{\mathcal{G}} J(r) \cdot \nabla \frac{1}{T}(r) \, dr = -\kappa \int_{\mathcal{G}} \nabla T \cdot \nabla \frac{1}{T}(r) \, dr \end{aligned} \quad (6.19)$$

Thus,  $S_{gen}(T)$  is a positive semidefinite function on  $\mathcal{D}$ . The steady state solution  $T^*$  of (6.15a) satisfy  $\kappa \nabla^2 T^* = 0$  and therefore have constant gradient  $\nabla T^*$ . For the case of the constant  $\nabla T^* \neq 0$ , the  $S_{gen}(T^*)$  does not vanish at the steady state  $T^*$ . For the case of the constant  $\nabla T^* = 0$ , the  $S_{gen}$  is equal to zero. Overall, the  $S_{gen}(T)$  is positive semidefinite on a neighborhood of  $T^*$ .

**Theorem 6.2** *Let  $\mathbb{E}$  denote the set of all thermodynamic equilibrium points of the system (6.15). Then any  $T^* \in \mathbb{E}$  is Lyapunov stable. Moreover, the function  $S_{gen} : \mathcal{D} \rightarrow \mathbb{R}$  defined by (6.19) is a Lyapunov function in the sense that  $S_{gen}$  is positive definite on a neighborhood of  $T^* \in \mathcal{D}$  and  $\dot{S}_{gen}(T) \leq 0$  for all  $T \in \mathcal{D}$ .*

**Proof.** Let  $T^* \in \mathbb{E}$  and consider  $S_{gen}$  defined in (6.19). Then  $S_{gen}(T) \geq 0$  and since  $T^*$  satisfies (6.13), it follows from (6.19) that  $S_{gen}(T^*) = 0$ .  $S_{gen}$  is therefore positive definite on a neighborhood of  $T^*$ . For any time varying temperature distribution  $T(r, t)$  we assume that the frozen time instant  $T(\cdot, t)$  belongs to  $\mathcal{D}$ . Then the composite function  $t \mapsto S_{gen}(T(r, t))$  is well defined and its derivative with respect to time is denote  $\dot{S}_{gen}(T)$  in the sequel. Then, applying the chain

rule, we find the time derivative  $\dot{S}_{gen}(T)$  of  $S_{gen}(T(r, t))$  along solutions of (6.15) is given by

$$\begin{aligned} \frac{dS_{gen}(T)}{dt} &:= \dot{S}_{gen}(T) = -\kappa \int_{\mathcal{G}} \left[ \left( \frac{\partial}{\partial t} \nabla T \right) \cdot \nabla \frac{1}{T} + \nabla T \cdot \left( \frac{\partial}{\partial t} \nabla \frac{1}{T} \right) \right] dr \\ &= -\kappa \int_{\mathcal{G}} \left[ \nabla \dot{T} \cdot \nabla \frac{1}{T} + \nabla T \cdot \nabla \left( \frac{\partial}{\partial t} \frac{1}{T} \right) \right] dr. \end{aligned} \quad (6.20)$$

where  $\dot{T}$  also refers to the time derivative  $\frac{d}{dt}T(r, t)$ . Using the Divergence theorem (integration by parts to higher dimensions) it follows that

$$\begin{aligned} \int_{\mathcal{G}} \nabla \dot{T} \cdot \nabla \frac{1}{T} dr &= \int_{\mathcal{G}} \nabla \cdot \left( \dot{T} \nabla \frac{1}{T} \right) dr - \int_{\mathcal{G}} \dot{T} \nabla^2 \frac{1}{T} dr, \\ &= \int_{\partial \mathcal{G}} \dot{T} \nabla \frac{1}{T} \cdot \bar{n} d\bar{r} - \int_{\mathcal{G}} \dot{T} \nabla^2 \frac{1}{T} dr \end{aligned} \quad (6.21)$$

$$\begin{aligned} \int_{\mathcal{G}} \nabla T \cdot \nabla \left( \frac{\partial}{\partial t} \frac{1}{T} \right) dr &= \int_{\mathcal{G}} \nabla \cdot \left( \frac{\partial}{\partial t} \frac{1}{T} \nabla T \right) dr - \int_{\mathcal{G}} \frac{\partial}{\partial t} \frac{1}{T} \nabla^2 T dr, \\ &= \int_{\partial \mathcal{G}} \frac{\partial}{\partial t} \frac{1}{T} \nabla T \cdot \bar{n} d\bar{r} - \int_{\mathcal{G}} \frac{\partial}{\partial t} \frac{1}{T} \nabla^2 T dr \end{aligned} \quad (6.22)$$

Substitute (6.21),(6.22) in (6.20) to infer that

$$\dot{S}_{gen}(T) = -\kappa \int_{\partial \mathcal{G}} \left[ \left( \frac{\partial}{\partial t} \frac{1}{T} \nabla T \right) + \left( \dot{T} \nabla \frac{1}{T} \right) \right] \cdot \bar{n} d\bar{r} + \kappa \int_{\mathcal{G}} \left[ \dot{T} \nabla^2 \frac{1}{T} + \left( \frac{\partial}{\partial t} \frac{1}{T} \right) \nabla^2 T \right] dr. \quad (6.23)$$

In the following, we take a closer look at the first three terms at the right-hand side of (6.23), using the Divergence theorem, it can be derived as

$$\begin{aligned} &-\kappa \int_{\partial \mathcal{G}} \left[ \frac{\partial}{\partial t} \frac{1}{T} \nabla T + \dot{T} \nabla \frac{1}{T} \right] \cdot \bar{n} d\bar{r} + \kappa \int_{\mathcal{G}} \dot{T} \nabla \cdot \left( \nabla \frac{1}{T} \right) dr \\ &= \kappa \int_{\partial \mathcal{G}} \frac{2}{|T|^2} \dot{T} \nabla T \cdot \bar{n} d\bar{r} - \kappa \int_{\mathcal{G}} \dot{T} \nabla \cdot \left( \frac{1}{|T|^2} \nabla T \right) dr \\ &= \kappa \int_{\partial \mathcal{G}} \frac{1}{|T|^2} \dot{T} \nabla T \cdot \bar{n} d\bar{r} \end{aligned} \quad (6.24)$$

which vanish because of the boundary condition (6.15b). Moving from this, (6.23) is therefore equal to

$$\dot{S}_{gen}(T) = \kappa \int_{\mathcal{G}} \left( \frac{\partial}{\partial t} \frac{1}{T} \right) \nabla^2 T dr = -\kappa \int_{\mathcal{G}} \frac{1}{|T|^2} \dot{T} \nabla^2 T dr.$$

By substituting the model equation (6.15a) this gives

$$\dot{S}_{gen}(T) = -\rho c_p \int_{\mathcal{G}} \left| \frac{\dot{T}}{T} \right|^2 dr = -\rho c_p \left\| \frac{d}{dt}(\ln T) \right\|^2 \quad (6.25)$$

or, equivalently,

$$\dot{S}_{gen}(T) = -\frac{\kappa^2}{\rho c_p} \left\| \frac{\nabla^2 T}{T} \right\|^2. \quad (6.26)$$

It follows that  $\dot{S}_{gen}(T) \leq 0$  for all  $T \in \mathcal{D}$  and  $\dot{S}_{gen}(T^*) = 0$ . We can conclude that  $S_{gen}(T)$  is a Lyapunov function proving Lyapunov stability of  $T^* \in \mathbb{E}$ .  $\square$

Theorem 6.2 proves stability of thermodynamic equilibrium points of the model (6.15), not their asymptotic stability. Indeed, the set  $\mathbb{E}$  generally does not consist of isolated equilibrium points, which means that  $\dot{S}_{gen}$  is not negative definite in an open neighborhood of  $T^* \in \mathbb{E}$ . The following result establishes uniform asymptotic stability to the set  $\mathbb{E}$ . For this, let the distance between  $T \in \mathcal{D}$  and the set  $\mathbb{E}$  of thermodynamic equilibria of (6.15) be defined by

$$\text{dist}(T, \mathbb{E}) := \inf_{T^* \in \mathbb{E}} \|T - T^*\|^2. \quad (6.27)$$

**Theorem 6.3** *Let  $\mathbb{E}$  denote the set of all thermodynamic equilibrium points of the system (6.15). Then  $\mathbb{E}$  is uniformly asymptotically stable in the sense that  $\lim_{t \rightarrow \infty} \text{dist}(T, \mathbb{E}) = 0$  for any solution  $T$  of (6.15). In particular, there exist functions  $\alpha_1, \alpha_2, \alpha_3$  of class  $\mathcal{K}$  such that*

$$\begin{aligned} \alpha_1(\text{dist}(T, \mathbb{E})) &\leq S_{gen}(T) \leq \alpha_2(\text{dist}(T, \mathbb{E})) \\ \dot{S}_{gen}(T) &\leq -\alpha_3(\text{dist}(T, \mathbb{E})) \end{aligned}$$

for all  $T \in \mathcal{D}$ .

**Proof:** The proof follows a similar reasoning as in [Theorem 4.1] [101]. The difference is that the upper and lower bound on  $S_{gen}$  are measures of  $\text{dist}(T, \mathbb{E})$ . We claim that the distance from  $T \in \mathcal{D}$  to the set  $\mathbb{E}$  of thermodynamic equilibria is  $\text{dist}(T, \mathbb{E}) = \|\nabla T\|^2$ . For  $\infty > T \geq 1$  and  $\kappa > 0$  we have

$$S_{gen}(T) = \kappa \left\| \frac{\nabla T}{T} \right\|^2 \leq \kappa \|\nabla T\|^2 \|T^{-1}\|^2 \leq \kappa \bar{\nu}^2 \text{dist}(T, \mathbb{E}) = \alpha_2(\text{dist}(T, \mathbb{E}))$$

where  $0 < \sup_{\infty > T \geq 1} \|T^{-1}\| = \bar{\nu} \leq 1$ . Similarly, for the lower bound, we have

$$S_{gen}(T) = \kappa \left\| \frac{\nabla T}{T} \right\|^2 \geq \kappa \|\nabla T \underline{\nu}\|^2 = \kappa \underline{\nu}^2 \|\nabla T\|^2 = \alpha_1(\text{dist}(T, \mathbb{E}))$$

where  $0 < \inf_{\infty > T \geq 1} \frac{1}{T} = \underline{\nu} \leq 1$ . By construction,  $\alpha_1(\text{dist}(T^*, \mathbb{E})) = \alpha_1(0) = 0$  and



$\alpha_2(\text{dist}(T^*, \mathbb{E})) = \alpha_2(0) = 0$ . Next, we claim for all real positive  $\rho C_p \in \mathbb{R}_{>0}$ ,

$$\begin{aligned} \dot{S}_{gen}(T) &= -\frac{\kappa^2}{\rho C_p} \left\| \frac{\nabla^2 T}{T} \right\|^2 \leq -\frac{\kappa^2}{\rho C_p} \|\nabla^2 T \underline{\nu}\|^2 \\ &= -\frac{\underline{\nu}^2 \kappa^2}{\rho C_p} \|\nabla^2 T\|^2 \\ &\leq -\frac{\underline{\nu}^2 \kappa^2 \lambda}{\rho C_p} \|\nabla T\|^2 \end{aligned} \quad (6.28)$$

We prove this claim as follows. Using integration by parts,  $\|\nabla T\|^2$  infers

$$\begin{aligned} \int_{\mathcal{G}} \nabla T \nabla T \, dr &= - \int_{\mathcal{G}} T \nabla^2 T \, dr + \int_{\mathcal{G}} \nabla \cdot (T \nabla T) \, dr \\ &= - \int_{\mathcal{G}} T \nabla^2 T \, dr + \int_{\partial \mathcal{G}} T \nabla T \cdot \bar{n} \, dr \\ &\leq - \int_{\mathcal{G}} T \nabla^2 T \, dr + \gamma \max\{T(0), T(L)\} \\ &\leq \int_{\mathcal{G}} -\nabla^2 T \, T \, dr. \end{aligned} \quad (6.29)$$

Notice that the  $\gamma \max\{T(0), T(L)\} \leq 0$  is derived by the divergence theorem and the  $\max\{T(0), T(L)\}$  denotes the highest boundary temperature. Using Holder's inequality, (6.29) can be rewritten as

$$\int_{\mathcal{G}} \nabla T \nabla T \, dr \leq \int_{\mathcal{G}} -\nabla^2 T \, T \, dr \leq \left( \int_{\mathcal{G}} |-\nabla^2 T|^2 \, dr \right)^{\frac{1}{2}} \left( \int_{\mathcal{G}} |T|^2 \, dr \right)^{\frac{1}{2}}.$$

We note that for the given boundary condition (6.15b), there exist such an eigenvalue

$$\lambda = \min_{T \in \mathcal{D}, T \neq 0} \frac{\int_{\mathcal{G}} |\nabla T|^2 \, dr}{\int_{\mathcal{G}} |T|^2 \, dr} \geq 0.$$

(The proof is given in [65] and the references therein). Consequently, we have

$$\|\nabla T\|^2 \leq \lambda^{-1} \|\nabla^2 T\|^2. \quad (6.30)$$

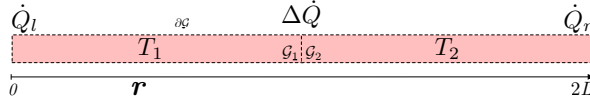
By substituting the (6.30) into (6.28), we prove the claim

$$\dot{S}_{gen}(T) \leq -\frac{\underline{\nu}^2 \kappa^2 \lambda}{\rho C_p} \|\nabla T\|^2 = -\alpha_3(\text{dist}(T, \mathbb{E})).$$

Therefore,  $\dot{S}_{gen}(T) \leq -\alpha_3(\text{dist}(T, \mathbb{E}))$  for a function  $\alpha_3$  of class  $\mathcal{K}$  is proved.  $\square$

### 6.2.b Equilibria in lumped composite systems

In this subsection we give a stability result for thermodynamic equilibria in systems described by ordinary differential equations. For this, consider a spatial partitioning of the geometry  $\mathcal{G}$  depicted in Figure 6.2, leading to the composite system shown in Figure 6.3 with two disjoint geometries  $\mathcal{G}_1$  and  $\mathcal{G}_2$ . In this formulation,



**Figure 6.3:** Composite system of two compartments  $\mathcal{G} := \mathcal{G}_1 \cup \mathcal{G}_2$ . The  $\dot{Q}_l$  and  $\dot{Q}_r$  denote the heat energy interaction at the left side and right side, respectively.

the system dynamics of the two compartments are described by the ordinary differential equations

$$\rho_1 c_{p1} \dot{T}_1(t) = \dot{Q}_l - \Delta \dot{Q} \quad \text{in } \mathcal{G}_1, \quad (6.31a)$$

$$\rho_2 c_{p2} \dot{T}_2(t) = \Delta \dot{Q} - \dot{Q}_r \quad \text{in } \mathcal{G}_2 \quad (6.31b)$$

with initial conditions

$$T_1(0) = T_{0,1}, \quad T_2(0) = T_{0,2}. \quad (6.32)$$

Here  $\Delta \dot{Q}$  denotes the rate of heat flow between  $\mathcal{G}_1$  and  $\mathcal{G}_2$ , represented by Fourier's law  $\Delta \dot{Q} = \kappa(T_1 - T_2)$  where  $\kappa > 0$ . The density and the specific heat capacity of each subsystem are given by  $\rho_i$  and  $c_{pi}$ ,  $i = 1, 2$ , respectively. Notice that the  $\dot{Q}_l$  in (6.31a) and the  $\dot{Q}_r$  in (6.31b) are considered as the heat inflow of  $\mathcal{G}_1$  and the heat outflow of  $\mathcal{G}_2$ .

From the governing equations, the entropy balance (6.8) can be derived by dividing (6.31) by  $T_1$  and  $T_2$ . This gives

$$\rho_1 c_{p1} \frac{\dot{T}_1}{T_1} = \frac{\dot{Q}_l}{T_1} - \frac{\Delta \dot{Q}}{T_1}, \quad \rho_2 c_{p2} \frac{\dot{T}_2}{T_2} = \frac{\Delta \dot{Q}}{T_2} - \frac{\dot{Q}_r}{T_2}. \quad (6.33)$$

The left-hand sides in (6.33) can be rewritten as

$$\rho_i c_{pi} \frac{\dot{T}_i}{T_i} = \rho_i c_{pi} \frac{d}{dt} [\ln T_i(t)].$$

This logarithmic form, also derived in [90], shares the form of Boltzmann's entropy formula. For the composite system, the first law of thermodynamics is obtained by adding the equations (6.31) which gives

$$\rho_1 c_{p1} \dot{T}_1 + \rho_2 c_{p2} \dot{T}_2 = \dot{Q}_l - \dot{Q}_r. \quad (6.34)$$

Similarly, the entropy balance is inferred by adding the equations (6.33),

$$\frac{dS_{sys}}{dt} = \frac{d}{dt} \left[ \ln(\rho_1 c_{p1} \dot{T}_1) + \ln(\rho_2 c_{p2} \dot{T}_2) \right] = \frac{\dot{Q}_l}{T_1} - \frac{\dot{Q}_r}{T_2} + S_{gen}(T_1, T_2) \quad (6.35)$$

where the entropy production is

$$S_{gen}(T_1, T_2) = \kappa \frac{(T_1 - T_2)^2}{T_1 T_2}. \quad (6.36)$$

In this derivation, we set  $\dot{Q}_l = \dot{Q}_r = 0$  as an input-free system. Notice that  $S_{gen}(T_1, T_2) > 0$  as long as  $T_1 \neq T_2$  and  $S_{gen}(T_1, T_2) = 0$  if and only if  $T_1 = T_2$ . Taking the time derivative of  $S_{gen}$  along solutions  $(T_1, T_2)$  of (6.31) gives

$$\dot{S}_{gen}(T_1, T_2) = -\kappa \frac{(T_1^2 - T_2^2)(T_1 \dot{T}_2 - T_2 \dot{T}_1)}{T_1^2 T_2^2} \quad (6.37)$$

Substitute (6.31) to infer that

$$\begin{aligned} \dot{S}_{gen}(T_1, T_2) = & -\frac{1}{2} \frac{T_1 + T_2}{T_1^2 T_2^2} \left( \dot{T}_1^2 \left( \rho_1 c_{p1} T_2 + \frac{\rho_1 c_{p1}}{2 \rho_2 c_{p2}} \right) \right. \\ & \left. + \dot{T}_2^2 \left( \rho_2 c_{p2} T_1 + \frac{\rho_2 c_{p2}}{2 \rho_1 c_{p1}} \right) \right) \end{aligned} \quad (6.38)$$

or, equivalently,

$$\dot{S}_{gen}(T_1, T_2) = -\kappa^2 \frac{(T_1 - T_2)^2 (T_1 + T_2)}{T_1^2 T_2^2} \left( \frac{T_1}{\rho_2 c_{p2}} + \frac{T_2}{\rho_1 c_{p1}} \right). \quad (6.39)$$

This proves that  $S_{gen}$  is a Lyapunov function for the thermodynamic equilibria of the composite system.

**Theorem 6.4** *The set  $\mathbb{E} = \{(T_1, T_2) \mid T_1 = T_2\}$  consist of the thermodynamic equilibrium points of the system (6.31). Any  $(T_1^*, T_2^*) \in \mathbb{E}$  is Lyapunov stable. Moreover, the function  $S_{gen}$  defined in (6.36) is a Lyapunov function.*

Next, we derive the equivalent of Theorem 6.3 for the composite system. Let

$$\text{dist}((T_1, T_2), \mathbb{E}) := \inf_{(T_1^*, T_2^*) \in \mathbb{E}} \omega_1 (T_1 - T_1^*)^2 + \omega_2 (T_2 - T_2^*)^2. \quad (6.40)$$

where  $\omega_i > 0, i = 1, 2$  is positive weight. The weighting factors  $\omega_i$  can be made material dependent and, for example, be chosen as  $\omega_i = \rho_i c_{pi}$ .

**Theorem 6.5** *Let  $\mathbb{E} := \{(T_1^*, T_2^*) \mid T_1^* = T_2^*\}$  be the set of all thermodynamic equilibrium points of (6.31). Then  $\mathbb{E}$  is uniformly asymptotically stable in the sense that  $\lim_{t \rightarrow \infty} \text{dist}(T_i(t), \mathbb{E}) = 0$  for any solution  $T_i(t)$  of (6.31). In particular, there exist*

functions  $\alpha_1, \alpha_2, \alpha_3$  of class  $\mathcal{K}$  such that

$$\begin{aligned}\alpha_1(\text{dist}((T_1, T_2), \mathcal{E})) &\leq S_{gen}(T_1, T_2) \leq \alpha_2(\text{dist}((T_1, T_2), \mathcal{E})) \\ \dot{S}_{gen}(T_1, T_2) &\leq -\alpha_3(\text{dist}((T_1, T_2), \mathcal{E}))\end{aligned}$$

for all  $T \in \mathcal{D}$ .

**Proof:** The proof follows the same lines as the proof of Theorem 6.3.

First, the thermodynamic equilibria with respect to the shortest distance can be obtained by solving the following optimization problem

$$\begin{aligned}\min_{T_1^*, T_2^* \in \mathbb{E}} \quad & \omega_1(T_1 - T_1^*)^2 + \omega_2(T_2 - T_2^*)^2 \\ \text{s.t.} \quad & \mathbb{E} := \{T^* | T_1^* = T_2^*\}, \quad T \in \mathcal{D}\end{aligned}$$

And the local thermodynamic equilibria with respect to the shortest distance is

$$\text{dist}(T, \mathbb{E}) = \frac{\omega_1 \omega_2}{\omega_1 + \omega_2} (T_1 - T_2)^2.$$

Then, the upper bound of the  $S_{gen}(T_1, T_2)$  is

$$S_{gen}(T_1, T_2) = \kappa \|T_1 - T_2\|^2 (T_1 T_2)^{-1} \leq \bar{\varphi} \kappa \frac{\omega_1 + \omega_2}{\omega_1 \omega_2} \text{dist}(T, \mathbb{E}) = \alpha_2(\text{dist}(T, \mathbb{E}))$$

where  $0 < \sup_{\infty > T_1, T_2 \geq 1} (T_1 T_2)^{-1} = \bar{\varphi} \leq 1$ . Similarly, the lower bound is

$$S_{gen}(T_1, T_2) \geq \underline{\varphi} \kappa \|T_1 - T_2\|^2 = \underline{\varphi} \kappa \frac{\omega_1 + \omega_2}{\omega_1 \omega_2} \text{dist}(T, \mathbb{E}) = \alpha_1(\text{dist}(T, \mathbb{E}))$$

where  $0 < \inf_{\infty > T_1, T_2 \geq 1} (T_1 T_2)^{-1} = \underline{\varphi} \leq 1$ . Under such construction,  $\alpha_1(\text{dist}(T^*, \mathbb{E})) = \alpha_1(0) = 0$  and  $\alpha_2(\text{dist}(T^*, \mathbb{E})) = \alpha_2(0) = 0$ . It is readily shown that there exists such a function  $\alpha_3$  of class  $\mathcal{K}$

$$\dot{S}_{gen} \leq -\psi \kappa^2 (T_1 - T_2)^2 = \alpha_3(\text{dist}(T, \mathbb{E}))$$

where  $0 < \psi = \inf_{1 \leq T_1, T_2 < \infty} (T_1^{-1} + T_2^{-2})((T_1 \rho_2 c_{p_2})^{-1} + T_2 \rho_1 c_{p_1})$ . Thus, the proof is completed.  $\square$

The proposed Lyapunov function is the irreversible thermal energy associated with the temperature gradient between any two nodes. One indication is that temperature difference inside the system boundary, induced by thermal energy transfer, results in increasing irreversible entropy that is considered as thermal energy loss. Furthermore, the construction of such Lyapunov functions is independent of the complexity of the thermal system. The extension to  $n$ -compartment system will be covered in the further research.

## 6.3 Numerical example

In this section we provide a computation of the temperature distribution in the 1D heat conduction example given in Figure 6.2. Consider the spatial geometry  $\mathcal{G} = [0, L]$  with  $L > 0$  the length of a beam. Let its temperature distribution be described by (6.15a) with Dirichlet/Neumann boundary condition and initial conditions given by

$$\text{Dirichlet \& IC: } \begin{cases} T(0, t) = T_0(t) \\ T(L, t) = T_L(t) \\ T(r, 0) = \beta \sin(\frac{\pi}{L}r) + T_{0L}(r) \end{cases} \quad (6.41)$$

$$\text{Neumann \& IC: } \begin{cases} \frac{\partial T(0, t)}{\partial r} = 0 \\ \frac{\partial T(L, t)}{\partial r} = 0 \\ T(r, 0) = \cos(\frac{\pi r}{L}) + \xi \end{cases} \quad (6.42)$$

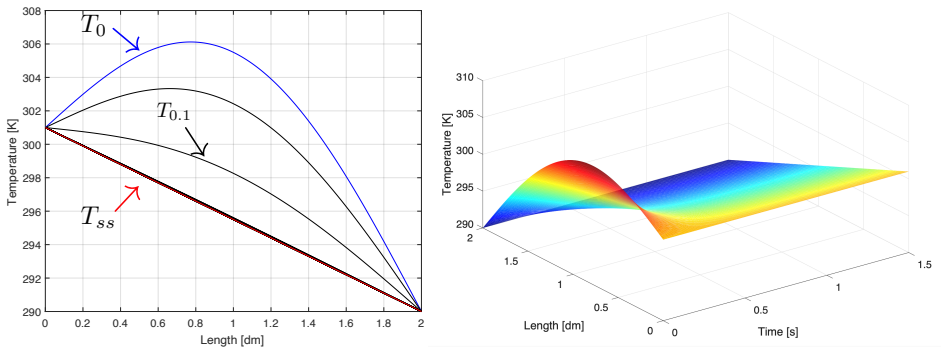
where  $T_{0L}(r) = T_0 + r \frac{T_L - T_0}{L}$  and  $\xi > 0$ . The solutions are obtained by separation of variables and read for the Dirichlet boundary condition

$$T(r, t) = T_{0L}(r) + \beta \sin(\frac{\pi r}{L}) \exp(-\frac{\pi^2}{L^2} \frac{\kappa}{\rho c_p} t)$$

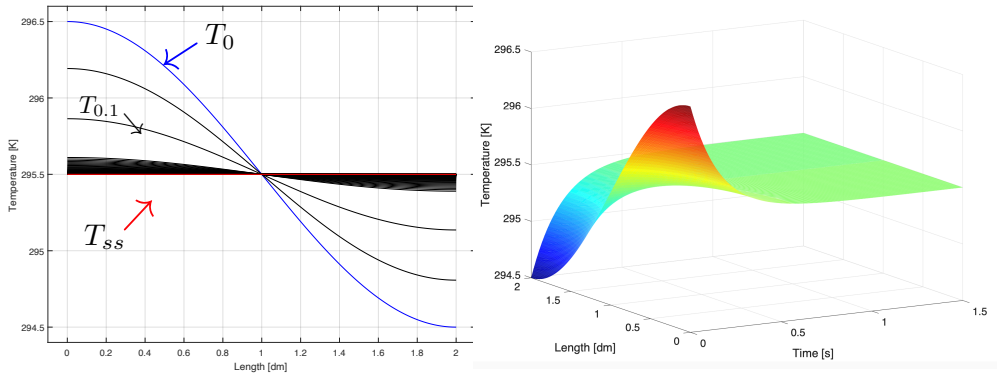
with  $\beta > 0$ , and for the Neumann boundary condition

$$T(r, t) = \xi + \cos(\frac{\pi r}{L}) \exp(-\frac{\pi^2}{L^2} \frac{\kappa}{\rho c_p} t).$$

For  $t \rightarrow \infty$  the two types of boundary conditions exhibit two different stable thermodynamic equilibria as depicted in Figure 6.4 and Figure 6.5.

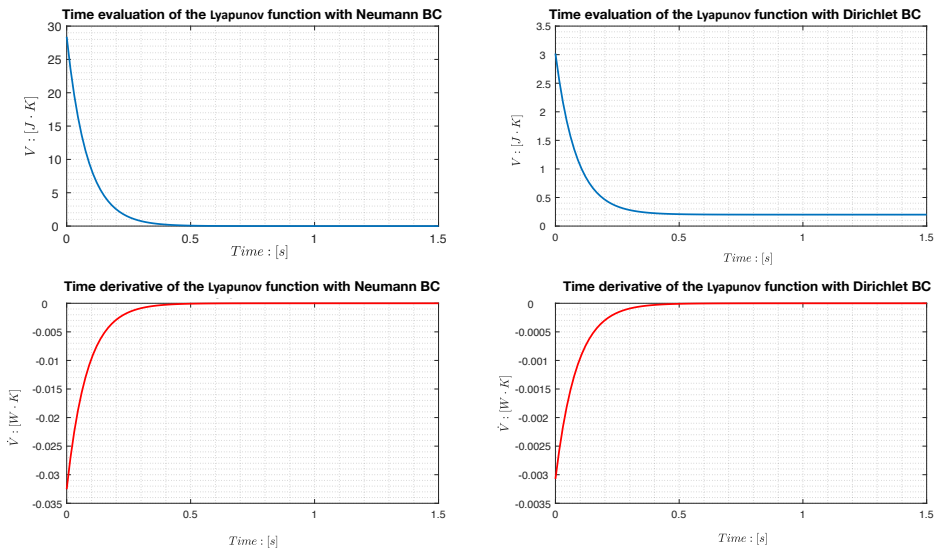


**Figure 6.4:** Spatial temperature profiles with Dirichlet boundary condition. The initial distribution  $t = 0$ , distribution at  $t = 0.1$  and the steady state solution are indicated by  $T_0$ ,  $T_{0.1}$  and  $T_{ss}$ , respectively. The equilibrium in this plot is not a thermodynamic equilibrium since the equipartition condition does not hold.



**Figure 6.5:** Spatial temperature profiles with Neumann boundary condition. The same conditions as in Figure 6.4. The steady state in this figure is a thermodynamic equilibrium.

Besides, the time evolution of the entropy generation function (the proposed Lyapunov function), together with the time derivatives plots for these two boundary conditions (6.41) and (6.42) are given in the following figures.



**Figure 6.6:** Time evolution of the Lyapunov function with two different boundary conditions.

## 6.4 Conclusion

This chapter considers thermodynamical systems beyond their quasi-static behavior. We make an important conceptual distinction between thermodynamic equilibria of a system and equilibria in the sense of constant (time-positive invariant) solutions of autonomous systems described by differential equations. We proved the stability of thermodynamic equilibria in a distributed thermodynamical model and in a composite lumped thermodynamical model. In both cases, a suitable Lyapunov function has been derived from the first principle properties of the entropy balance equation of the system. It is shown that this function represents entropy generation for irreversible thermal processes and decays along the solutions of the system towards thermodynamic equilibria. The set of thermodynamic equilibria is proven to be uniform asymptotically stable.

## 6.A Derivation of entropy balance equation

We present the detailed derivation of the entropy balance equation (6.18). Start with partial differential equation of (6.15a)

$$\rho c_p \frac{\partial T(r, t)}{\partial t} = -\nabla \cdot J(r, t) \quad \text{in } \mathcal{G},$$

The Divergence theorem yields

$$\int_{\mathcal{G}} \rho c_p \frac{\partial T(r, t)}{\partial t} dr = \int_{\mathcal{G}} -\nabla \cdot J(r, t) dr.$$

By considering the internal energy  $U$ , the above equation can be rewritten as

$$\frac{dU}{dt} = \int_{\mathcal{G}} \frac{\partial u(r, t)}{\partial t} dr, \quad \frac{\partial u(r, t)}{\partial t} = \rho c_p \frac{\partial T(r, t)}{\partial t} \quad (6.43)$$

here  $u(r, t)$  denotes the local internal energy at location  $r$  and time  $t$ . Using the local form [103] of the Gibbs' relation in (6.6) and the relation in (6.43), we have

$$\frac{\partial u(r, t)}{\partial t} = T(r) \frac{\partial s(r, t)}{\partial t} \iff \frac{\partial s(r, t)}{\partial t} = \frac{1}{T(r)} \frac{\partial u(r, t)}{\partial t} = \frac{1}{T(r)} \rho c_p \frac{\partial T(r, t)}{\partial t}$$

here the  $s$  represents the local entropy of the system. By substituting  $\rho c_p \frac{\partial T}{\partial t}$  with  $\nabla \cdot J$  in (6.15a), it reads

$$\frac{\partial s(r, t)}{\partial t} = -\frac{1}{T(r)} \nabla \cdot J(r) \quad \text{in } \mathcal{G}.$$

Using the Divergence theorem, the entropy of the 1-D solid system in Figure 6.2 is

$$\begin{aligned}
 \frac{dS_{sys}(r, t)}{dt} &= \int_{\mathcal{G}} \frac{\partial s(r, t)}{\partial t} dr = - \int_{\mathcal{G}} \frac{1}{T(r)} \nabla \cdot J(r) dr \\
 &= - \int_{\mathcal{G}} \nabla \left( \frac{J(r)}{T(r)} \right) dr + \int_{\mathcal{G}} J(r) \cdot \nabla \frac{1}{T(r)} dr \\
 &= - \int_{\partial \mathcal{G}} \frac{J(r)}{T(r)} \cdot \bar{n} d\bar{r} + \int_{\mathcal{G}} J(r) \cdot \nabla \frac{1}{T(r)} dr,
 \end{aligned}$$

here  $\bar{r}$  refers the location over the  $\partial \mathcal{G}$ . Thus, the (6.18) is derived.



# Dissipative properties of thermal systems

---

**T**his chapter embeds thermal systems in the classical theory of dissipative dynamical systems. This theory views dissipation as a generic concept in which supply functions represent power and power exchange among system components and it treats storage functions as the generalized notion of energy. We introduce a storage function that shows that a class of thermal systems is dissipative for both time-independent and time-dependent boundary conditions. Based on the entropy generation minimization theorem, we extend the storage function formulation to systems where thermodynamic work is considered. Using the maximum and the minimum work efficiency, we define a minimal entropy generation and a maximal entropy generation as the upper bound and the lower bound for this storage function. For a 2D heat conduction case study with a time-varying boundary condition, we present the time profile of both the storage function and supply function. In accordance with the dissipation property, the results show that the rate of supply delivered to the system always exceeds the change of internal storage.

---

## 7.1 Introduction

In the development of high-precision systems, the thermal effects are identified as one of the dominant factors that determines the system performance [41, 151]. With the increasing accuracy requirements, high-precision systems become more interconnected and sophisticated. This trend implies that models need to be created where more compartments and/or subsystems need to be interconnected to assess the dynamical properties of a complex system. Consequently, thermal

effects are becoming even more influential. This is even more so when such interactions are coupled not only in their steady-state but also in their transient dynamics. One way to characterize these large-scale interconnected systems, which include interconnections with its surroundings is by using dissipativity theory. The main idea behind the notion of a dissipative system is to establish a relationship between the energy storage in the system and the supply rate via the interaction. With this relationship, an energy-based description of the input-output description for physical systems can be characterized. Additionally, this theory provides the possibility to view a system with large-scale interconnections as an aggregation of modules, each of which is dissipative. The dissipativity theory has witnessed remarkable progress in system analysis and control design for different physical systems, e.g., mechanical systems [122, 162, 118], electrical systems [117, 128], and chemical systems [108, 175, 19]. However, the increasing complexity and accuracy demands of thermal systems still pose challenges to the analysis and the management of large-scale and high-precision systems.

Dissipative systems were introduced by Willems in [171, 172]. Since their introduction, different researchers have addressed the problem to cast thermodynamical systems in the context of dissipative systems. The critical question is the choice of *supply function* and the choice of the *storage function* for thermal systems, which is central for dissipativity theory. In [172], Willems proposed a candidate that is the system entropy with negative sign  $-S$ . Although it satisfies the dissipation inequality, this candidate loses its physical interpretation as an energy-based notion for thermodynamics. Additionally, as pointed in [5], *entropy is not bounded from above* and therefore the choice  $-S$  may not provide a lower bound as an energy function. An alternative proposal is to consider the irreversible entropy production as the *storage function*. This matches the work of Onsager's reciprocal relation [126, 127], the minimal entropy generation principle introduced by Prigogine [139] and the work of irreversible thermodynamics from de Groot and Mazur [54]. In this context, the proposed irreversible entropy rate, also named *entropy production*, has been proved be a suitable candidate to prove the stability of a closed system. In addition to the concept of irreversible entropy rate, the developments have been extended to open systems where the passivity properties of a plasma reaction model [76] and multi-physical systems [165] are discussed. It is worth mentioning that the use of irreversible entropy production as the storage function is different from the conventional way for constructing the storage function: different from finding a *preserved* energy representation of the system, e.g., kinetic energy and potential energy, the irreversible entropy production describes the energy which crosses the system's boundary is dissipated as a loss. Another approach is to use thermodynamic availability as the *storage function* [1, 4, 175]. In this way, the irreversible entropy rate is implicitly considered in the formulation, and the proposed availability has a lower bound using thermodynamic potential [5].

Other work connects the dissipative systems theory with thermodynamics on top of the methods mentioned above. An approach in [90] is studied based on constructing an *ectropy* as storage function for stability, dissipativity and passivity analysis. Another work introduces an energy notion of *meta-energy* in [58, 59] that

allows the dissipativity theory to incorporate thermodynamical systems. Furthermore, port-Hamiltonian systems have also attracted much attention in the last decades. Numerous papers have been published on reformulating thermodynamic problems as port-hamiltonian [57, 160, 92, 119, 161]. Despite the outstanding work carried out by these researchers, it appears that even for simple high-precision thermal systems, model-based controller synthesis, observer synthesis, or uncertainty descriptions using thermodynamical properties are hard to establish and generalize [76, 95].

In this chapter, the main contributions can be summarized as follows: 1) We study the dissipative properties of a class of thermal systems using the Lyapunov function proposed in Chapter 6 as a storage function when the system interacts with its surroundings. Specifically, when such interactions are time-dependent and can be considered as control actions. 2) Based on the entropy generation minimization, a bounded storage function is proposed in such a way that the storage function has an upper bound and a lower bound.

The remainder of the chapter is organized in the following manner. In Section 7.2 we introduce the fundamentals of dissipativity theory and non-equilibrium thermodynamics. An analysis is given in Section 7.3 for the dissipative properties of a thermal system for both time-dependent and time-independent boundary conditions. In Section 7.4, We aim to connect the minimal/maximal entropy generation with an efficiency criterion that expresses minimal/maximal loss of mechanical working rate in the system. Then, in Section 7.5, a 2-D heat conduction example is given and the profile of the corresponding storage function and supply rate are defined. The conclusion is given in Section 7.6.

## 7.2 Fundamentals

In the sequel, we first briefly introduce the dissipativity theory, which includes the definition of storage function, supply rate and the dissipation inequality. For the detailed exposition of the dissipation theory presented in this paper, we refer to [171, 172]. Second, some fundamentals of non-equilibrium thermodynamics together with the entropy balance equation, also called the second law of thermodynamics for open systems, will be given.

### 7.2.a Dissipative systems

Consider a dynamical system  $\Sigma$  described in the state space form and given by the equation

$$\Sigma : \quad \dot{x} = f(x, u), \quad y = h(x, u) \quad (7.1)$$

where  $x(t) \in \mathbb{X} = \mathbb{R}^n$ ,  $u(t) \in \mathbb{U} = \mathbb{R}^m$  and  $y(t) \in \mathbb{Y} = \mathbb{R}^q$  denote the state, input and output. Furthermore,  $f$  is assumed to be a smooth function defined on  $x \times u$ , and  $h$  is the output map of the system. Then, let

$$w : \mathbb{U} \times \mathbb{Y} \rightarrow \mathbb{R} \quad (7.2)$$

be a real-valued function, called the *supply rate*. Throughout, we assume that the mapping  $t \mapsto w(u(t), y(t))$  is locally integrable for any solution  $(u, x, y)$  of (7.1) in the sense that

$$\int_{t_0}^{t_1} |w(u(\tau), y(\tau))| d\tau < \infty$$

for all solutions  $(u, x, y)$  of (7.1) and all  $t_0 \leq t_1$ . The dissipation property for system (7.1) is given in the following definition

**Definition 7.1** (Adapted from [171]) *The system  $\Sigma$  is said to satisfy the dissipation inequality with respect to the supply rate  $w$  and the storage function  $S : \mathbb{X} \rightarrow \mathbb{R}$  if*

$$S(x(t_2)) - S(x(t_1)) \leq \int_{t_1}^{t_2} w(u(t), y(t)) dt$$

*holds for all  $(x, u, y)$  that satisfies (7.1) and for all  $t_1, t_2 \in \mathbb{R}$ , with  $t_2 \geq t_1$ .*

**Definition 7.2** *The system  $\Sigma$  is dissipative w.r.t. the supply function  $w$  if there exists a storage function  $S : \mathbb{X} \rightarrow \mathbb{R}$  so that  $\Sigma$  satisfies the dissipation inequality in Definition 7.1 w.r.t. supply rate  $w$  and storage function  $S$ .*

With a differentiable storage function  $S(\cdot)$ , we obtain the following equivalent differential dissipation inequality.

**Definition 7.3** *If the storage function  $S(\cdot)$  is differentiable, then the system  $\Sigma$  satisfies the dissipation inequality w.r.t. supply  $w$  and storage function  $S$  if and only if*

$$\frac{d}{dt} S(x(t)) \leq w(u(t), y(t))$$

*for all trajectories  $(u, x, y)$  satisfying (7.1) and for time  $t$ .*

The above dissipation inequality expresses the concept of an open dynamical system  $\Sigma$ , which interacts with its surroundings through  $u$  and  $y$ , the rate at which internal storage changes cannot exceed the amount which has been supplied to the system. The power supplied to the system is denoted by a function  $w(u, y)$  and is a function of the input and the output, which has the meaning of the rate at which a relevant quantity (heat flow, mass flow, power flow) flows in and out of the system. Specifically, the  $w$  is counted as positive when power flows into the system and negative if power flows out of the system (is delivered to its environment). The power supply  $S(x)$  is a function of the state of the system and expresses the amount of storage when the system finds itself in state  $x$ . The dissipation is defined as the difference between what is supplied and what is stored.

**Remark 7.1** *The dissipation inequality in Definition 7.1 does not require the storage function to be non-negative (It is required to be non-negative  $S(x) \in \mathbb{R}_{\geq 0}$  in [171, 172]). The requirement of non-negativity imposes an extra lower bound to the storage function. There are some arguments against adding the non-negativity in the definition [169]. In the context of thermodynamics in this chapter, we will assume the storage function to be any function  $S : \mathbb{X} \rightarrow \mathbb{R}$ , and therefore do not assume lower bounds on storage functions.*

In the following, we introduce two special storage functions the *available storage*  $S_a(x)$  and the *required supply*  $S_r(x)$ . We assume that the state space model (7.1) has a specific ground state (or equilibrium state)  $x^* \in X$  where we declare storage to be neutral or normalized to 0. Typically,  $x^* = 0$ .

**Definition 7.4** The available storage,  $S_a : \mathbb{X} \rightarrow \mathbb{R} \cup \{\infty\}$  of a dynamical system (7.1) with supply rate  $w$  is defined by

$$S_a(x_0) := \sup \left\{ - \int_0^{t_+} w(u(\tau), y(\tau)) d\tau \mid t_+ \geq 0, x(0) = x_0, x(t_+) = x^* \right\}$$

where the supremum is taken over all input trajectories  $u : [0, t_+] \rightarrow \mathbb{R}^m$  and time instants  $t_+ \geq 0$  that denoted  $x(t; u)$  as defined in (7.1) with  $x(0) = x_0$  at time  $t = 0$  to  $x(t_+; u) = x^*$  at time  $t = t_+$ .

When the system absorbs energy from surroundings or dissipates energy to the surroundings, there are limitations in the maximal amount of energy  $S_a(x)$  that can be extracted from the system while moving from state  $x$  some reference state  $x^*$ . Similarly, there are limitations on the minimal amount of storage necessary to bring the system from this reference state  $x^*$  to  $x$ . Similarly, the required supply  $S_r$  is defined as

**Definition 7.5** The required supply,  $S_r : \mathbb{X} \rightarrow \mathbb{R} \cup \{-\infty\}$  of a dynamical system (7.1) with supply rate is defined by

$$S_r(x_0) := \inf \left\{ \int_{t_-}^0 w(u(\tau), y(\tau)) d\tau \mid t_- \leq 0, x(t_-) = x^*, x(0) = x_0 \right\}$$

where the infimum is taken over all input trajectories  $u : [0, t_+] \rightarrow \mathbb{R}^m$  and time instants  $t_- \leq 0$  that denoted  $x(t; u)$  as defined in (7.1) with  $x(t_-) = x^*$  at time  $t = t_-$  to  $x(0) = x_0$  at time  $t = 0$ .

With the theorem introduced in [148], the following inequality can be obtained

**Theorem 7.6** Suppose that  $\Sigma$  is controllable. Then  $\Sigma$  is dissipative with supply  $w$  (according to Definition 7.2) if and only if  $S_a(x) < \infty$  for all  $x$  and if and only if  $S_r(x) > -\infty$  for all  $x$ . Moreover, for any storage function  $S(x)$  that satisfies  $S(x^*) = 0$  the inequality

$$S_a(x) \leq S(x) \leq S_r(x)$$

holds for all  $x$ .

If there is no dissipation in a system, then the system can be defined as conservative or lossless, meaning that the inequalities in Definition 7.2 and Definition 7.5 are, in fact, equalities.

**Definition 7.7** A dynamical system (7.1) is said to be lossless with respect to supply rate  $w$  and storage function  $S$  if

$$S(x(t_2)) = \int_{t_1}^{t_2} w(u(\tau), y(\tau)) d\tau + S(x(t_1))$$

for all  $(u, x, y)$  that satisfying (7.1) and for all  $t_1 \leq t_2 \in \mathbb{R}$ .

## 7.2.b Irreversible thermodynamics and balance equations

The thermodynamical system considered in this chapter is assumed to be in a local equilibrium and it is characterized by the *extensive* and *intensive* variables. Its behaviour is a subset

$$\mathcal{T} \subset \mathcal{X}^{\text{ext}} \times \mathcal{X}^{\text{int}},$$

where the extensive variables  $x^{\text{ext}} = \text{col}(U, S, M, N_1, \dots, N_r) \in \mathcal{X}^{\text{ext}}$  consist of internal energy  $U$ , entropy  $S$ , mass  $M$  and the mole numbers  $N_i$ . The intensive variables  $x^{\text{int}} = \text{col}(T, P, \mu_1, \dots, \mu_r) \in \mathcal{X}^{\text{int}}$  consist of temperature  $T$ , pressure  $P$  and chemical potential  $\mu_i$  of each component, which is obtained as the first differential of the fundamental equation that writes the extensive variable  $U$  as explicit function  $U = U(S, V, N_1, \dots, N_r)$  of the remaining extensive variables: where the intensive variables are defined by the following relations

$$T := \left( \frac{\partial U}{\partial S} \right)_{V, N_1, \dots, N_r}, P := - \left( \frac{\partial U}{\partial V} \right)_{S, N_1, \dots, N_r}, \mu_i := \left( \frac{\partial U}{\partial N_i} \right)_{S, V, \dots, N_k, \dots}.$$

Similarly, one can write the extensive variable  $S$  as function  $S = S(U, V, N_1, \dots, N_r)$  of the remaining extensive variables (the so called entropic representation). Leading to the equivalent

$$\frac{1}{T} := \left( \frac{\partial S}{\partial U} \right)_{V, N_1, \dots, N_r}, -\frac{P}{T} := - \left( \frac{\partial S}{\partial V} \right)_{S, N_1, \dots, N_r}, \frac{\mu_i}{T} := \left( \frac{\partial S}{\partial N_i} \right)_{S, V, \dots, N_k, \dots}.$$

Based on the fundamental postulate of thermodynamics [54], the entropy balance equation of a system can be characterized via

$$\frac{dS_{\text{sys}}}{dt} = \frac{dS_e}{dt} + \frac{dS_i}{dt}, \quad (7.3)$$

where  $S_e$  is the total entropy supplied by its surrounding across the external boundary of the system and  $S_i$  denotes the total entropy production due to processes in the interior of the system. In this chapter, we consider the system including irreversible process, in which can be characterized by  $S_{\text{gen}}$

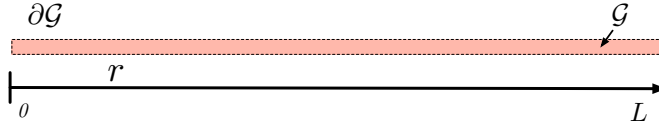
$$\frac{dS_i}{dt} = S_{\text{gen}}. \quad (7.4)$$

so that  $S_{\text{gen}} = \frac{dS_{\text{sys}}}{dt} - \frac{dS_e}{dt}$  is the difference of entropy rates between system and

total supplied entropy, i.e., entropy rate only due to changes in the interior of the system. Note that  $S_{gen} = 0$  for reversible process, and  $S_{gen} > 0$  for irreversible process, and the term  $S_e$  can be negative, positive or zero if there is no interaction with the system [54].

### 7.3 Dissipative properties of thermal systems

Consider a thermal system of 1-dimensional heat conduction in a solid, depicted in Figure 7.1, whose governing equation and boundary condition are given by



**Figure 7.1:** 1-D Heat conduction in a solid with geometry  $\mathcal{G}$  and boundary  $\partial\mathcal{G}$  (dashed line).

$$\rho c_p \frac{\partial T(r, t)}{\partial t} = -\nabla \cdot J(r, t), \quad \text{in } \mathcal{G}, \quad (7.5a)$$

$$\nabla T(r, t) \cdot \bar{n} = \gamma(t) \quad \text{on } \partial\mathcal{G}, \quad (7.5b)$$

$$T(r, 0) = T_0(r) \quad \text{at initial time } t = 0. \quad (7.5c)$$

Here,  $T : \mathcal{G} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{>0}$  is an analytic function of the 1-dimensional domain  $\mathcal{G} = [0, L]$  with  $L > 0$  that represents temperature. Density and heat capacity of the solid are defined by  $\rho$  and the  $c_p$ , respectively. The heat flux is  $J(r, t) := -\kappa \nabla T(r, t)$ ,  $r \in \mathcal{G}$  where the heat transfer coefficient  $\kappa > 0$  is assumed constant and temperature independent. Note that compared to (6.15) the boundary condition is relaxed to time-dependent function  $\gamma(t)$ . As in the previous chapter, the set of all temperature  $T(\cdot, t)$  is assumed to be a mapping from  $\mathcal{G}$  to  $\mathbb{R}_{>0}$  for any  $t \geq 0$ . The set of all such mappings is denoted by  $\mathcal{D}$ . Finally,  $T_0 \in \mathcal{D}$  is a given initial temperature profile at time  $t = 0$ .

In this chapter we write the distributed parameter thermo-dynamical system (7.5) as a dissipative dynamical system. This requires the introduction of a suitable supply function and a suitable storage function so as to verify the dissipation inequality in Definition 7.1 for the dynamics given in (7.5). For this, we propose to consider the entropy generation function  $S_{gen}$  as a suitable storage function that the thermodynamic system becomes dissipative. Under the assumption of local thermodynamic equilibrium [46], the irreversible entropy can be written as into a linear mapping between the thermodynamic flux  $J$  and the thermodynamic force  $F$ .

**Theorem 7.8** *Assume the system (7.5) that the irreversible entropy rate can be expressed as a linear relation between the thermodynamic flux  $J = -\kappa \nabla T$ ,  $\kappa \in \mathbb{R}_{>0}$  and the force*

$F = \nabla \frac{1}{T}$ . Let the irreversible entropy rate (entropy generation function)  $S_{gen}$  be the storage function, then this system is dissipative with respect to the supply rate  $w$  if

$$S_{gen}(T(\cdot, t_2)) - S_{gen}(T(\cdot, t_1)) \leq \int_{t_1}^{t_2} w(u(\tau), y(\tau)) d\tau$$

holds for all  $t_2 \geq t_1$ , with  $t_1, t_2 \in \mathbb{R}$ , here the supply rate  $w$  is given as

$$w(u, y) = \int_{\partial\mathcal{G}} \left( J \cdot \frac{\partial}{\partial t} \frac{1}{T} \right) \cdot \bar{n},$$

here  $\partial\mathcal{G}$  denotes the boundary of the system,  $(J, \frac{\partial}{\partial t} \frac{1}{T})$  denote the input and the output of the system.

**Proof:** The entropy generation for the system (7.5) over the domain  $\mathcal{G}$  is

$$S_{gen}(T) = \int_{\mathcal{G}} J(T) \cdot F(T) dr = \int_{\mathcal{G}} -\kappa \nabla T \cdot \nabla \frac{1}{T} dr.$$

Take the time derivative of the composite function  $t \mapsto S_{gen}(T(r, t))$  and denote this derivative as  $\dot{S}_{gen}$  along the solution of (7.5a) and it reads

$$\begin{aligned} \dot{S}_{gen}(T) &= \int_{\mathcal{G}} \frac{\partial}{\partial t} J(T) \cdot F(T) + J(T) \cdot \frac{\partial}{\partial t} F(T) dr \\ &= -\kappa \int_{\mathcal{G}} \left[ \nabla \dot{T} \cdot \nabla \frac{1}{T} + \nabla T \cdot \nabla \left( \frac{\partial}{\partial t} \frac{1}{T} \right) \right] dr \end{aligned}$$

Using the Divergence theorem and it follows that

$$\begin{aligned} \dot{S}_{gen}(T) &= -\kappa \int_{\mathcal{G}} \left[ \nabla \cdot \left( \frac{\partial}{\partial t} \frac{1}{T} \nabla T \right) + \nabla \cdot \left( \dot{T} \nabla \frac{1}{T} \right) - \dot{T} \nabla^2 \frac{1}{T} - \left( \frac{\partial}{\partial t} \frac{1}{T} \right) \nabla^2 T \right] dr \\ &= \kappa \int_{\partial\mathcal{G}} \frac{1}{T^2} \dot{T} \nabla T \cdot \bar{n} d\bar{r} + \kappa \int_{\mathcal{G}} \left( \frac{\partial}{\partial t} \frac{1}{T} \right) \nabla^2 T dr \\ &= \int_{\partial\mathcal{G}} \left( J \cdot \frac{\partial}{\partial t} \frac{1}{T} \right) \cdot \bar{n} d\bar{r} - \int_{\mathcal{G}} \nabla J \cdot \left( \frac{\partial}{\partial t} \frac{1}{T} \right) dr \\ &= \kappa \int_{\partial\mathcal{G}} \frac{1}{|T|^2} \dot{T} \nabla T \cdot \bar{n} d\bar{r} - \int_{\mathcal{G}} \nabla J \cdot \left( \frac{\partial}{\partial t} \frac{1}{T} \right) dr \end{aligned} \quad (7.6)$$

here  $\bar{r}$  refers the location over the boundary  $\partial\mathcal{G}$ .

- 1) For the case of time-independent boundary condition on the boundary  $\partial\mathcal{G}$ , the first term on the right-hand side of (7.6) vanishes which is proved in (6.24);
- 2) For the case of time-dependent boundary condition in (7.5b), the first term on the right-hand side of (7.6) denotes the irreversible entropy and results from the interactions with surroundings.



In this proof, we focus on the second case 2) of a time-dependent boundary condition. By substituting (7.5a), the second term on the right-hand side of (7.6) is rewritten as

$$-\int_{\mathcal{G}} \nabla J \cdot \left( \frac{\partial}{\partial t} \frac{1}{T} \right) dr = -\rho c_p \int_{\mathcal{G}} \left| \frac{\dot{T}}{T} \right|^2 dr \leq 0.$$

Moving this second term on the right-hand of the equation  $\int_{\partial G}(\cdot)$  in (7.6) to the left-hand side of the equality, we obtain the following relation

$$\begin{aligned} \dot{S}_{gen}(T) + \int_{\mathcal{G}} \nabla J \cdot \left( \frac{\partial}{\partial t} \frac{1}{T} \right) dr &= \int_{\partial \mathcal{G}} \left( J \cdot \frac{\partial}{\partial t} \frac{1}{T} \right) \cdot \bar{n} d\bar{r} \\ \implies \dot{S}_{gen}(T) &\leq \int_{\partial \mathcal{G}} \left( J \cdot \frac{\partial}{\partial t} \frac{1}{T} \right) \cdot \bar{n} d\bar{r}. \end{aligned}$$

Since by integration from  $t_1$  to  $t_2$ , one has

$$S_{gen}(T(\cdot, t_2)) - S_{gen}(T(\cdot, t_1)) \leq \int_{t_1}^{t_2} w(\tau) d\tau$$

with the supply rate

$$w(\tau) = \int_{\partial \mathcal{G}} \left( J(\tau, \bar{r}) \cdot \frac{\partial}{\partial t} \frac{1}{T}(\tau, \bar{r}) \right) \cdot \bar{n} d\bar{r}.$$

Thus, the system (7.5) is proved to be dissipative with respect to the supply rate function  $w$  of the input and the output  $(J, \frac{\partial}{\partial t} \frac{1}{T})$ .  $\square$

For the detailed derivation involved in (7.6) we refer to the proof of Theorem 6.2.

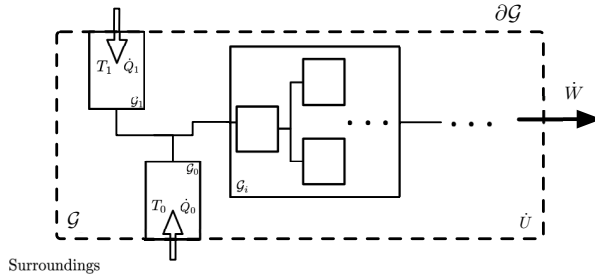
## 7.4 Dissipative properties of interconnected systems

From the discussion of the dissipation inequality using the proposed storage function  $S_{gen}$  for a single system in (7.5), one remaining question is if there exists a physically interpretable upper bound and a lower bound for the storage functions in a thermodynamical system. In [5], authors introduce the idea of a supporting hyperplane using one arbitrary reference state for the thermodynamic availability (as the storage function). In such a formulation the *minimal available* storage function is defined within the projection of the hyperplane through a reference state. However, this bound, presented in [5], is a local property that is reference state-dependent. This *minimal available* storage is not physically interpretable in the sense that the bridge between the system behaviours and the proposed minimal available storage energy is not linked. In the section, we attempt to connect the Gouy-Stodola theorem and entropy generation minimization in [27] which are easily interpretable with system behaviours, to the formulation of the upper bound and the lower bound for the storage function.

### 7.4.a $S_{gen}$ extension to interconnected systems

In the previous part, only a single system is considered. In this section we consider a system composed of multiple interconnected subsystems. As shown in Figure 7.2, this system consists of the internal energy  $\dot{U}$  within the domain  $\mathcal{G}$ , the transfer of rate of work  $\dot{W}$ , the heat flow transfer rate  $\dot{Q}_i$  with contact temperature  $T_i$  through the boundary  $\partial\mathcal{G}$ . We suppose that the entire system consists of the interaction of  $\ell$  components (or subsystems). Figure 7.2 illustrates the idea and indicates two thermal interconnections with temperature  $T_0, T_1$  and with heat rates  $\dot{Q}_0$  and  $\dot{Q}_1$ . Mechanical work such as thermal deformation in high-precision machines can be denoted by the rate of work  $\dot{W}$ .

We consider a complex system that consists of a number of subsystems that are thermally interconnected, where the total interconnected system  $\mathcal{G}$  has a boundary  $\partial\mathcal{G}$  and where the heat flow rate and contact temperature of the  $i$ th subsystem with its surrounding are denoted by  $\{\dot{Q}_i, T_i\}$ .



**Figure 7.2:** The system consists of multiple subsystems which are interconnected with each other, the domain of each subsystem is denoted by  $\mathcal{G}_i, \mathcal{G}_i \subset \mathcal{G}$  where  $\mathcal{G}$  represents the whole system domain, the dashed line denotes the boundary  $\partial\mathcal{G}$  of the system. The arrow with  $\dot{W}$  characterizes the work done by the system on the surroundings, the arrow points towards subsystems  $\mathcal{G}_0$  or  $\mathcal{G}_1$  denotes the heat flow rate  $\dot{Q}_0$  or  $\dot{Q}_1$  flow into the system. The heat flow rate is the integral over the surface where the heat flux rate  $J_i$  crosses the surface  $\dot{Q}_i = \int_{\mathcal{S}_i} J_i$  with  $J_i = -\kappa_i \nabla T_i$  and  $\mathcal{S}_i$  the boundary or the surrounding of the  $i$ th subsystem.

The first law of thermodynamics for the system in Figure 7.2 reads

$$\frac{dU}{dt} = \sum_{i=0}^{\ell} \frac{dQ_i}{dt} - \frac{dW}{dt} \quad (7.7)$$

where the heat flow rate  $\frac{dQ_i}{dt} = \int_{\text{surf}} J_i$  is the integral of the heat flux  $J_i$  over the surface. The internal energy  $U$  is a function of temperature  $T(r, t)$  with  $r \in \mathcal{G}, t \geq 0$ .  $T(r, t)$  is assumed to a mapping from  $\mathcal{G}$  to  $\mathbb{R}_{>0}$ . The set of all such mapping is denoted by  $\mathcal{D}$ . Specifically, the boundary condition is defined as  $\nabla T(r, t) = \gamma(t)$  which is time-dependent function. In the following, we abuse the  $\frac{dQ}{dt} = (\dot{Q})$ ,  $\frac{dU}{dt} = \dot{U}$  and  $\frac{dW}{dt} = \dot{W}$  for simplicity of notation. The second law of thermodynamics of

this system is

$$\begin{aligned}\frac{dS_{sys}}{dt}(t) &= \sum_{i=0}^{\ell} \frac{\dot{Q}_i}{T_i}(t) + S_{gen}(T(t)) \\ &= \frac{\dot{Q}_0}{T_0}(t) + \sum_{i=1}^{\ell} \frac{\dot{Q}_i}{T_i}(t) + S_{gen}(T(t))\end{aligned}\quad (7.8)$$

here  $S_{gen}(T)$  denotes the irreversible entropy generation rate (7.4) due to thermal interaction. Notice that the  $T$  in  $S_{gen}(T(t))$  represents a set of temperature  $[T_0, T_1, \dots, T_{\ell}]$ . Under the assumption that the irreversible entropy of the system  $\mathcal{G}$  is only caused by the thermal interactions  $\{\dot{Q}_0, \dots, \dot{Q}_{\ell}\}$  across the boundary, then the entropy generation rate  $S_{gen}(T)$  of the system, with the supply function  $w$  defined previously, is

$$S_{gen}(T) = \int_{\mathcal{G}_0} J_0 \cdot \nabla \frac{1}{T_0} dr_0 + \dots + \int_{\mathcal{G}_{\ell}} J_{\ell} \cdot \nabla \frac{1}{T_{\ell}} dr_{\ell} \quad (7.9)$$

It is notable that the  $S_{gen}$  is a composite function of temperature  $T(r, t)$  which is temporal dependent, the notation  $S_{gen}(t)$  refers to  $S_{gen}(T(t))$  if it is not specified.

#### 7.4.b The upper bound and the lower bound of $S_{gen}$

We take a closer look at the entropy balance equation (7.8) and replace the  $\dot{Q}_0$  in (7.8) from (7.7) to infer that

$$\dot{W}(t) = \frac{d}{dt}(T_0 S_{sys}(t) - U(t)) + \sum_{i=1}^{\ell} \left(1 - \frac{T_0}{T_i}\right) \dot{Q}_i(t) - T_0 S_{gen}(t). \quad (7.10)$$

Note that the pair of  $(T_0, Q_0)$  is chosen as the reference temperature and the heat flow for the analysis of  $S_{gen}$ , respectively. In the case where all work is reversible, we replace  $\dot{W}$  by  $\dot{W}_{rev}$  and the (7.7) can be rewritten as:

$$\dot{W}_{rev}(t) = \frac{d}{dt}(T_0 S_{sys}(t) - U(t)) + \sum_{i=1}^{\ell} \left(1 - \frac{T_0}{T_i}\right) \dot{Q}_i(t) \quad (7.11)$$

where the entropy generation rate is  $S_{gen}(t) = 0$  for all  $t$ . Subtracting (7.10) from (7.11) we arrive at the formula

$$\dot{W}_{rev}(t) - \dot{W}(t) = T_0 S_{gen}(t) \quad (7.12)$$

or

$$\frac{1}{T_0} (\dot{W}_{rev}(t) - \dot{W}(t)) = S_{gen}(t) \quad (7.13)$$

which is also known as *Gouy-Stodola theorem* [78, 79, 157] in thermodynamics. By defining the rate of mechanical loss, say  $\dot{W}_{loss}$ , as the difference between rate of

work of the reversible part  $\dot{W}_{rev}$  and the rate of (the actual) work  $\dot{W}$ , i.e,  $\dot{W}_{loss} = \dot{W}_{rev} - \dot{W}$ , the equation (7.12) and the equation 7.13 can be rewritten as

$$\dot{W}_{loss}(t) = T_0 S_{gen}(t) \quad (7.14)$$

or

$$\frac{1}{T_0} \dot{W}_{loss}(t) = S_{gen}(t) \quad (7.15)$$

here  $\dot{W}_{loss}$  represents the rate of mechanical work loss done by the system that can not be recovered by the system. Remark that regardless of the system described in Figure 7.2 is a power generator  $\dot{W} > 0$  (e.g., power plant) or a power consumer  $\dot{W} < 0$  (e.g., computer), the  $\dot{W}_{loss}$  is always non-negative. In this chapter, we only consider the case of  $\dot{W} > 0$  if it is not specified.

The idea of introducing the reversible work rate  $\dot{W}_{rev}$  is inspired by the research of chemical processes where the chemical reaction leads to reversible and irreversible work. For the case of the system described in Figure 7.2, the work rate  $\dot{W}(t)$  can be contributed from the thermo-induced mechanical deformation in which we may distinguish between plastic deformation and elastic deformation. The elastic deformation is reversible, and the plastic deformation can be considered as energy loss [129, 96] which is irreversible. Using the idea of  $\dot{W}_{rev}$ , we can connect the storage function to the work efficiency. By defining the work rate efficiency

$$\eta := 1 - \frac{\dot{W}}{\dot{W}_{rev}} = \frac{\dot{W}_{rev} - \dot{W}}{\dot{W}_{rev}} = \frac{\dot{W}_{loss}}{\dot{W}_{rev}}, \quad \text{with} \quad \frac{\dot{W}}{\dot{W}_{rev}} \geq 0. \quad (7.16)$$

Here the  $\dot{W}, \dot{W}_{rev}$  can be positive or negative but share the same sign; this leads to the non-negative  $\eta$ .

Once the idea of work rate efficiency is presented, we start introducing the available entropy generation: the lower bound of the storage function  $S_{gen}$  which is allowed to be extracted from a dynamical system at any time.

**Definition 7.9** *The minimal entropy generation  $S_{gen}^-$  of a dynamical system (as in Figure 7.2) is defined as*

$$S_{gen}^-(t) := \min_{T(t): T \in \mathcal{T}} S_{gen}(T(t))$$

where

$$\mathcal{T} := \{T(t) \in \mathcal{D} \mid t_0 \leq t \leq t_+, T(t_0) = \tilde{T}, T(t_+) = T^*\}$$

and

$$\mathcal{D} := \{T : \mathcal{G} \rightarrow \mathbb{R}_{>0}\}$$

where the minimum is taken over all possible state trajectories  $T(t) \in \mathcal{T}$ . Let  $T_{opt}^-$  be a function of  $t$  and satisfies

$$S_{gen}^-(t) = S_{gen}(T_{opt}^-(t)) \leq S_{gen}(T(t)), \quad \text{for } \forall t, \forall T(t) \in \mathcal{T}.$$

**Theorem 7.10** *For a dynamical system in Figure 7.2 with a storage function is given by the entropy generation  $S_{gen}$ , the minimal entropy generation  $S_{gen}^-$  satisfies*

$$0 \leq S_{gen}^-(t) \leq S_{gen}(t)$$

for all  $t_- \leq t \leq t_+$ .

**Proof:** By Definition 7.9, the  $S_{gen}^-$  is defined as the lower bound of the entropy generation

$$S_{gen}^-(t) = S_{gen}(T_{opt}^-(t)) \leq S_{gen}(t),$$

additionally, the  $S_{gen}(t)$  is defined as the non-negative and  $S_{gen}^-(t) \geq 0, \forall t$ . Thus, the inequality is proved.  $\square$

By considering the work loss formulation (7.14), the entropy generation can be represented as the function of work loss rate  $\dot{W}_{loss}(t)$ . Using the definition of work rate efficiency, the  $S_{gen}(t)$  is

$$S_{gen}(t) = \frac{1}{T_0} \dot{W}_{rev}(t) \cdot \eta(t)$$

here  $\dot{W}_{rev}(t)$  denotes the ideal work rate for every time instance  $t$  and can be considered as the reference for the  $\dot{W}(t)$  in (7.16). Notice that the above representation indicates that the entropy generation  $S_{gen}(t)$  is a function of the work efficiency rate  $\eta(t)$  and this leads to the following conjecture

**Conjecture 7.11** *If work rate efficiency of a dynamical system in Figure 7.2 is defined as in (7.16), then the minimal entropy generation satisfies*

$$S_{gen}^-(t) := \min_{\eta(t) \rightarrow, t_0 \leq t \leq t_+} \frac{1}{T_0} \dot{W}_{rev}(t) \cdot \eta(t) \quad (7.17)$$

where  $\eta(t) \rightarrow$  denotes the infimum over all point-wise work rate efficiency from  $t = t_0$  to  $t = t_+$ . Moreover, there exists a function  $\eta_{opt}^-(t)$  that satisfies (7.17)

$$S_{gen}^-(t) = \frac{1}{T_0} \dot{W}_{rev}(t) \cdot \eta_{opt}^-(t) \leq S_{gen}(t) = \frac{1}{T_0} \dot{W}_{rev}(t) \cdot \eta(t),$$

and

$$S_{gen}^-(t) > 0 \quad \text{for all } t.$$

In the next, the maximal entropy generation  $S_{gen}^+$  is given: the upper bound of the storage function  $S_{gen}$  for which is allowed to provide at any time from a dynamical system.

**Definition 7.12** *The maximal entropy generation  $S_{gen}^+$  of a dynamical system (as in Figure 7.2) is defined as*

$$\begin{aligned}
S_{gen}^+(t) : &= \max_{T(t): T \in \mathcal{T}} S_{gen}(T(t)) \\
&= S_{gen}(T_{opt}^+(t)) \geq S_{gen}(t), \text{ for } \forall t, \forall T(t) \in \mathcal{T}
\end{aligned}$$

where

$$\mathcal{T} := \{T(t) \in \mathcal{D} \mid t_0 \leq t \leq t_+, T(t_0) = \tilde{T}; T(t_+) = T^*\}$$

and

$$\mathcal{D} := \{T : \mathcal{G} \rightarrow \mathbb{R}_{>0}\}$$

where the maximum is taken over all possible state trajectories  $T(t) \in \mathcal{T}$  from  $t_0$  to  $t_+$ .

By connecting to the work rate efficiency, the maximal entropy generation also can be formulated as

**Conjecture 7.13** *If work rate efficiency of a dynamical system in Figure 7.2 is defined as in (7.16), then the required entropy generation satisfies*

$$S_{gen}^+(t) = \max_{\eta(t) \rightarrow, t_- \leq t \leq t_0} \frac{1}{T_0} \dot{W}_{rev}(t) \cdot \eta(t)$$

where  $\eta(t) \rightarrow$  denotes the supremum over all point-wise work rate efficiency from  $t = t_-$  to  $t = t_0$ . Moreover, there exists a function  $\eta_{opt}^+(t)$  that satisfies the above equation

$$S_{gen}^+(t) = \frac{1}{T_0} \dot{W}_{rev}(t) \cdot \eta_{opt}^+(t) \geq S_{gen}(t) = \frac{1}{T_0} \dot{W}_{rev}(t) \cdot \eta(t), \text{ for all } t$$

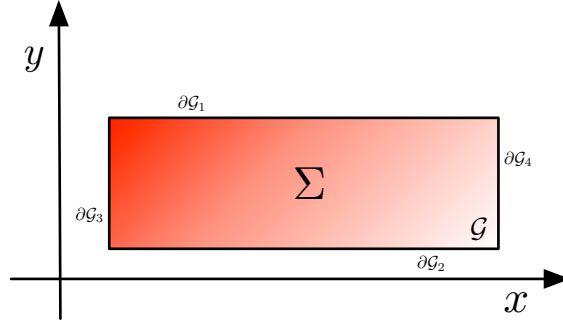
**Remark 7.2** *For a dynamical system which the irreversible entropy is considered as the storage function  $S_{gen}$ , with the minimal entropy generation  $S_{gen}^-$  in Theorem 7.10 and the maximal entropy generation  $S_{gen}^+(t)$  defined in Conjecture 7.13, we can obtain the following inequality*

$$0 \leq S_{gen}^- \leq S_{gen} \leq S_{gen}^+.$$

*In particular, by considering the work rate, the  $S_{gen}^-$  has physically-relevant interpretations: the lower bound of the storage function represents the least loss process when  $S_{gen}$  is allowed to extract from a dynamical system, and the  $S_{gen}^+$ : the upper bound of the storage function is the most loss process when a dynamical system can provide.*

## 7.5 Numerical example: 2-D heat conduction

In this section, we provide a computation of the temperature distribution of a 2-D heat conduction example with time-dependent boundary conditions, depicted in Figure 7.3. The temperature distribution is described by (7.5a) with  $J = -\kappa \nabla T$ . The coefficients are defined as  $\{\rho, c_p, \kappa\} = \{2.7 \times 10^3, 0.91, 209\}$  with proper units.



**Figure 7.3:** 2D heat conduction in solids with time-dependent boundary conditions.

The boundary conditions are given by

$$\begin{aligned}\bar{n} \cdot \nabla T(x, y, t) &= 0 \quad \text{on } \partial G_1, \partial G_2 \\ \bar{n} \cdot \nabla T(x, y, t) &= 2 \sin(2t) \quad \text{on } \partial G_3 \\ \bar{n} \cdot \nabla T(x, y, t) &= -\cos(2t) \quad \text{on } \partial G_4\end{aligned}\tag{7.18}$$

which represents an isolated boundary on  $\partial G_1$  and  $\partial G_2$  and a time varying heat flux at the boundaries  $\partial G_3$  and  $\partial G_4$ . The initial condition are defined as

$$T(x, y, 0) = 273.15[K] \quad \text{in } G.$$

The geometry of the system is given by a rectangle with  $0 \leq x \leq 1.5[\text{m}]$  and  $0 \leq y \leq 0.2[\text{m}]$ . The solutions are obtained by applying a spatial discretization method which generates 46 cells, and reads the temperature profiles of all spatial locations in Figure 7.4.

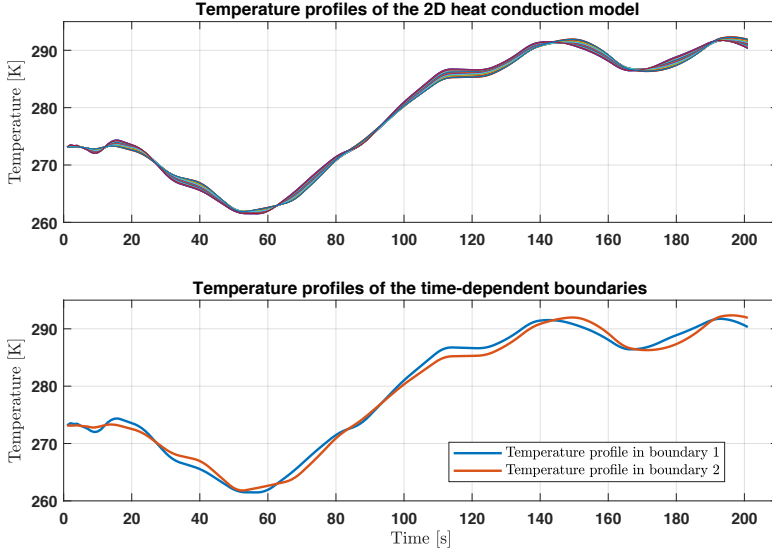
As defined in Theorem 7.8, the storage function rate  $\dot{S}_{gen}(T)$  is obtained by computing the storage function

$$S_{gen}(T(r, t)) = \int_G -\kappa \nabla T(r, t) \cdot \nabla \frac{1}{T(r, t)} dr$$

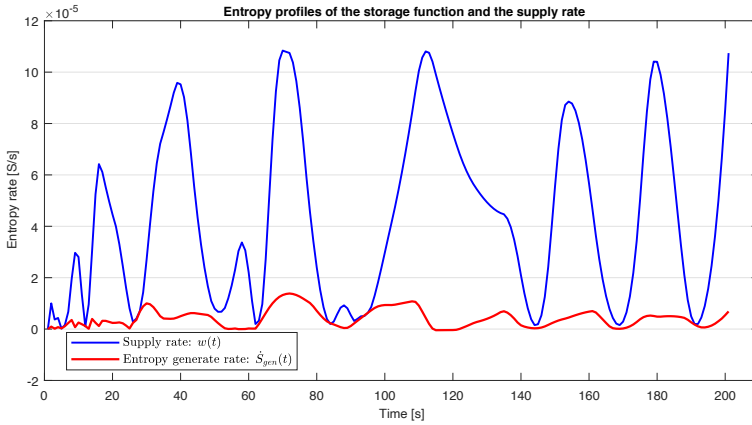
for  $t \in [0, 200]$ , and taking the time derivatives with respect to the time step 0.05. The supply rate is calculated via the equation

$$w(\tau) = \int_{\partial G} (J(\bar{r}, \tau) \cdot \frac{\partial}{\partial \tau} \frac{1}{T(\bar{r}, \tau)}) \cdot \bar{n} d\bar{r},$$

since the temperature profiles of cells at the boundary are known and the time derivative of  $1/T(\bar{r}, t)$  is also computable. Note that  $\partial G$  denotes the boundaries  $\partial G_1 \times \partial G_2 \times \partial G_3 \times \partial G_4$  and we only consider the  $T(\bar{r})$  at the cells on these boundaries for computing  $w(\tau)$ . As indicated in the plot, for the whole period, the storage function rate is always less or equal to the supply rate  $\dot{S}_{gen}(t) \leq w(t)$  for all time



**Figure 7.4:** Temperature profiles of the 2D heat conduction in solid with time-dependent boundary conditions. The upper plot represents the temperature of all 46 cells over the simulation time  $[0, 200]$ . The lower plot shows the temperature at the boundaries  $[\partial\mathcal{G}_1, \partial\mathcal{G}_2]$ .



**Figure 7.5:** Entropy profiles of the 2D heat conduction in solids with time-dependent boundary conditions.



instances  $t \geq 0$ .

## 7.6 Conclusion

In this chapter, we present the dissipation theory using the irreversible entropy as the storage function. Time-independent and time-dependent boundary conditions are discussed in the dissipation formulation. By considering a thermal complex system which the thermal behaviours leads to possible energy lost, an energy-based upper bound and a lower bound for the proposed storage function are derived. This chapter aims to embed thermal dynamical systems in the classical approach of open dissipative dynamical systems, open in the sense of allowing external influence and excitation via mechanical work on the system. It is shown how the dissipation inequality is established with  $S_{gen}$  as the generated internal entropy function as storage function. A derivation of the mechanical losses in the system is provided and an example of a 2D system driven by time-varying boundary conditions is given to illustrate the theory.



## **Part III**

# **Tooling of parameterized model order reduction**



# Tooling of parameterized model order reduction for large-scale systems

---

**T**his chapter presents an overview of the tooling that has been developed on the basis of the multi-parameter and multi-frequency moment-matching method proposed in Chapter 2. A dedicated software toolbox named paraMOR, parameterized Model Order Reduction, has been developed for the reduction of parametric and non-parametric large-scale dynamic systems and their visualization. We briefly introduce the underlying programming concept of the tooling and present the key supporting functions for model order reduction and the visualization of results. A number of case studies are provided.

---

Model reduction techniques can be informally formulated as the problem of finding a less complex mathematical model, while preserving relevant properties of the original model in the abstracted reduced model. Typically, the relevance of the reduction method is largely determined by the properties which are preserved or inherited in the reduced model. For many of the SVD-based methods, stability and passivity of a model can be preserved. The Krylov subspace method can maintain the input-output relationship in a well-defined sense for the class of linear systems. For nonlinear systems, data-driven methods introduced in [146] have been proven to be a useful option since the reduced model interpolates the original system at the range where the nonlinear dynamics is dominating.

Apart from the properties which need to be preserved in the reduction procedure, the computational complexity and speed of computation of the reduced order model is critical for many implementations and applications of reduced order models. As the dimensionality and complexity of engineered systems invariably increases, there is a persistent need for computationally efficient methods to meet the demands on computationally feasible models in industry. To

the best knowledge of the author, most of the existing MATLAB-based toolboxes for model reduction [141, 31, 105, 51] target for small-scale (state dimension  $< 100$ ) and middle-scale LTI systems (state dimension  $100 - 1000$ ) or reduced-basis method based [88] Only one tool [164] approximates LPV systems using moment-matching based method.

Part of the research reported in this thesis involves the development of software tooling for model order reduction of dynamical systems. This has led to the development of the **paraMOR** toolbox, referring to **parameterized Model Order Reduction**. Based on the work in Chapter 2, we develop tooling that is compatible with MathWorks MATLAB on both Linux and Windows operating systems. In this software, we focus on the approximation of LTI systems of the form

$$\Sigma := \begin{cases} E\dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \quad (8.1)$$

where  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$  and  $y(t) \in \mathbb{R}^q$  denote, respectively, the state vector, the input and the output, and the reduction of the LPV systems of form

$$\Sigma(\theta) := \begin{cases} E(\theta)\dot{x}(t) = A(\theta)x(t) + B(\theta)u(t) \\ y(t) = C(\theta)x(t) + D(\theta)u(t) \end{cases} \quad (8.2)$$

here states  $x(t) \in \mathbb{R}^n$ , input  $u(t) \in \mathbb{R}^m$  and output  $y(t) \in \mathbb{R}^q$ . Here  $\theta$  denotes a vector of real parameters of dimension  $\ell$ . It is assumed that the system matrices depend in an affine manner on  $\theta$ . For  $E$ , for example, this means that  $E(\theta) = E_0 + E_1\theta_1 + \dots + E_\ell\theta_\ell$ , where  $E_i$  are real-valued matrices and where  $\theta_i$  denotes the  $i$ -th entry of  $\theta$ . The matrices  $A(\theta)$ ,  $B(\theta)$ ,  $C(\theta)$ ,  $D(\theta)$  admit the same structure.

The remaining of the chapter is organized as follows. In Section 8.1, the functionality overview of paraMOR is given. Afterwards, Section 8.2 presents the structure of paraMOR together with the workflow to use paraMOR. Specifically, some programming principles used in the tooling are described in Section 8.3. In Section 8.4, two examples using the paraMOR toolbox are given as case studies. This chapter is concluded in Section 8.5

## 8.1 Overview of paraMOR tooling

The goal of paraMOR is to provide a complete workflow. This starts from constructing the full-order model (8.1) or (8.2), then generate the reduced-order model and end with presenting the results of the reduced-order model in comparison with the full-order model. Specifically, this tooling enables the reading and construction of system matrices in (8.1) and (8.2) from third party modelling software such as COMSOL Multiphysics and ANSYS. The following Table 8.1 summarizes the functionalities of the paraMOR tooling.

Some remarks on the functionalities in Table 8.1 are given in order:

1. With the use of Livelink provided by COMSOL Multiphysics and the *APDL marc* supported by ANSYS, the two routines the `ssExtraction.m` and the

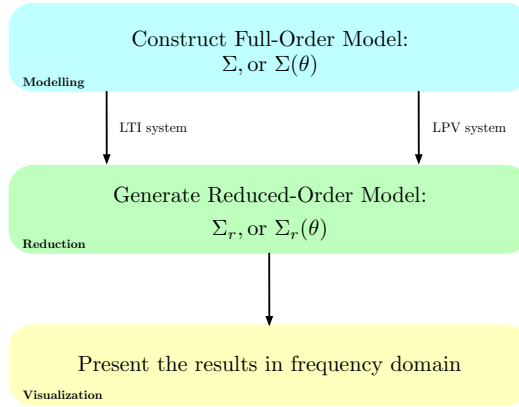
**Table 8.1:** Supported functions of the tooling

	System Class	Routine Name
<b>Modelling</b>	COMSOL model ANSYS model	
extract model	$\Sigma$ $\Sigma(\theta)$	ssExtraction.m pssExtraction.m
<b>Reduction</b>	$\Sigma \rightarrow \Sigma_r$ $\Sigma(\theta) \rightarrow \Sigma_r(\theta)$	krylovMOR(E, {}, A, {}, B, {}, s0, Opt) krylovMOR(Ep, p, Ap, p, Bp, p, s0, Opt)
<b>Visualization</b>	Freq. resp. $\Sigma$ or $\Sigma(\theta)$ Freq. resp. $\Sigma_r$ or $\Sigma_r(\theta)$	MIMOBodeplot

pssExtraction.m enable the construction of a linear time-invariant system (8.1) or a parameterized linear time invariant system (8.2) from a FEM model that is implemented in either COMSOL or ANSYS. The routines extract the state space matrices in (8.1) or (8.2) from the implicit model representations in these multi-physical system models. Note that this function is applicable to different physical systems and multi-physical systems.

2. The routine pssExtraction.m supports the parameterization of physical parameters such as density, heat capacitance or thermal conductivity for thermal systems, but not geometric parameters that define the spatial geometry of the system.
3. The reduction routine  $[V, W, \text{sysr}] = \text{krylovMOR}(\{\}, \{\}, \{\}, \{\}, \{\}, \{\}, \{\}, \{\})$  delivers the projection matrices  $V, W$  and the state space representation of the reduced model  $\text{sysr}$ . Besides the input matrices  $E, A, B$  or  $E_p, A_p, B_p$  and the parameters  $p$ , the  $s0$  includes the expansion points and the associated moments.
4. The option `Opt` describes the numerical configuration of model order reduction methods used for paraMOR. For example, the matrix type of the system representation `Opt.mtype = 'full', 'sparse'`, the orthogonalization methods for generating the Krylov subspace `Opt.orth = 'gs', 'mgs', 'rmgs'`, the orthogonality tolerance and the algorithm for solving the inverse of  $Ax = B$  problem `Opt.lse = 'dLU', 'sLU', 'chol'`.
5. The function MIMOBodeplot allows to plot the frequency response of multiple systems within one figure.

More information can be obtained on the specific use of these functions by using the build-in help functionality. The workflow mentioned above is illustrated in Figure 8.1.



**Figure 8.1:** General paraMOR workflow in which a FOM is constructed, then reduced and finally the results are analyzed.

## 8.2 Structure of the paraMOR

Based on the functionality of paraMOR, the functions of paraMOR are organized in the naming scheme such that it allows users to have easy access and interpret for specific usage. The toolbox is divided into the following sub-structures

<b>/src/Modelling</b>	Contains functions used to extract the system matrices from third-party software and construct them into state-space representation.
<b>/src/MOR</b>	Contains functions used to arrange the (multiple) expansion points with the associated moments and perform the model reduction.
<b>/src/Visua</b>	Contains functions used to compute and present the bode plot of the ROM and the FOM, compatible with SISO and MIMO systems.
<b>/demo</b>	Contains example code that explains the different main functionalities of the toolbox.
<b>/doc</b>	Contains a manual and a presentation file that details the theoretical background of the toolbox and programming implementation.

## 8.3 Programming principles

The default data type for matrix in MATLAB is stored in `dense` format. However, this format is not suitable for the large-scale sparse matrices ( $A \in \mathbb{R}^{n \times n}$ ,  $n \approx 5 \times 10^5$ ). In the paraMOR toolbox, all matrices are stored in `sparse` format.



As discussed in [49], the orthogonality of the projection matrix/matrices is critical to the fidelity of the reduced-order model. The Krylov subspaces associated with the image of the projection matrices ( $V, W$ ) defined in (2.21) and (2.22) can be generated by implicit and by explicit moment-matching. As has been observed in various applications [28], the explicit computation of the moments of the transfer function is numerically unstable, and the accuracy of the approximation method cannot be improved after 8 iterations. To remedy this, an orthonormal basis of the subspace spanned by the moment vectors can be obtained using an implicit calculation, so that a more accurate and numerically stable reduced-order model can be derived. In the paramOR toolbox, we provide three options for constructing the orthonormal basis or the Krylov subspace:

Method	Routine Name
Gram-Schmidt	<code>Opt.orth = 'gs'</code>
Modified Gram-Schmidt	<code>Opt.orth = 'mgs'</code>
Modified Gram-Schmidt with reorthogonalization	<code>Opt.orth = 'rmgs'</code>

**Table 8.2:** Supported orthogonal methods

Note that the option `Opt.orth = 'rmgs'` delivers the most accurate results on orthonormalization but at the expense of cost of extra computation time [77].

For moment-matching based model order reduction methods, the main computational burden lies in solving  $(s_0 E - A)^{-1}$  for constructing the Krylov subspace. In the paramOR toolbox, three types of matrix decomposition methods are provided for determining the inverse of  $(s_0 E - A)$ :

Method	Routine Name
LU decomposition for dense matrix	<code>Opt.lse = 'dLU'</code>
LU decomposition for sparse matrix	<code>Opt.lse = 'sLU'</code>
Cholesky decomposition	<code>Opt.lse = 'chol'</code>

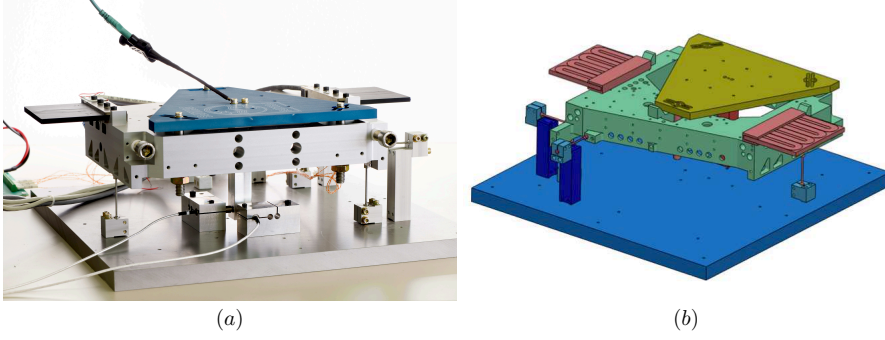
**Table 8.3:** Supported matrix decomposition methods

## 8.4 Case studies

In this section, we present the complete workflow using paramOR for two examples: one for an LTI system  $\Sigma$  of the form (8.1) and one for an LPV system  $\Sigma(\theta)$  of the form (8.2). The physical model is based on the PSA introduced in Section 4.5.

First, we give the case study of the LTI system based on the PSA setup using the paramOR toolbox.

1. The model is built in COMSOL Multiphysics,



**Figure 8.2:** (a): the real application of PSA; (b): the digitalized model of the PSA in COMSOL Multiphysics.

2. Extract the LTI state space representation using `ssExtraction.m`,
3. Obtain the system matrices  $\Sigma = \{E, A, B, C\}$  in (8.1) with  $E, A \in \mathbb{R}^{n \times n}$ ,  $n = 38476$  and  $B \in \mathbb{R}^{n \times m}$ ,  $C \in \mathbb{R}^{q \times n}$ ,  $m = 2$ ,  $q = 2$ ,
4. Define the frequency expansion point and the associated moment using `s0input = s0gen[m, q, s0]` with `s0 = [0; 30]`,
5. Define the numerical configuration `Opt`: e.g., `Opt.mtype = 'sparse'`, `Opt.orth = 'mgs'` and `Opt.lse = 'sLU'`,
6. With the preparation above, perform the reduction and obtain the reduced model  $\Sigma_r$  using `[~,~,sysr] = krylovMOR(E, {}, A, {}, B, {}, s0input, Opt)`,
7. Visualize and analyze the  $\Sigma$  and  $\Sigma_r$  in frequency domain.

We applied these steps for the model in Figure 8.2. The two inputs are the heat flux to the linear motor coil, and the two outputs are the temperature of NTC sensor 1 and NTC sensor 2, which are defined in Figure 4.4. Since the dynamics are predominately low-frequent, the frequency expansion point has been chosen as  $s_0 = 0$  with 30 orders. With this expansion point, the ROM preserves the low frequency of the FOM. The ROM has a state dimension of 30. The frequency response of the ROM are given in Figure 8.3 and Figure 8.4. The computation of the above steps are performed in a PC with 2.60GHz CPU and 8GB RAM, and the computation time is summarized below

Orthogonality error	CPU Time
$1.62 \times 10^{-16}$	0.92[s]

Second, the case study of LPV system based on the PSA setup using paramOR toolbox is given:

1. The FEM model is built in COMSOL Multiphysics with 38476 nodes,

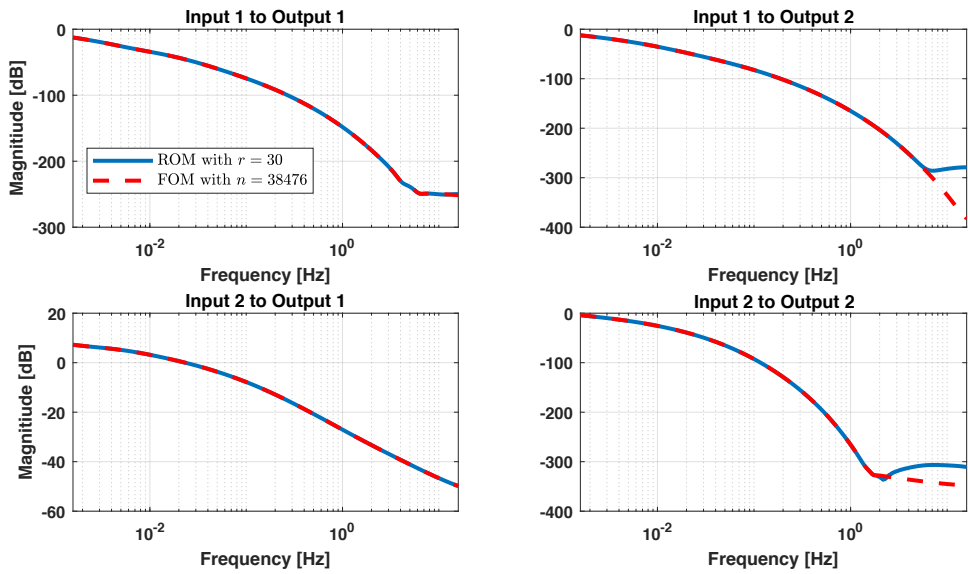


Figure 8.3: Magnitude plot of the ROM and the FOM with  $2 \times 2$  input and output.

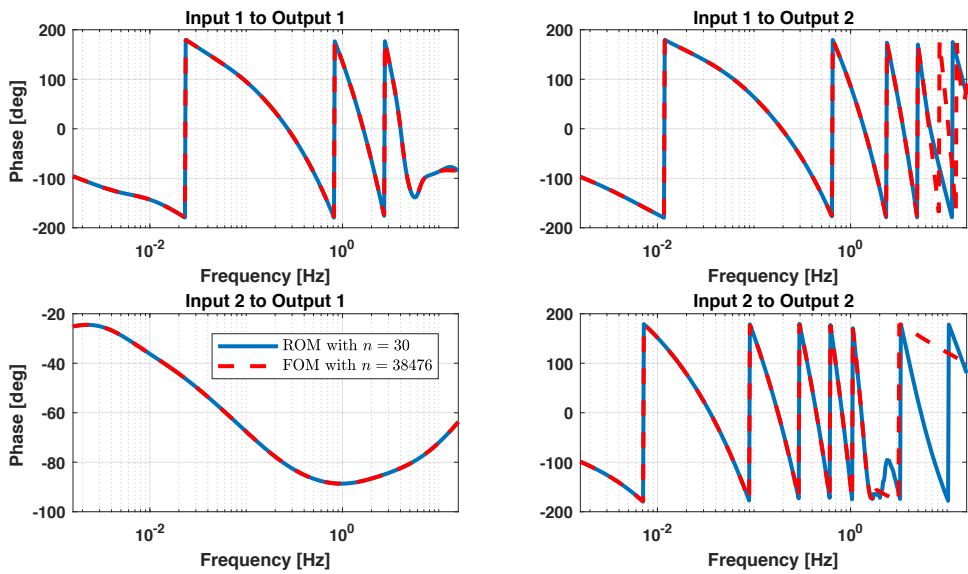


Figure 8.4: Phase plot of the ROM and the FOM with  $2 \times 2$  input and output.

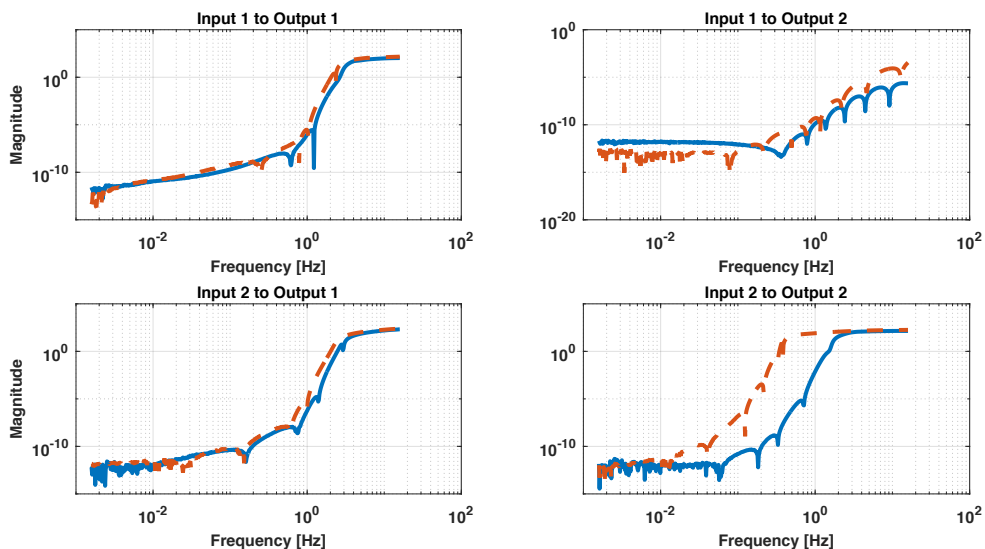
2. Define the heat capacitance of the PSA as the  $\theta_1$  to be parameterized or linearized,
3. Extract the LPV state space representation using `pssExtraction.m`,
4. Obtain the system matrices  $\Sigma = \{E(\theta), A, B, C\}$  with  $E(\theta) = E_0 + E_1\theta_1$ ,  $E_0, E_1, A \in \mathbb{R}^{n \times n}$ ,  $n = 38476$  and  $B \in \mathbb{R}^{n \times m}$ ,  $C \in \mathbb{R}^{q \times n}$ ,  $m = 2$ ,  $q = 2$  and  $\ell = 1$ ,
5. Define the frequency expansion points, parameter expansion points  $p$  and the associated moments using `ps0input = s0gen[m, q, ps0]`, the frequency expansion point is  $s_0 = 0$  with 100 moments, and the parameter expansion point is  $p = 475$  with 2 moments,
6. Define the numerical configuration Opt: e.g., `Opt.mtype = 'sparse'`, `Opt.orth = 'rmgs'` and `Opt.lse = 'sLU'`,
7. With the preparation above, perform the reduction and obtain the reduced model  $\Sigma_r$  using  
`[V, ~, psysr] = krylovMOR([E0, E1], p, A, {}, B, {}, ps0input, Opt),`
8. Construct the parameterized reduced order model using  $V$  for parameter at different values  $[0.5p, 2p]$ ,
9. Visualize and analyze the  $\{\Sigma(0.5p), \Sigma_r(0.5p)\}$  and  $\{\Sigma(2p), \Sigma_r(2p)\}$  in frequency domain.

The input and output configuration is the same as for the previous example: two inputs denote the heat flux to the two linear motor coils, and the outputs represent the temperature of two NTC sensors. The parameterized reduced model  $\Sigma_r(\theta)$  has the state number of 200, and the original full-order parameterized model  $\Sigma(\theta)$  has the state dimension of 38476.

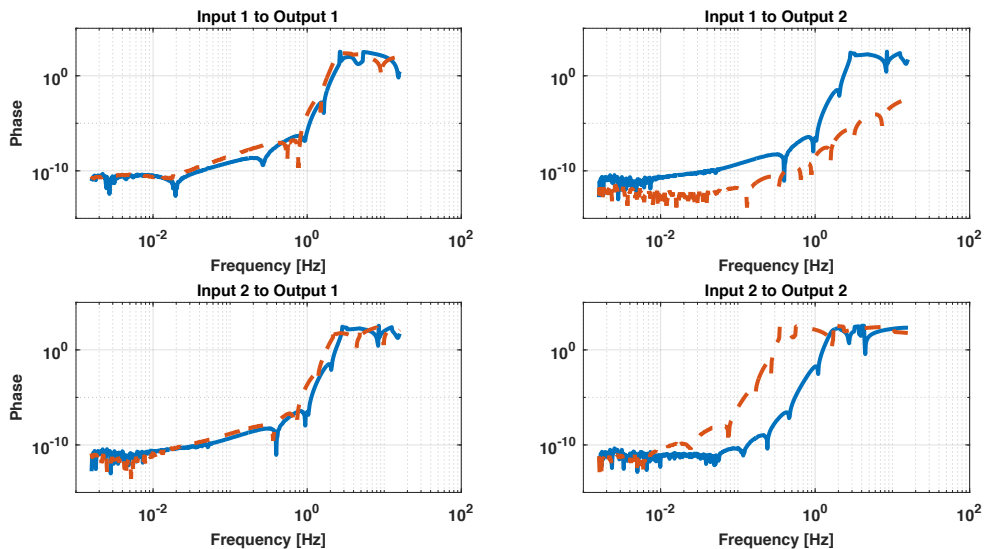
The magnitude and the phase plots between the FOM  $\Sigma(\theta)$  at the parameter  $\theta = 0.5p$  and  $\theta = 2p$ , the ROM  $\Sigma_r(\theta)$  at  $\theta = 0.5p$  and  $\theta = 2p$  are given in Figure 8.5 and Figure 8.6. The computation of the above steps are performed in a PC with 2.60GHz CPU and 8GB RAM, and the computation time is summarized below

Orthogonality error	CPU Time
$1.62 \times 10^{-13}$	3.67[s]

The blue lines in Figure 8.5 and Figure 8.6 show the difference(error) between the FOM  $\Sigma(\theta)$  and ROM  $\Sigma(\theta)$  at the  $\theta = 0.5p$  and the red lines in those two figures present the difference(error) between the FOM  $\Sigma(\theta)$  and ROM  $\Sigma(\theta)$  at  $\theta = 2p$ . These two figures show that both the magnitude and phase errors are relatively smaller at the lower frequency range. The error at both magnitude and phase plots is increasing at a higher frequency range. It is due to the choices of the expansion points. By adding higher frequency expansion points, we can further improve the ROM fidelity.



**Figure 8.5:** The red line denotes the magnitude error between the FOM  $\Sigma(2p)$  and the ROM  $\Sigma_r(2p)$ . The blue line represents the magnitude error between the FOM  $\Sigma(0.5p)$  and the ROM  $\Sigma_r(0.5p)$ .



**Figure 8.6:** The red line denotes the phase error between the FOM  $\Sigma(2p)$  and the ROM  $\Sigma_r(2p)$ . The blue line represents the phase error between the FOM  $\Sigma(0.5p)$  and the ROM  $\Sigma_r(0.5p)$ .

## 8.5 Conclusion

We presented the paraMOR toolbox as an efficient software solution for model order reduction for large-scale LTI and LPV systems. The toolbox facilitates the extraction of the affine parametric model directly from the FEM models that are used in the COMSOL or ANSYS software tools for the simulation of spatial-temporal and multi-physics systems. An overview of the toolbox and the structure of scripts are given, as well as the supported orthogonal methods and matrix decomposition methods. Two examples of both parameterized and non-parameterized systems are given to illustrate how paraMOR can be used to truncate the complexity. In the first example, paraMOR provides an efficient solution for the LTI system to visualize the results. In the second example, we used paraMOR to generate a projection matrix that can construct a parameterized reduced order model with different parameter values, and the error analysis is also attached.

The toolbox proves suitable for models of order up to  $n = 5 \times 10^6$  and gives computation times for the reduction process of the order of magnitude to minutes (e.g., a test model with state dimension of  $4.5 \times 10^6$  and 5 frequency expansion points and 10 parameter expansion points are chosen for the reduction, and it takes less than 10 [mins] to construct the projection matrices), depending on the number of expansion points, the optimization procedure and the requirements precision in the orthogonalization process. The toolbox has been applied to industrial applications via industrial collaborators.

# Conclusions and future research directions

---

## 9.1 Conclusions

In conclusion, this thesis has studied the analysis and the approximation of thermal systems. This is burdened by its typical large state and parameter dimensionality, together with the complexity of its inner interconnection structure. To solve this problem, we have tackled three sub-problems.

In the first part of the thesis, we have focused on the first subproblem, which entails the development of efficient model approximation methods for large-scale thermal systems. In Chapter 2, we have introduced a parameterized model order method that can construct physically relevant parameterized reduced-order models. Moreover, an a-priori error bound has been quantified for this method. This pMOR method has been successfully tested on industrial applications with state dimensions in the order of  $10^5$ . In addition, in Chapter 3, we also have explored the problem of parameter reduction. This helps to achieve higher precision requirements for which the dimension of the parameter space is also higher. By evaluating the system Gramian over parametric variation, we show that we can obtain a truncated parameter space where the approximation error is minimal. Two case studies have been carried out, and the results indicate the benefits of this approach. Chapter 4 leverages the pMOR method of Chapter 2 to achieve parameter calibration for large-scale thermal models. This enables parameter optimization for models for which this is time-consuming or even infeasible due to their heavy computational and memory loads. We have shown that the proposed scheme guarantees the equivalent first and second-order interpolation conditions for the cost function between the full-order model and the reduced-order model. In addition, this scheme is inherently adaptive to minimize the calibration error. More precisely, this algorithm finds the most sensitive frequency points for the parameters and automatically adapts the approximation of the full-order model

to it. This algorithm has been tested with a Precision Stage Application (PSA), and it has been shown to outperform the current conventional method. Building on the results of previous chapters, Chapter 5 has developed a complete methodology for LFT systems that combines the state and parameter approximation. We focused on the LTI systems and the uncertain LPV systems in LFT representation. For these systems, the proposed methodology successfully truncates the state dimension while preserving the LFT structure. Furthermore, we show that the state reduction and parameter reduction influence each other. In such a way, the parameter reduction can lead to further state reduction, while the fidelity of the reduced-order model remains the same. Overall, we have answered the sub-problem on the pMOR for large-scale thermal systems in Part I.

In the second part of the thesis, we targeted the system analysis for thermal systems. Chapter 6 investigated the stability analysis of these systems. We have shown the use of a Lyapunov function based on the irreversible entropy concept for stability analysis. On top of this, a conceptual distinction has been made between thermodynamic equilibrium and equilibria in control theory. Sequentially, Chapter 7 extended this Lyapunov function to analyze the dissipation properties of thermal systems, in which the boundary condition is relaxed from time-independence to time-dependence. We have illustrated with simulations that the Lyapunov function satisfies the dissipativity inequality.

The last part of this thesis presents a tool that we have developed as a MATLAB based software package for non-parametrized and parameterized model order reduction. Beyond the implementation of results delivered from Chapter 2 - Chapter 5, this toolbox contains functions that enable to start with model extraction, go to model reduction and end with analysis and visualizations.

## 9.2 Future research directions

In this work, we have solved three problems with the goal to facilitate system design, design optimization and calibration. Apart from these problems, several potential questions appeared during our journey that deserve further investigation.

Firstly, we would like to emphasize once more that moment-matching, as used in this thesis, is a local technique providing accurate models on user-defined bandwidths. More precisely, the ROM accuracy is inherently chosen for user-defined bandwidths. This provides the advantage of flexibility, but its local error bound lacks computability and it further misses a global error bound or other global properties. Therefore, Chapter 2 gives for now only a posterior local error bound in this user-defined frequency range. Recently, some preliminary work [102, 70] on error bounds for moment-matching based methods has been published. Therefore, future work may focus on finding a computable local a-prior error bound or even global a-prior error bound for large-scale models.

Secondly, Chapter 4 considers the parameter calibration problem, and we proposed an adaptive scheme to minimize the approximation error. In each iteration,



the new expansion point was found by solving the Jacobian matrix. This whole process may lead to a local minimum of the approximation error. Further investigation into finding the expansion points that provide a global minimum of the approximation can therefore improve the calibration results.

Thirdly, the definition of the moment used in this thesis is based on the transfer function of dynamical systems. Therefore, the moments preserved in the reduced model can reflect the essential dynamics of the original model. One possible future research direction can be to generalize the moments towards different system functions. For instance, one may define the moment for the impulse response of a nonlinear system. In such a way, the nonlinear dynamical systems can be approximated using the moment-matching method. Additionally, we also can further extend the model-based moment to a data-based moment, e.g., the data-drive model reduction technique Loewner framework [16] is a special case of a two-sided moment-matching procedure.

Furthermore, from the perspective of software implementation, it would be of particular interest to investigate the effect of industrial-approved tooling. The most common MOR tooling developed by academia is based on MATLAB, which imposes many limitations in exploiting the computational resources. In addition, parallel techniques have been widely used in many research domains while not much studied by the MOR community. One potential research direction is in the field of parallelization of MOR techniques using high-performance programming language. For instance, the paraMOR toolbox, which generates the projection matrix in Chapter 8, only utilizes one core of a multiprocess CPU, and the projection matrix is constructed in sequential order. Further research is required to find an efficient MOR algorithm such that each part of the projection matrix is computed in a separated processor, and all processors are deployed simultaneously.

As a final research direction, an investigation into energy-based control for thermo-relevant systems could be done. This would start from the results in Chapter 6 and Chapter 7, in which the description of the entropy generation function represents the system energy.



# Bibliography

- 
- [1] A. A. Alonso, C. V. Fernandez, and J. R. Banga. Dissipative systems: from physics to robust nonlinear control. *International Journal of Robust and Non-linear Control: IFAC-Affiliated Journal*, 14(2):157–179, 2004.
  - [2] R. A. Adams and C. Essex. *Calculus: a complete course*, volume 4. Addison-Wesley Boston, 1999.
  - [3] D. Akçay Perdahcioğlu, M. H. M. Ellenbroek, H. J. Geijselaers, and A. de Boer. Updating the craig–bampton reduction basis for efficient structural reanalysis. *International journal for numerical methods in engineering*, 85(5):607–624, 2011.
  - [4] A. A. Alonso and B. E. Ydstie. Process systems, passivity and the second law of thermodynamics. *Computers & chemical engineering*, 20:S1119–S1124, 1996.
  - [5] A. A. Alonso and B. E. Ydstie. Stabilization of distributed systems using irreversible thermodynamics. *Automatica*, 37(11):1739–1755, 2001.
  - [6] D. Amsallem, J. Cortial, and C. Farhat. On-demand cfd-based aeroelastic predictions using a database of reduced-order bases and models. In *47th AIAA Aerospace Sciences Meeting Including The New Horizons Forum and Aerospace Exposition*, page 800, 2009.
  - [7] D. Amsallem and C. Farhat. Interpolation method for adapting reduced-order models and application to aeroelasticity. *AIAA journal*, 46(7):1803–1813, 2008.
  - [8] D. Amsallem and C. Farhat. An online method for interpolating linear parametric reduced-order models. *SIAM Journal on Scientific Computing*, 33(5):2169–2198, 2011.
  - [9] D. Amsallem, C. Farhat, and T. Lieu. Aeroelastic analysis of f-16 and f-18/a configurations using adapted cfd-based reduced-order models. In *48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, page 2364, 2007.
  - [10] L. Andersson and A. Rantzer. Frequency-dependent error bounds for uncertain linear models. *IEEE Transactions on Automatic Control*, 44(11):2094–2098, 1999.

- [11] L. Andersson, A. Rantzer, and C. Beck. Model comparison and simplification. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 9(3):157–181, 1999.
- [12] A. C. Antoulas. On recursiveness and related topics in linear systems. *IEEE Transactions on Automatic Control*, 31(12):1121–1135, 1986.
- [13] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. SIAM, 2005.
- [14] A. C. Antoulas. An overview of approximation methods for large-scale dynamical systems. *Annual reviews in Control*, 29(2):181–190, 2005.
- [15] A. C. Antoulas, C. A. Beattie, and S. Gugercin. Interpolatory model reduction of large-scale dynamical systems. In *Efficient modeling and control of large-scale systems*, pages 3–58. Springer, 2010.
- [16] A. C. Antoulas, S. Lefteriu, A. C. Ionita, P. Benner, and A. Cohen. A tutorial introduction to the loewner framework for model reduction. *Model Reduction and Approximation: Theory and Algorithms*, 15:335, 2017.
- [17] A. C. Antoulas, D. C. Sorensen, and S. Gugercin. A survey of model reduction methods for large-scale systems. Technical report, 2000.
- [18] Z. Bai. Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. *Applied numerical mathematics*, 43(1-2):9–44, 2002.
- [19] J. Bao and P. L. Lee. *Process control: the passive systems approach*. Springer Science & Business Media, 2007.
- [20] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. An ‘empirical interpolation’ method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique*, 339(9):667–672, 2004.
- [21] U. Baur, C. Beattie, P. Benner, and S. Gugercin. Interpolatory projection methods for parameterized model reduction. *SIAM Journal on Scientific Computing*, 33(5):2489–2518, 2011.
- [22] U. Baur and P. Benner. Modellreduktion für parametrisierte systeme durch balanciertes abschneiden und interpolation. *Automatisierungstechnik*, 57(8):411–419, 2009.
- [23] U. Baur, P. Benner, A. Greiner, J. G. Korvink, J. Lienemann, and C. Moosmann. Parameter preserving model order reduction for mems applications. *Mathematical and Computer Modelling of Dynamical Systems*, 17(4):297–317, 2011.
- [24] C. A. Beattie and S. Gugercin. Model reduction by rational interpolation. *Model Reduction and Algorithms: Theory and Applications*, P. Benner, A. Cohen, M. Ohlberger, and K. Willcox, eds., *Comput. Sci. Engrg*, 15:297–334, 2014.
- [25] C. Beck. Coprime factors reduction methods for linear parameter varying and uncertain systems. *Systems & control letters*, 55(3):199–213, 2006.

- [26] C. L. Beck, J. Doyle, and K. Glover. Model reduction of multidimensional and uncertain systems. *IEEE Transactions on Automatic Control*, 41(10):1466–1477, 1996.
- [27] A. Bejan. Entropy generation minimization: The new thermodynamics of finite-size devices and finite-time processes. *Journal of Applied Physics*, 79(3):1191–1218, 1996.
- [28] P. Benner and L. H. Feng. A robust algorithm for parametric model order reduction based on implicit moment matching. In A. Quarteroni and G. Rozza, editors, *Reduced Order Methods for Modeling and Computational Reduction*, chapter 6, pages 159–186. Springer International Publishing Switzerland, 9 edition, 2014.
- [29] P. Benner, S. Gugercin, and K. Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM review*, 57(4):483–531, 2015.
- [30] P. Benner, E. S. Quintana-Ortí, and G. Quintana-Ortí. Balanced truncation model reduction of large-scale dense systems on parallel computers. *Mathematical and Computer Modelling of Dynamical Systems*, 6(4):383–405, 2000.
- [31] P. Benner and S. W. R. Werner. MORLAB – Model Order Reduction LABORatory (version 5.0), Aug. 2019. see also: <http://www.mpi-magdeburg.mpg.de/projects/morlab>.
- [32] D. S. Bernstein and D. C. Hyland. Compartmental modeling and second-moment analysis of state space systems. *SIAM journal on matrix analysis and applications*, 14(3):880–901, 1993.
- [33] B. Besselink, U. Tabak, A. Lutowska, N. Van de Wouw, H. Nijmeijer, D. J. Rixen, M. Hochstenbach, and W. Schilders. A comparison of model reduction techniques from structural dynamics, numerical mathematics and systems and control. *Journal of Sound and Vibration*, 332(19):4403–4422, 2013.
- [34] R. Bhatia. *Matrix analysis*, volume 169. Springer Science & Business Media, 2013.
- [35] C. Bickora and S. Weiland. Parameter-oriented reduction strategies for a thermo-mechanical model of extreme ultraviolet reticles. In *2016 IEEE Conference on Control Applications (CCA)*, pages 1149–1154. IEEE, 2016.
- [36] D. L. Boley. Krylov space methods on state-space control models. *Circuits, Systems, and Signal Processing*, 13(6):733–758, 1994.
- [37] B. Bond and L. Daniel. Parameterized model order reduction of nonlinear dynamical systems. In *ICCAD-2005. IEEE/ACM International Conference on Computer-Aided Design, 2005.*, pages 487–494. IEEE, 2005.
- [38] J. Borggaard, K. R. Pond, and L. Zietsman. Parametric reduced order models using adaptive sampling and interpolation. *IFAC Proceedings Volumes*, 47(3):7773–7778, 2014.

- [39] G. E. Box and N. R. Draper. *Empirical model-building and response surfaces*. John Wiley & Sons, 1987.
- [40] S. Brunsch. *Parametric Model Order Reduction Using Sparse Grids*. PhD thesis, Universitäts- und Landesbibliothek Bonn, 2017.
- [41] J. Bryan. International status of thermal error research (1990). *CIRP annals*, 39(2):645–656, 1990.
- [42] T. Bui-Thanh, K. E. Willcox, and O. Ghattas. Model reduction for large-scale systems with high-dimensional parametric input space. *SIAM Journal on Scientific Computing*, 30(6):3270–3288, 2008.
- [43] S. Burgard, O. Farle, and R. Dyczij-Edlinger. An  $h$  adaptive sub-domain framework for parametric order reduction. *IEEE Transactions on Magnetics*, 51(3):1–4, 2015.
- [44] T. Burke. Advanced thermal control consortium. <http://www.ibspe.com/page/new-consortium-in-advanced-thermal-control.htm/>, 2016.
- [45] O. Burkovska, K. Glau, M. Mahlstedt, and B. Wohlmuth. Model reduction for calibration of american options. *arXiv preprint arXiv:1611.06452*, 2016.
- [46] H. B. Callen. *Thermodynamics and an introduction to thermostatistics*, 1998.
- [47] K. Carlberg, C. Bou-Mosleh, and C. Farhat. Efficient non-linear model reduction via a least-squares petrov–galerkin projection and compressive tensor approximations. *International Journal for numerical methods in engineering*, 86(2):155–181, 2011.
- [48] K. Carlberg, C. Farhat, J. Cortial, and D. Amsallem. The gnat method for nonlinear model reduction: effective implementation and application to computational fluid dynamics and turbulent flows. *Journal of Computational Physics*, 242:623–647, 2013.
- [49] K. Carlberg, V. Forstall, and R. Tuminaro. Krylov-subspace recycling via the pod-augmented conjugate-gradient method. *SIAM Journal on Matrix Analysis and Applications*, 37(3):1304–1336, 2016.
- [50] A. Castagnotto, C. Beattie, and S. Gugercin. Interpolatory methods for  $\mathcal{H}_\infty$  model reduction of multi-input/multi-output systems. In *Model Reduction of Parametrized Systems*, pages 349–365. Springer, 2017.
- [51] A. Castagnotto, M. Cruz Varona, and B. Lohmann. sss & sssmor-analysis and reduction of large-scale dynamic systems in matlab. In *KoMSO Challenge Workshop on Reduced-Order Modeling for Simulation and Optimization*, 2016.
- [52] S. Chaturantabut and D. C. Sorensen. Nonlinear model reduction via discrete empirical interpolation. *SIAM Journal on Scientific Computing*, 32(5):2737–2764, 2010.

- [53] L. Daniel, O. C. Siong, L. S. Chay, K. H. Lee, and J. White. A multiparameter moment-matching model-reduction approach for generating geometrically parameterized interconnect performance models. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 23(5):678–693, 2004.
- [54] S. De Groot and P. Mazur. Nonequilibrium thermodynamics, 1sted, 1962.
- [55] S. R. De Groot. Thermodynamics of irreversible processes. 1952.
- [56] J. Degroote, J. Vierendeels, and K. 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.
- [57] J.-C. Delvenne and H. Sandberg. Finite-time thermodynamics of port-hamiltonian systems. *Physica D: Nonlinear Phenomena*, 267:123–132, 2014.
- [58] J.-C. Delvenne and H. Sandberg. Deriving thermodynamics from linear dissipativity theory. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 537–542. IEEE, 2015.
- [59] J.-C. Delvenne and H. Sandberg. Dissipative open systems theory as a foundation for the thermodynamics of linear systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375(2088):20160218, 2017.
- [60] Y. Demirel and S. I. Sandler. Nonequilibrium thermodynamics in engineering and science. *The Journal of Physical Chemistry B*, 108(1):31–43, 2004.
- [61] K. Denbigh. The thermodynamics of the steady state, methuen & co. Ltd., London, 84, 1951.
- [62] M. Dettori and C. W. Scherer. LPV design for a CD player : an experimental evaluation of performance. *Selected Topics in Signals, Systems and Control*, 12(September), 2001.
- [63] V. Duindam, A. Macchelli, S. Stramigioli, and H. Bruyninckx. *Modeling and control of complex physical systems: the port-Hamiltonian approach*. Springer Science & Business Media, 2009.
- [64] R. Eid, B. Salimbahrami, B. Lohmann, E. B. Rudnyi, and J. G. Korvink. Parametric order reduction of proportionally damped second order systems. Technical report, Lehrstuhl für Regelungstechnik, 2006.
- [65] L. C. Evans. *Partial differential equations*, volume 19. American Mathematical Soc., 2010.
- [66] E. Evers. *Identification and active thermomechanical control in precision mechatronics*. PhD thesis, Mechanical Engineering, Jan. 2021. Proefschrift.

- [67] P. Feldmann and R. W. Freund. Efficient linear circuit analysis by padé approximation via the lanczos process. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 14(5):639–649, 1995.
- [68] P. Feldmann and R. W. Freund. Reduced-order modeling of large linear subcircuits via a block lanczos algorithm. In *Proceedings of the 32nd annual ACM/IEEE Design Automation Conference*, pages 474–479. ACM, 1995.
- [69] L. Feng. Parameter independent model order reduction. *Mathematics and Computers in Simulation*, 68(3):221–234, 2005.
- [70] L. Feng, A. C. Antoulas, and P. Benner. Some a posteriori error bounds for reduced-order modelling of (non-) parametrized linear systems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 51(6):2127–2158, 2017.
- [71] L. H. Feng, E. B. Rudnyi, and J. G. Korvink. Preserving the film coefficient as a parameter in the compact thermal model for fast electrothermal simulation. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 24(12):1838–1847, 2005.
- [72] Z. Feng and P. Li. Performance-Oriented Statistical Parameter Reduction of Parameterized Systems via Reduced Rank Regression. In *2006 IEEE/ACM International Conference on Computer Aided Design*, pages 868–875, 2006.
- [73] D. Fitts. Nonequilibrium thermodynamics, a phenomenological theory of irreversible processes in fluid systems, mcgraw-hill, new york. 1962.
- [74] M. Frangos and I. M. Jaimoukha. Adaptive rational interpolation: Arnoldi and lanczos-like equations. 2008.
- [75] R. W. Freund. Model reduction methods based on krylov subspaces. *Acta Numerica*, 12:267–319, 2003.
- [76] J. P. García-Sandoval, V. González-Álvarez, and C. Calderón. Stability analysis and passivity properties for a class of chemical reactors: Internal entropy production approach. *Computers & chemical engineering*, 75:184–195, 2015.
- [77] L. Giraud, J. Langou, and M. Rozloznik. The loss of orthogonality in the gram-schmidt orthogonalization process. *Computers & Mathematics with Applications*, 50(7):1069–1075, 2005.
- [78] G. Gouy. Sur les transformation et l'équilibre en thermodynamique. *CR Acad. Sci. Paris*, 108(10):507–509, 1889.
- [79] G. Gouy et al. Sur l'énergie utilisable. *Journal de physique*, 8(2nd edition):501, 1889.
- [80] M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 41(3):575–605, 2007.



- [81] E. J. Grimme. *Krylov projection methods for model reduction*. PhD thesis, University of Illinois at Urbana-Champaign Urbana-Champaign, IL, 1997.
- [82] S. Gugercin and A. C. Antoulas. A survey of model reduction by balanced truncation and some new results. *International Journal of Control*, 77(8):748–766, 2004.
- [83] S. Gugercin and A. C. Antoulas. Model reduction of large-scale systems by least squares. *Linear algebra and its applications*, 415(2-3):290–321, 2006.
- [84] S. Gugercin, A. C. Antoulas, and C. Beattie.  $\mathcal{H}_2$  Model Reduction for Large-Scale Linear Dynamical Systems. *SIAM Journal on Matrix Analysis and Applications*, June 2008.
- [85] S. Gugercin, C. Beattie, and A. Antoulas. Rational krylov methods for optimal h2 model reduction. *submitted for publication*, 2006.
- [86] P. Gunupudi, R. Khazaka, and M. Nakhla. Analysis of transmission line circuits using multidimensional model reduction techniques. *IEEE Transactions on Advanced Packaging*, 25(2):174–180, 2002.
- [87] P. K. Gunupudi, R. Khazaka, M. S. Nakhla, T. Smy, and D. Celo. Passive parameterized time-domain macromodels for high-speed transmission-line networks. *IEEE Transactions on Microwave Theory and Techniques*, 51(12):2347–2354, 2003.
- [88] B. Haasdonk. Reduced basis methods for parametrized pdes—a tutorial introduction for stationary and instationary problems. *Model reduction and approximation: theory and algorithms*, 15:65, 2017.
- [89] B. Haasdonk, M. Ohlberger, and G. Rozza. A reduced basis method for evolution schemes with parameter-dependent explicit operators. *ETNA, Electronic Transactions on Numerical Analysis*, 32(ARTICLE):145–168, 2008.
- [90] W. M. Haddad, V. Chellaboina, and S. G. Nersisov. *Thermodynamics: A dynamical systems approach*, volume 23. Princeton University Press, 2009.
- [91] J. S. Han, E. B. Rudnyi, and J. G. Korvink. Efficient optimization of transient dynamic problems in mems devices using model order reduction. *Journal of Micromechanics and Microengineering*, 15(4):822, 2005.
- [92] S.-A. Hauschild, N. Marheineke, V. Mehrmann, J. Mohring, A. M. Badlyan, M. Rein, and M. Schmidt. Port-hamiltonian modeling of district heating networks. In *Progress in Differential-Algebraic Equations II*, pages 333–355. Springer, 2020.
- [93] C. Himpe and M. Ohlberger. A unified software framework for empirical gramians. *Journal of Mathematics*, 2013, 2013.
- [94] C. Himpe and M. Ohlberger. Cross-gramian-based combined state and parameter reduction for large-scale control systems. *Mathematical Problems in Engineering*, 2014, 2014.

- [95] H. Hoang, F. Couenne, Y. Le Gorrec, and D. Dochain. Thermodynamics based stabiliztion of cstr networks. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 6352–6357. IEEE, 2012.
- [96] M. Huang, P. Rivera-Díaz-del Castillo, O. Bouaziz, and S. Van Der Zwaag. Irreversible thermodynamics modelling of plastic deformation of metals. *Materials Science and Technology*, 24(4):495–500, 2008.
- [97] L. Iapichino, A. Quarteroni, and G. Rozza. A reduced basis hybrid method for the coupling of parametrized domains represented by fluidic networks. *Computer Methods in Applied Mechanics and Engineering*, 221:63–82, 2012.
- [98] L. Iapichino, A. Quarteroni, and G. Rozza. Reduced basis method and domain decomposition for elliptic problems in networks and complex parametrized geometries. *Computers & Mathematics with Applications*, 71(1):408–430, 2016.
- [99] K. Ito and S. Ravindran. A reduced basis method for control problems governed by pdes. In *Control and estimation of distributed parameter systems*, pages 153–168. Springer, 1998.
- [100] J. Keizer. Thermodynamics at nonequilibrium steady states. *The Journal of Chemical Physics*, 69(6):2609–2620, 1978.
- [101] H. K. Khalil. Nonlinear systems. *Upper Saddle River*, 2002.
- [102] M. A. Khatkhat, M. I. Ahmad, L. Feng, and P. Benner. Multivariate moment matching for model order reduction of quadratic-bilinear systems using error bounds. *arXiv preprint arXiv:2105.12966*, 2021.
- [103] S. Kjelstrup, D. Bedeaux, E. Johannessen, and J. Gross. *Non-equilibrium thermodynamics for engineers*. World Scientific, 2010.
- [104] D. Kondepudi and I. Prigogine. *Modern thermodynamics: from heat engines to dissipative structures*. John Wiley & Sons, 2014.
- [105] P. Koutsovasilis and M. Beitelschmidt. Model order reduction package for coupling rigid and elastic multibody dynamics. In *NAFEMS World Congress*, 2009.
- [106] G. D. Kuiken and G. D. Kuiken. *Thermodynamics of irreversible processes: applications to diffusion and rheology*, volume 49. Wiley New York, 1994.
- [107] P. T. Landsberg. Foundations of thermodynamics. *Reviews of modern physics*, 28(4):363, 1956.
- [108] Y. Le Gorrec, B. Maschke, J. Villegas, and H. Zwart. Dissipative boundary control systems with application to distributed parameters reactors. In *2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control*, pages 668–673. IEEE, 2006.

- [109] Y. Li, W. Zhao, S. Lan, J. Ni, W. Wu, and B. Lu. A review on spindle thermal error compensation in machine tools. *International Journal of Machine Tools and Manufacture*, 95:20–38, 2015.
- [110] C. Lieberman, K. Willcox, and O. Ghattas. Parameter and state model reduction for large-scale statistical inverse problems. *SIAM Journal on Scientific Computing*, 32(5):2523–2542, 2010.
- [111] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, pages 284–289, Taipei, 2004.
- [112] B. Lohmann and R. Eid. Efficient order reduction of parametric and nonlinear models by superposition of locally reduced models. In *Methoden und Anwendungen der Regelungstechnik. Erlangen-Münchener Workshops*, pages 27–36, 2007.
- [113] D. Lou and S. Weiland. Parametric model order reduction for large-scale and complex thermal systems. In *2018 European Control Conference (ECC)*, pages 2593–2598. IEEE, 2018.
- [114] D. Lou and S. Weiland. Stability analysis of thermodynamic systems: Heat conduction in solids. In *21st IFAC World Congress*, 2020.
- [115] R. H. Lyon. Statistical energy analysis of dynamical systems. *Theory and Applications*, 1975.
- [116] A. Marcos and G. J. Balas. Development of Linear Parameter Varying Models for Aircraft. *Journal of Guidance, Control and Dynamics*, 27(2):218–228, 2004.
- [117] B. Maschke, R. Ortega, and A. J. van der Schaft. Energy-based lyapunov functions for forced hamiltonian systems with dissipation. *IEEE Transactions on automatic control*, 45(8):1498–1502, 2000.
- [118] B. Maschke and A. J. van der Schaft. A hamiltonian approach to stabilization of nonholonomic mechanical systems. In *Proceedings of 1994 33rd IEEE Conference on Decision and Control*, volume 3, pages 2950–2954. IEEE, 1994.
- [119] B. Maschke and A. J. van der Schaft. Port-thermodynamic systems and the assignment of their structure by feedback. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 8067–8072. IEEE, 2019.
- [120] G. Masson, B. A. Brik, S. Cogan, and N. Bouhaddi. Component mode synthesis (cms) based on an enriched ritz approach for efficient structural optimization. *Journal of sound and vibration*, 296(4-5):845–860, 2006.
- [121] J. Mockus, V. Tiesis, and A. Zilinskas. The application of bayesian methods for seeking the extremum. *Towards global optimization*, 2(117-129):2, 1978.
- [122] M. C. Muñoz-Lecanda and F. J. Yániz-Fernández. Dissipative control of mechanical systems: a geometric approach. *SIAM journal on control and optimization*, 40(5):1505–1516, 2002.

- [123] F. Negri, A. Manzoni, and D. Amsallem. Efficient model reduction of parametrized systems by matrix discrete empirical interpolation. *Journal of Computational Physics*, 303:431–454, 2015.
- [124] C. Oguz, T. Laomettachit, K. C. Chen, L. T. Watson, W. T. Baumann, and J. J. Tyson. Optimization and model reduction in the high dimensional parameter space of a budding yeast cell cycle model. *BMC systems biology*, 7(1):53, 2013.
- [125] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on automatic control*, 49(9):1520–1533, 2004.
- [126] L. Onsager. Reciprocal relations in irreversible processes. i. *Physical review*, 37(4):405, 1931.
- [127] L. Onsager. Reciprocal relations in irreversible processes. ii. *Physical review*, 38(12):2265, 1931.
- [128] R. Ortega, D. Jeltsema, and J. M. Scherpen. Power shaping: A new paradigm for stabilization of nonlinear rlc circuits. *IEEE Transactions on Automatic Control*, 48(10):1762–1767, 2003.
- [129] M. Ortiz and A. Pandolfi. Finite-deformation irreversible cohesive elements for three-dimensional crack-propagation analysis. *International journal for numerical methods in engineering*, 44(9):1267–1282, 1999.
- [130] H. Panzer, J. Mohring, R. Eid, and B. Lohmann. Parametric model order reduction by matrix interpolation. *at-Automatisierungstechnik Methoden und Anwendungen der Steuerungs-, Regelungs-und Informationstechnik*, 58(8):475–484, 2010.
- [131] V. Papakos and I. M. Jaimoukha. Model reduction via an lft-based explicitly restarted nonsymmetric lanczos algorithm. *MTNS-02*, 2002.
- [132] A. Paul-Dubois-Taine and D. Amsallem. An adaptive and efficient greedy procedure for the optimal training of parametric reduced-order models. *International Journal for Numerical Methods in Engineering*, 102(5):1262–1292, 2015.
- [133] R. K. Pearson and T. L. Johnson. Energy equipartition and fluctuation-dissipation theorems for damped flexible structures. *Quarterly of applied mathematics*, 45(2):223–238, 1987.
- [134] B. Peherstorfer, D. Butnaru, K. Willcox, and H.-J. Bungartz. Localized discrete empirical interpolation method. *SIAM Journal on Scientific Computing*, 36(1):A168–A192, 2014.
- [135] D. Peumans, C. Busschots, G. Vandersteen, and R. Pintelon. Improved frf measurements of lightly damped systems using local rational models. *IEEE Transactions on Instrumentation and Measurement*, 67(7):1749–1759, 2018.

- [136] L. T. Pillage and R. A. Rohrer. Asymptotic waveform evaluation for timing analysis. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 9(4):352–366, 1990.
- [137] R. Pintelon and J. Schoukens. *System identification: a frequency domain approach*. John Wiley & Sons, 2012.
- [138] C. Poussot-Vassal and C. Roos. Generation of a reduced-order lpv/lft model from a set of large-scale mimo lti flexible aircraft models. *Control Engineering Practice*, 20(9):919–930, 2012.
- [139] I. Prigogine. *Étude thermodynamique des phénomènes irréversibles*. 1947.
- [140] S. Z. Rizvi, J. Mohammadpour, R. Tóth, and N. Meskin. A kernel-based PCA approach to model reduction of linear parameter-varying systems. *IEEE Transactions on Control Systems Technology*, 24(5):1883–1891, 2016.
- [141] J. Rommes. *Methods for eigenvalue problems with applications in model order reduction*. PhD thesis, 2007.
- [142] A. Romo Hernandez. *A thermodynamic approach to modeling and stability analysis of multiphase systems*. PhD thesis, UCL-Université Catholique de Louvain, 2019.
- [143] G. Rozza, D. B. P. Huynh, and A. T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, 15(3):1, 2007.
- [144] J. Saak, M. Köhler, and P. Benner. M-M.E.S.S.-2.0.1 – the matrix equations sparse solvers library. DOI:10.5281/zenodo.3606345, Feb. 2020. see also: [www.mpi-magdeburg.mpg.de/projects/mess](http://www.mpi-magdeburg.mpg.de/projects/mess).
- [145] S. I. Sandler. *Chemical, biochemical, and engineering thermodynamics*. John Wiley & Sons, 2017.
- [146] G. Scarcioffi and A. Astolfi. Nonlinear model reduction by moment matching. 2017.
- [147] C. Scherer. Theory of robust control. *Delft University of Technology*, pages 1–160, 2001.
- [148] C. Scherer and S. Weiland. Linear matrix inequalities in control. *Lecture Notes, Dutch Institute for Systems and Control, Delft, The Netherlands*, 3(2), 2000.
- [149] W. H. Schilders, H. A. Van der Vorst, and J. Rommes. *Model order reduction: theory, research aspects and applications*, volume 13. Springer, 2008.
- [150] S. Schouten, D. Lou, and S. Weiland. Model reduction for linear parameter-varying systems through parameter projection. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 7800–7805. IEEE, 2019.

- [151] H. Schwenke, W. Knapp, H. Haitjema, A. Weckenmann, R. Schmitt, and F. Delbressine. Geometric error measurement and compensation of machines-an update. *CIRP annals*, 57(2):660–675, 2008.
- [152] L. M. Silveira, M. Kamon, and J. White. Efficient reduced-order modeling of frequency-dependent coupling inductances associated with 3-d interconnect structures. *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part B*, 19(2):283–288, 1996.
- [153] J. M. Smith. Introduction to chemical engineering thermodynamics, 1950.
- [154] N. T. Son. A real time procedure for affinely dependent parametric model order reduction using interpolation on grassmann manifolds. *International Journal for Numerical Methods in Engineering*, 93(8):818–833, 2013.
- [155] N. T. Son, P.-Y. Gousenbourger, E. Massart, and T. Stykel. Balanced truncation for parametric linear systems using interpolation of gramians: a comparison of algebraic and geometric approaches. *arXiv preprint arXiv:2003.04577*, 2020.
- [156] N. T. Son and T. Stykel. Solving parameter-dependent lyapunov equations using the reduced basis method with application to parametric model order reduction. *SIAM Journal on Matrix Analysis and Applications*, 38(2):478–504, 2017.
- [157] A. Stodola. *Steam and gas turbines: with a supplement on the prospects of the thermal prime mover*, volume 2. McGraw-Hill, 1927.
- [158] C. Sun and J. Hahn. Parameter reduction for stable dynamical systems based on Hankel singular values and sensitivity analysis. *Chemical Engineering Science*, 61(16):5393–5403, 2006.
- [159] P. M. J. Van den Hof, J. F. Van Doren, and S. G. Douma. Identification of parameters in large scale physical model structures, for the purpose of model-based operations. In *Model-Based Control*., pages 125–143. Springer, 2009.
- [160] A. J. van der Schaft and D. Jeltsema. Port-hamiltonian systems theory: An introductory overview. *Foundations and Trends in Systems and Control*, 1(2-3):173–378, 2014.
- [161] A. J. van der Schaft and B. Maschke. About some system-theoretic properties of port-thermodynamic systems. In *International Conference on Geometric Science of Information*, pages 228–238. Springer, 2019.
- [162] A. J. van der Schaft and B. M. Maschke. On the hamiltonian formulation of nonholonomic mechanical systems. *Reports on mathematical physics*, 34(2):225–233, 1994.
- [163] A. J. van der Schaft and B. M. Maschke. Hamiltonian formulation of distributed-parameter systems with boundary energy flow. *Journal of Geometry and physics*, 42(1-2):166–194, 2002.

- [164] M. C. Varona, B. Lohmann, et al. Automatic adaptive sampling in parametric model order reduction by matrix interpolation. In *2017 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 472–477. IEEE, 2017.
- [165] B. Vincent, N. Hudon, L. Lefèvre, and D. Dochain. Passivity and stability properties of multi-physics systems using the entropy production. *IFAC-PapersOnLine*, 51(3):13–18, 2018.
- [166] S. Walton, O. Hassan, and K. Morgan. Reduced order modelling for unsteady fluid flow using proper orthogonal decomposition and radial basis functions. *Applied Mathematical Modelling*, 37(20-21):8930–8945, 2013.
- [167] J. M. Wang, O. A. Hafiz, and J. Li. A linear fractional transform (lft) based model for interconnect parametric uncertainty. In *Proceedings of the 41st annual Design Automation Conference*, pages 375–380, 2004.
- [168] W. Wang, J. Doyle, C. Beck, and K. Glover. Model reduction of lft systems. 1991.
- [169] S. Weiland and J. C. WILLEMS. Dissipative dynamical systems in a behavioral context. *Mathematical Models and Methods in Applied Sciences*, 1(01):1–25, 1991.
- [170] D. Weile, E. Michielssen, E. Grimme, and K. Gallivan. A method for generating rational interpolant reduced order models of two-parameter linear systems. *Applied Mathematics Letters*, 12(5):93–102, 1999.
- [171] J. C. Willems. Dissipative dynamical systems part i: General theory. *Archive for rational mechanics and analysis*, 45(5):321–351, 1972.
- [172] J. C. Willems. Dissipative dynamical systems part ii: Linear systems with quadratic supply rates. *Archive for rational mechanics and analysis*, 45(5):352–393, 1972.
- [173] P. Wittmuess, C. Tarin, A. Keck, E. Arnold, and O. Sawodny. Parametric model order reduction via balanced truncation with taylor series representation. *IEEE Transactions on Automatic Control*, 61(11):3438–3451, 2016.
- [174] G. D. Wood. *Control of Parameter-Dependent Mechanical Systems*. Phd, Cambridge University, 1995.
- [175] B. E. Ydstie and A. A. Alonso. Process systems and passivity via the clausius-planck inequality. *Systems & control letters*, 30(5):253–264, 1997.
- [176] Y. Yue and K. Meerbergen. Accelerating optimization of parametric linear systems by model order reduction. *SIAM Journal on Optimization*, 23(2):1344–1370, 2013.
- [177] Y. Yue and K. Meerbergen. Parametric model order reduction of damped mechanical systems via the block arnoldi process. *Applied Mathematics Letters*, 26(6):643–648, 2013.

- [178] R. Zhou, B. Gressick, J. T. Wen, M. Jensen, J. Frankel, G. Lerner, and M. Unrath. Active thermal management for precision positioning. In *2007 IEEE International Conference on Automation Science and Engineering*, pages 45–50. IEEE, 2007.
- [179] R. Zimmermann. A locally parametrized reduced-order model for the linear frequency domain approach to time-accurate computational fluid dynamics. *SIAM Journal on Scientific Computing*, 36(3):B508–B537, 2014.



# Acknowledgments

---

During the journey of this thesis, I received much appreciated support from many people and partners who contributed to this research and my life. This work would not have been possible without their support.

First of all, I would like to express my sincere gratitude to my supervisor Prof. Siep Weiland for his thoughtful advice and constant guidance that he has given me in the past six years. After being the supervisor of my master's thesis, he continued to be the promotor of my Ph.D. thesis for another five years. His enthusiasm, vision and motivation have deeply inspired me. He has taught me the clarity of thought over a scientific statement. I still remember our spirited discussion on the thermodynamic equilibrium while writing over the whole whiteboard. It was a great pleasure to be his student. I extend my appreciation to my co-promotor dr. Leyla Ozkan. She has always motivated me to stay on course and maintain focus on the essence, especially during the age of turbulence. I also would like to thank Will Hendrix for his effort that he has spent on the hardware and the software implementation of the setup.

I greatly appreciate my Ph.D. committee members: prof.dr.ir. Nathan van de Wouw, prof.dr. Arjan van der Schaft, prof. dr.-Ing. Jan Lunze, dr.ir. Marc van de Wal and dr. Jos Maubach. I highly appreciate the time and effort they put into evaluating the quality of my thesis, and I am thankful for their valuable feedbacks.

The research presented in this dissertation was supported by the Advanced Thermal Control Consortium, which is greatly appreciated. The committee of the ATC provided tremendous technical support and user applications to this research: Wim Symens, Wouter Aangenent, Ronald Lamers, Jack van der Sanden, Nic Dirks, Dennis Heck, Henny Spaan. Rob van Gils, thanks for sharing the precision stage application with us and assisting the knowledge transfer session. Theresa Burke, I appreciate your great efforts on this consortium and provided us with the opportunity to promote our work in the VDMA, Germany. A special word of thanks goes to Marc van de Wal, he has provided continuous support, insightful discussion and comprehensive feedback for the thesis. Additionally, he has helped to organize the knowledge transfer event. I also would like to thank another two joint partners in the project, Enzo and Max, I very much appreciate the time we spent in the annual events, Frankfurt and the discussions we had.

My time as a Ph.D. student has been really enjoyable, I would like to thank my friends and colleagues at the CS group over the past years. Special thanks go to my officemate: Bep, Paul, Feye, Bahadir, Henrik, Zuan, David, Tuan, Amritam, for creating such a wonderful working atmosphere. I want to thank Dhruv

for sharing a lot of experiences and memories since the master program. David, who was also there when I need some advice and suggestion. Tuan, I really enjoy our conversation and our gym hours. Xiaodong, I appreciate our scientific debates and brainstorming, which broaden my vision on model order reduction. Maarten, I am grateful for your support and enjoy our discussions. The time we spent in the coffee-break-conversation, conference, workshops, and dinner is a precious memory for me. Thanks Harm, Ruben, Giulio, Pepijn, Tom, Karthik, Shengling, Zhiyong, Clarissa, Carlos, Alejandro, Marcella, Constantijn, Alina, Ruxandra. Special thanks goes to Diana, Hiltje and Barbara, I am grateful for their kindness and supports during my stay in the group. I would like to thank Paul for his support and effort to provide such a great working environment. I have been fortunate and proud to be one of the CS group.

Since the day I arrived in the Netherlands, I have been lucky to have lots of great friends. In Guannan, I found a reliable source of emotional support and valuable food recommendations. I would like to thank Guibin for his long-standing friendship, and even he is back in China. Michelle, impressed by her strength and thanks for the time we have spent. I extend my heartfelt thanks for the companionship of my Chinese friends here: Li Yue, Baobao, Kevin, Elly, Alex Zhao, Enzo, Wang shixiong, Yifei, Zhangcheng, Stephen. I have been very much enjoyed the basketball time with them.

My life in the Netherlands could not be so delightful without the support and love from my family-in-law: Dirk, Carla, Stefanie and Maarten. They have provided me with a warm family. Special thanks goes to Carla who provided the great design of the cover page.

I would like to express my great gratitude to my parents, my sister, brother-in-law, and my lovely niece. Your unconditional support and love have always been my source of strength.

Above all I would like to thank my wife Sofie for her love and constant support, for keeping me sane over the past year. She has stood by me through all my travails and my impatience. She gave me support and help, discussed ideas and provided valuable suggestions.

*Daming  
Nuenen, 2021*

# Curriculum Vitae

Daming Lou was born in Hangzhou, China, in 1989. He received his B.Sc. degree in Electrical Engineering from the Shanghai Maritime University with distinction in 2012, China. From 2013 to 2016, he followed the M.Sc program in Electrical Engineering at the Eindhoven University of Technology (TU/e), the Netherlands. He completed his M.Sc thesis, titled "Closed-loop controller design for a 3D printer using model-based control".

From June 2016, he started his PhD project at the Control Systems Group of TU/e, under the supervision of prof. dr. Siep Weiland, on the research topic "*next generation model reduction techniques for complex systems*". This project was funded by Advanced Thermal Control Consortium.

Since January 2021, he has started his career at MSC Software, Netherlands, as a development engineer on finite element modelling and high-performance computing.

Daming's research interests cover computational mathematics for control systems and model approximation or reduction techniques.

