Robust and Linear Parameter-Varying Control of Aeroservoelastic Systems

Vom Promotionsausschuss der Technischen Universität Hamburg-Harburg zur Erlangung des akademischen Grades Doktor-Ingenieur (Dr.-Ing.) genehmigte Dissertation

von
Julian Theis
aus
Hamburg

2018
Erster Gutachter:
Prof. Dr. Herbert Werner
Institut für Regelungstechnik
Technische Universität Hamburg-Harburg

Zweiter Gutachter:
Prof. Dr. Peter J. Seiler
Aerospace Engineering and Mechanics Department
University of Minnesota, Minneapolis

Vorsitzender des Prüfungsauschusses:
Prof. Dr.-Ing. Gerhard Bauch
Institut für Nachrichtentechnik
Technische Universität Hamburg-Harburg

Eingereicht am: 10.07.2017
Tag der mündlichen Prüfung: 27.02.2018
The king is gone, but he’s not forgotten.
Preface

The present dissertation is the result of three years of work at the Institute of Control Systems at Hamburg University of Technology and 14 months at the Aerospace Engineering and Mechanics Department at the University of Minnesota, Minneapolis. First, I like to thank my advisor Prof. Herbert Werner for his constant support, the amount of trust he put in me, and for the tremendous leeway. I just as much like to thank Prof. Pete Seiler for co-advising my thesis and for providing invaluable guidance for my research. My thanks also go to Prof. Gerhard Bauch for chairing the committee.

This thesis would not have been possible without the persistent mentorship of Prof. Harald Pfifer. I cannot express nearly enough gratitude for the provided opportunities and for believing in my capacity to make use of them, for the countless discussions on robust control, and for the friendship that evolved with them. A special thanks also goes to my friend and colleague Daniel Ossmann for just as many relevant discussions, for keeping me motivated, and most of all for always putting things into perspective. I thank my colleagues at the Institute of Control Systems for making work as enjoyable as it was, pars pro toto and in particular, Simon Wollnack, Michael Heuer, Sophia Feder, and Antonio Mendez. I would like to thank Prof. Edwin Kreuzer for giving me confidence in my abilities. I also thank Marc-André Pick, Christian Radisch, Axel Hackbarth, and Eugen Solowjow for their support and companionship. I would also like to acknowledge that I am still indebted to Florian Saupe for his support during my first few months and like to thank Andreas Knoblauch for interesting and fun discussions. Further, I would like to thank my colleagues at the University of Minnesota for making my stays a very pleasant and memorable experience. In particular, thanks go to David Escobar, Aditya Kotikalpudi, Abhineet Gupta, Claudia Moreno, Sally Ann Keyes, Raghu Venkataraman, Brian Taylor, as well as Béla Takarics and Tamás Peni.

I would like to give credit to Prof. Dave Schmidt, Arnar Hjartarson, Jen Annoni, Brian Danowski and Prof. Dale Enns. They contributed to this thesis by providing models and helpful insights. I would also like to express my eternal gratitude to Prof. Andy Packard. Even though he was not directly involved with the work related to this thesis, I realize more and more how essential his teaching was in developing my understanding of control theory. Finally, I am infinitely grateful for having had the opportunity to work with Prof. Gary Balas. Meeting him had a very positive impact on my life and he will always be an inspiration for me. His unparalleled enthusiasm brought out the best in everyone. I would like to dedicate this dissertation to his memory.
Summary

The present thesis considers the control of aeroservoelastic systems characterized by a strong interaction of rigid-body dynamics, structural dynamics, aerodynamics, and feedback control systems. Increased structural flexibility, desirable from a design perspective, e.g., as in high aspect ratio large span aircraft for improved fuel efficiency, often results in adverse system dynamics. Lightly damped oscillations, increased structural loads, and even instability are possible consequences. Feedback control is hence becoming increasingly important to mitigate these adverse effects in a wide range of applications subject to aeroelastic deformations, including small unmanned aircraft, large manned aircraft, and wind turbines.

Aeroservoelastic systems are often modeled as dynamic systems with a parametric dependence on the surrounding fluid, e.g., in the form of airspeed. The framework of linear parameter-varying (LPV) systems is hence a natural approach to aeroservoelastic control. Building on linear robust control techniques, it provides a formal way to design gain-scheduled control systems and extends norm-optimal control theory to linear systems that depend continuously on external operating conditions. Both, a large body of available theoretical results and computational tools make the framework attractive.

LPV models of aeroservoelastic systems are often of high order which prohibits their immediate use for controller synthesis, whose complexity scales badly with the model order. The first major contribution of this thesis is therefore to develop algorithms to obtain accurate reduced-order LPV models for control design. The nonlinear dependence on time-varying external parameters complicates this task and state-of-the-art approaches suffer from the same computational limitations as controller synthesis algorithms. Two different model reduction methods are proposed: An interpolation method based on local modal decomposition and a subspace method based on local measures of controllability and observability. The methods are demonstrated on models of aeroelastic unmanned aircraft and a model of the aerodynamics of a wind turbine.

The robust control approach formulates objectives in terms of sensitivity reduction and is shown to be ideally suited for aeroservoelastic control where damping augmentation under large model uncertainty with frequency-limited control activity is the main objective. The second major contribution of this thesis is to provide an in-depth review of mixed sensitivity loopshaping design techniques with an emphasis on the importance of a two-degrees-of-freedom structure. A generic problem formulation with a manageable number of comprehensible design parameters is proposed and systematic design guidelines are developed along the lines of a unifying review of the loopshaping paradigm. Two realistic application examples are considered in detail: active flutter suppression on a small research drone and the integration of aeroservoelastic control objectives into a more general flight control system on a large flexible aircraft.
# Contents

Notation iii  

Acronymns and Abbreviations v  

1 Introduction and Motivation 1  
1.1 Models of Aeroservoelastic Systems ........................................ 2  
1.2 Aeroservoelastic Control .................................................... 3  
1.3 Model Order Reduction ....................................................... 5  
1.4 Contributions ................................................................. 6  
1.5 Thesis Structure ............................................................... 8  

2 Linear-Parameter Varying Systems 9  
2.1 LPV Models of Nonlinear Dynamic Systems ............................... 10  
2.2 Preliminary Concepts ......................................................... 17  
2.2.1 Fundamental Properties .................................................. 17  
2.2.2 Induced $L_2$-norm ...................................................... 19  

3 Model Order Reduction for LPV Systems 23  
3.1 Preliminary Model Order Reduction Concepts ............................. 23  
3.1.1 State Space Manipulations .............................................. 24  
3.1.2 State-of-the-Art Method: LPV Balancing and Truncation ............. 29  
3.1.3 Related Problems .......................................................... 34  
3.1.4 Evaluation of Reduced-Order Models .................................. 37  
3.2 Approximation by Modal Interpolation ..................................... 38  
3.2.1 Local Model Reduction .................................................. 38  
3.2.2 Modal Matching ............................................................ 39  
3.2.3 Stability Considerations and Limitations .............................. 41  
3.2.4 Numerical Example ....................................................... 42  
3.3 Parameter-Varying Oblique Projection .................................... 44  
3.3.1 Parameter-Varying Petrov-Galerkin Conditions ....................... 44  
3.3.2 Basis Space and Test Space Construction .............................. 46  
3.3.3 Stability Considerations and Limitations .............................. 50  
3.3.4 Numerical Example ....................................................... 52
4 Control Design by Loopshaping

4.1 Control System Structure ............................................ 53
4.2 Loopshaping Design ................................................... 58
  4.2.1 Classical Loopshaping .............................................. 60
  4.2.2 Optimal Loopshaping ............................................... 61
  4.2.3 Mixed Sensitivity Loopshaping .................................... 63
4.3 Synthesis Machinery .................................................. 71
  4.3.1 $H_{\infty}$ Controller Synthesis .................................... 71
  4.3.2 LPV Controller Synthesis .......................................... 73
  4.3.3 Suboptimal Synthesis and Implementation ....................... 74
  4.3.4 Discretization .................................................... 75
4.4 Robustness Margins .................................................. 76
  4.4.1 Disk Margins ..................................................... 76
  4.4.2 Robustness in Mixed Sensitivity Design ......................... 79

5 Application Examples .................................................. 83

5.1 Model Order Reduction for Aeroservoelastic Aircraft .............. 83
  5.1.1 Body Freedom Flutter Vehicle .................................... 83
  5.1.2 X56A MUTT ........................................................ 92
5.2 Model Order Reduction for Far Wakes of a Wind Turbine .......... 97
5.3 Active Flutter Suppression ............................................ 100
  5.3.1 Model of the miniMUTT Aircraft ................................ 100
  5.3.2 Control Law Design ............................................... 105
  5.3.3 Control Law Verification ......................................... 111
5.4 Integrated Aeroservoelastic Control ................................ 114
  5.4.1 Model of the B-1 Aircraft ....................................... 115
  5.4.2 Control Law Design ............................................... 117
  5.4.3 Control Law Verification ......................................... 122

6 Conclusions and Future Research Opportunities ....................... 135

Bibliography .................................................................... 139

List of Publications ....................................................... 1

Curriculum Vitae .......................................................... III
# Notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{R}$</td>
<td>field of real numbers</td>
</tr>
<tr>
<td>$\mathbb{R}_+$</td>
<td>field of non-negative real numbers</td>
</tr>
<tr>
<td>$\mathbb{R}^{n \times m}$</td>
<td>set of $n$-by-$m$ matrices with elements in $\mathbb{R}$</td>
</tr>
<tr>
<td>$\mathbb{C}$</td>
<td>field of complex numbers</td>
</tr>
<tr>
<td>$\mathbb{C}^{n \times m}$</td>
<td>set of $n$-by-$m$ matrices with elements in $\mathbb{C}$</td>
</tr>
<tr>
<td>$X^\perp$</td>
<td>orthogonal complement of a vector space $X$</td>
</tr>
<tr>
<td>$j$</td>
<td>imaginary number $\sqrt{-1}$</td>
</tr>
<tr>
<td>$s$</td>
<td>Laplace variable</td>
</tr>
<tr>
<td>$I_n$</td>
<td>$n$-dimensional identity matrix</td>
</tr>
<tr>
<td>$0_{n \times m}$</td>
<td>$n$-by-$m$ matrix with zero entries</td>
</tr>
<tr>
<td>$X^T$</td>
<td>transpose of a matrix $X$</td>
</tr>
<tr>
<td>$X^{-1}$</td>
<td>inverse of a matrix $X$</td>
</tr>
<tr>
<td>$X^\dagger$</td>
<td>pseudo-inverse of a matrix $X$</td>
</tr>
<tr>
<td>$X \prec 0$</td>
<td>$X$ is a symmetric negative definite matrix</td>
</tr>
<tr>
<td>$X \preceq 0$</td>
<td>$X$ is a symmetric negative semi-definite matrix</td>
</tr>
<tr>
<td>$X \succ 0$</td>
<td>$X$ is a symmetric positive definite matrix</td>
</tr>
<tr>
<td>$X \succeq 0$</td>
<td>$X$ is a symmetric positive semi-definite matrix</td>
</tr>
<tr>
<td>$X \prec Y$</td>
<td>$X - Y$ is a symmetric negative definite matrix</td>
</tr>
<tr>
<td>$X \succ Y$</td>
<td>$X - Y$ is a symmetric positive definite matrix</td>
</tr>
<tr>
<td>$\text{diag}(X,Y)$</td>
<td>block diagonal concatenation of matrices $X$ and $Y$, i.e., $[ X \ 0 \ Y ]$</td>
</tr>
<tr>
<td>$\text{span}(X)$</td>
<td>span or range space of a matrix $X$</td>
</tr>
<tr>
<td>$\ker(X)$</td>
<td>kernel or null space of a matrix $X$</td>
</tr>
<tr>
<td>$\lambda_{\text{max}}(X)$</td>
<td>maximum eigenvalue of a matrix $X$</td>
</tr>
<tr>
<td>$\lambda_{\text{min}}(X)$</td>
<td>minimum eigenvalue of a matrix $X$</td>
</tr>
<tr>
<td>$\sigma_{\text{max}}(X)$</td>
<td>maximum singular value of a matrix $X$</td>
</tr>
<tr>
<td>$\sigma_{\text{min}}(X)$</td>
<td>minimum singular value of a matrix $X$</td>
</tr>
</tbody>
</table>

$x := y$  
$x$ is defined as $y$

$\dot{x}$  
$\frac{dx}{dt}$  
$\frac{\partial}{\partial y} x$  
partial derivative of a function $x$ with respect to the variable $y$

$x |_y$  
evaluation of a function $x$ at argument $y$

$|x|$  
$\|x\|$  
absolute value of $x$  
norm of $x$
Notation

■ end of theorem
□ end of proof
▲ end of definition
△ end of example
Acronyms and Abbreviations

DLM  doublet lattice method
GM   gain margin
HALE high altitude long endurance
IMU  inertial measurement unit
LFT  linear fractional transformation
LMI  linear matrix inequality
LPV  linear parameter-varying
LQG  linear quadratic Gaussian
LTI  linear time-invariant
LTV  linear time-varying
MIMO multi-input multi-output
MUTT multi utility technology testbed
PM   phase margin
PID  proportional-integral-derivative
PWM  pulse width modulation
SAS  stability augmentation system
SISO single-input single-output
SMCS structural mode control system
SVD  singular value decomposition
VLM  vortex lattice method

cf.  compare (confer)
e.g. for example (exempli gratia)
i.e. that is (id est)
ibid. see previous reference (in the same place, ibidem)
1 Introduction and Motivation

Current developments in the aerospace industry are driven by the need to increase fuel efficiency to reduce operating cost and to meet increasingly demanding regulations to protect the environment. This trend is illustrated in Figure 1.1 for several commercial aircraft over the last 60 years.

![Figure 1.1: Fuel consumption of commercial aircraft [Knoblach 2015, p. 1, based on the Internal Energy Agency (IEA) report 2009].](image)

The consequences of this trend for aircraft design can be exemplified with the help of Breguet’s range equation

\[ R = V I \frac{L}{D} \ln \left( \frac{m_{\text{takeoff}}}{m_{\text{landing}}} \right), \tag{1.1} \]

where \( V \) denotes speed, \( I \) denotes propulsion efficiency, the ratio \( L/D \) of aerodynamic lift and aerodynamic drag is also known as the glide number, and \( m \) denotes mass. The
difference between mass at take-off and at landing for a transport aircraft is the burned fuel mass. Besides the possibility to increase the propulsion efficiency of the engines, there thus 
remains essentially only one design factor to improve fuel efficiency: increasing aerodynamic efficiency. Doing so requires to reduce the drag $D$, which can be decomposed into two 
components: profile drag and (lift-)induced drag. The induced drag is proportional to $\Lambda^2/\Lambda$, where $\Lambda$ is the wing aspect ratio, calculated by dividing the squared wing span by the 
wing area. The profile drag depends on the specific form of the airfoil, but generally a low 
thickness-to-chord ratio is desirable. Thus, “aerodynamically efficient” can be translated 
to large span, slender wings. This trend is apparent in state-of-the-art aircraft such as the 
Boeing 787 and Airbus A350 aircraft and is expected to continue for future generations of 
commercial aircraft. Other important applications include high altitude long endurance (HALE) unmanned aerial systems that are thought to provide a low-cost alternative 
to satellites for broadcasting and surveillance purposes in the foreseeable future. The 
42,438 km flight of the (manned) “Solar Impulse 2” aircraft in 2015–2016 using only solar 
power can be seen as a successful proof of concept for such HALE systems.

The benefit of reduced drag, however, comes at a price. Stress caused by bending is 
inversely proportional to the second moment of area, so that such wing structures are 
usually subject to larger loads. Further, they are much more flexible with significantly 
lower natural frequencies and increased in-flight deformations. This can cause undesired 
coupling of rigid body dynamics and elastic deformation through aerodynamic forces and 
feedback control systems. This complex of mutual interaction is termed aeroservoelasticity 
[e.g. Tewari 2015]. A loss of control effectiveness and degradation of handling qualities 
are commonly encountered as adverse consequences of aeroservoelastic coupling. Further, 
such coupling can even lead to an instability known as flutter beyond a critical airspeed 
[e.g., Försching 1974, Cha. 6, Wright & Cooper 2015, Cha. 10]. Finally, aeroservoelasticity 
remains one of the major obstacles in the realization of advanced aircraft design concepts 
such as joined-wing (Prandtl) planes or adaptive conformal wings [cf. Chambers 2005]. The use of active control concepts to counter the adverse effects of aeroelasticity is thus 
an important field of research and the present thesis seeks to contribute towards this aim.

1.1 Models of Aeroservoelastic Systems

As the flexibility increases, it becomes necessary to consider aeroelastic effects explicitly 
in aircraft models. Structural dynamics models can be expressed using a large number 
of nodes with individual degrees of freedom to spatially discretize the aircraft as is done, 
e.g., in finite element analysis. Alternatively, they can be expressed in terms of “assumed 
shapes”, e.g., eigenforms of the structure [Wright & Cooper 2015, Cha. 3]. Such shapes 
can be obtained either from ground vibration tests or finite element models and provide a 
spatial basis for the structural dynamics, i.e., they restrict the motion of individual nodes 
relative to each other. Assumed shapes therefore result in lower-order models, but often 
still contain several tens of modes. Modeling aeroelasticity further requires aerodynamics 
models. A common modeling method is the doublet lattice method (DLM), a panel 
method that calculates the flow across a harmonically oscillating lifting surface and hence
extends the vortex lattice method (VLM) by considering unsteady aerodynamic effects. A
dynamic system that describes this flow is then usually obtained through rational function
approximation and possibly involves thousands of state variables [e.g., Knoblach 2015,
Cha. 2, Tewari 2015, Cha. 4]. Approaches from computational fluid dynamics result in
even higher order models with millions of state variables [e.g., Farhat et al. 1995].

The structural dynamics and aerodynamics models need to be adequately coupled to
the rigid-body flight mechanics model of an aircraft. One popular approach is based on a
mean axes reference frame [Waszak & Schmidt 1988, Schmidt 2012, Sec. 4.2]. The mean
axes constitute a floating frame whose origin is chosen at the instantaneous center of
mass and thus not fixed to a material point on the aircraft. As a consequence, structural
deformations can be modeled as in-vacuo (free-free) vibrations with respect to these axes,
which can considerably simplify the integration of assumed shapes structural models [e.g.,
Looye 2008, Cha. 3, Moreno 2015, Sec. 2.1].

Thus, aeroelastic models are often of high order with the aerodynamics model being the
main driver. Even more importantly, their predominant characteristic is that their dynamics
vary substantially with the physical properties and stream velocity of the surrounding
fluid. Linear parameter-varying (LPV) systems, first proposed by Shamma [1988], are
linear dynamic systems that depend continuously on time-varying external parameters. As
they can explicitly account for dynamic variations due to changing operating conditions,
they are a natural modeling framework for aeroservoelastic systems. The LPV framework
is very useful for the analysis and design of gain-scheduled control systems and particularly
popular for aerospace applications [cf. Balas 2002b]. Its attractiveness is based on the
availability of a large body of theoretical results for analysis and synthesis [Becker et al.
tools [Hjartarson et al. 2013, 2014, Balas et al. 2015].

1.2 Aeroservoelastic Control

Classical flight control depends largely on one-loop-at-a-time design procedures and cas-
caded layers of control systems [e.g., Schmidt 2012, Brockhaus et al. 2013]. Inner loops
provide desirable dynamic behavior and are known as stability augmentation systems
(SAs). They are usually gain-scheduled so that the augmented dynamics remain similar
across the flight envelope. Outer loops for maneuver demand and autopilot functions are
then concatenated on the inner loops. As described, active mitigation of adverse aeroser-
voelastic effects is becoming an increasingly important aspect of modern flight control
systems. With present-day fly-by-wire systems, including objectives such as reduction of
structural vibrations, ride quality enhancement, and gust load alleviation into the flight
control system has become feasible and can be found on several of today’s commercial
aircraft [Balas 2003]. Usually, such systems form an extra inner control layer and a strict
frequency separation of the individual control systems is prevalent to avoid excitation of
structural modes by the primary control systems and limit interference of the secondary
controllers with handling qualities [Caldwell et al. 2000]. The increasing flexibility of wings, however, renders such strategies obsolete: Elastic and rigid-body dynamics are becoming too close in frequency to be separately addressed. Figure 1.2 illustrates this effect and further shows that the activity of aeroservoelastic control systems extends well into the frequency range of aeroelastic dynamics. This necessitates both accurate models of these dynamics as well as robust control systems that can tolerate the inevitable uncertainties in these models.

![Frequency regimes in aeroservoelastic control.](image)

(a) Sufficient frequency separation for classical control.  
(b) Insufficient frequency separation in aeroservoelastic systems.

Early research on aeroservoelastic control relied to a large extent on what is known as *collocated feedback* within the structural control community. Collocated feedback employs sensors and actuators in the same location. The special property of such feedback loops is the presence of a complex pair of zeros in the immediate vicinity of the lightly damped poles of the structural mode [e.g., Preumont 2002, Cha. 4–5]. Collocated controllers are, in general, easily designed using root-locus analysis and have favorable robustness properties, at least in terms of gain margins\(^1\). These advantages, however, come at the expense of a limited achievable damping augmentation, as the lightly damped open-loop poles are drawn towards the similarly lightly damped open-loop zeros under feedback. A closely related approach, termed the concept of *identically located force and acceleration* by Wykes [1968], was successfully applied to address the damping of structural modes on the B-1 aircraft [Wykes et al. 1977, 1980]. A similar configuration was also used in the first flight test\(^2\) beyond flutter speed, conducted in 1973 on a modified B-52 aircraft [Roger et al. 1975]. The control system on that aircraft involved two single feedback loops which fed back filtered vertical acceleration signals, acquired on the wing, to nearby located outboard ailerons and flaperons. Collocated acceleration feedback is also proposed in various other publications concerned with flutter suppression, e.g., by Adams et al. [1992], Waszak & Srinathkumar

---

\(^1\) Cf., e.g., Hanel [2001, Sec. D.4] and Balas [1990, Sec. 9] for a critical assessment of the robustness of collocated control.

\(^2\) A historical overview of flutter flight testing can be found in Kehoe [1995]. An Overview of some of NASA’s more recent research activities can be found in Perry et al. [1992] and Chambers [2005].
The restrictive structure and limited achievable damping are, however, disadvantages of collocated control. Recent research, e.g., by the Lockheed Martin company [Holm-Hansen et al. 2010] and NASA [Ryan et al. 2014] focuses on the use of multivariable linear quadratic Gaussian (LQG) controllers. These controllers require gain-scheduling and blending techniques in order to adapt the controller over the flight envelope and maintain stability and performance.

The use of LPV control techniques to directly incorporate these controller adaptations therefore seems natural. Several researchers have investigated the use of robust and LPV control techniques for aeroservoelastic control in the past with encouraging results, e.g., van Etten et al. [1999], Barker et al. [1999], Barker & Balas [2000], Hanel [2001], Waszak [2001], and Hjartarson et al. [2013].

1.3 Model Order Reduction

For LPV systems, both analysis and synthesis techniques require the solution of convex optimization problems that are described mathematically by linear matrix inequalities (LMIs). The required computation for the solution of these problems grows rapidly with increasing state space dimension. The applicability of available LPV techniques therefore remains limited to models with relatively few state variables. With current tools, models with an order of about 50 state variables are tractable. For many engineering problems, directly obtaining models from first principles with such a low number of state variables is not easy. In particular, models of aeroservoelastic systems that combine structural dynamics and unsteady aerodynamics as described in Section 1.1 are particularly prone to be of high order. In these cases, model simplification is required to make the problem amenable to existing methods and tools.

LPV model order reduction was first addressed by Wood [1995, Cha. 7] with an application to flexible helicopter blades. The method is a generalization of the concept of balancing and truncation [Moore 1981, Pernebo & Silverman 1982]. Balancing and truncation consists of a state space coordinate transformation so that the state variables can be related to their contribution to input-output behavior in terms of a norm bound, followed by removing state variables that are considered negligible in the new coordinates. For LPV systems, the procedure requires itself the solution of an optimization problem that involves LMIs and thus suffers from the same computational limitations as analysis and synthesis methods. Furthermore, converting an LPV model to a form that is suitable for reduction, i.e., the separation of “important” and “unimportant” state variables, generally involves a parameter-dependent state transformation. Such a parameter-varying transformation acknowledges that the importance of state variables varies over the parameter domain, but leads to a parameter-dependent state space basis and therefore introduces additional rate terms. This is usually undesired, since the complexity of the model is increased by properly retaining this additional dependence. In order to avoid this rate dependence, the search space for the optimization problem has to be restricted to parameter-independent transformations. This restriction may cause the solution to fail or deteriorate the quality of the approximation.
1 Introduction and Motivation

Recently, research to avoid LMIs in model simplification of LPV systems has gained increased attention. Aeroservoelastic systems can be identified as a major driver for this trend, as applications such as flexible aircraft wings [e.g., Amsallem & Farhat 2011, Poussot-Vassal & Roos 2011, 2012, Moreno et al. 2014, Wang et al. 2016, Luspay et al. 2018] and flexible structures in wind turbines [Adegas et al. 2013] are outside the scope of existing techniques. Probably the simplest approach to avoid LMIs is to calculate a state space transformation from local (frozen-parameter) models. Such an approach can be seen to conceptually draw from model order reduction for general nonlinear systems [Lall et al. 2002, Rathinam & Petzold 2003]. In particular, transformations can be constructed to meet local approximation criteria at certain points within the parameter domain [e.g., Balas 2002a, Poussot-Vassal & Roos 2012, Moreno 2015, Sec. 4.3.2, Annoni & Seiler 2016]. Achieving such a goal with a parameter-independent transformation can, however, require a large number of state variables to capture systems subject to large dynamics variations with sufficient accuracy and can therefore not result in the desired order reduction. Another approach is therefore to calculate local transformations at various parameter values and to interpolate between grid points. Doing so is closely related to the problem of parametric model reduction, which has also received considerable attention in recent years [e.g., Benner et al. 2013, Baur et al. 2011, Panzer et al. 2010]. The fundamental difference is, however, that parametric model reduction considers constant parameter values with the goal to approximate a family of parameterized linear time-invariant (LTI) models. Contrary, the LPV model order reduction problem requires to account for time-varying parameter values to successfully approximate an LPV system. Nevertheless, the field of parametric model reduction has had some impact also on LPV model order reduction. Specifically, methods were developed that use LTI techniques for frozen-parameter models to locally reduce LPV systems and then seek to interpolate the reduced-order models for time-varying parameters [e.g., Amsallem & Farhat 2011, De Caigny et al. 2011, Adegas et al. 2013, Gözse et al. 2016, Luspay et al. 2018]. These techniques all have in common that they construct local approximations of the LPV system, i.e., they find reduced models for fixed parameter-values and then construct a new LPV model through interpolation of these local models. A problem that frequently arises in this context is that of state-space consistency, as local models are usually described in different state space bases. These approaches draw conceptually from classical gain-scheduling ideas and suffer from the same fundamental limitations, i.e., validity between individual points in the parameter domain can usually not be assured. Thus, questions about validity and the quality of the approximation remain largely unanswered.

1.4 Contributions

The present dissertation leverages the LPV systems framework for new and challenging applications that arise in the field of aeroservoelastic systems. Models of aeroservoelastic systems are often of high order which limits the applicability of existing LPV analysis and synthesis tools. Further, state-of-the-art methods for model order reduction of LPV systems suffer from the same computational limitations. The thesis addresses this problem
and contributes two novel model order reduction methods that are applicable to models which are well outside the scope of the existing methods. As LMI solutions quickly become numerically intractable, even for moderate state dimensions, the novel methods seek to avoid LMI solutions and instead rely on local information from individual points within the parameter space. They are hence heuristic in nature and trade off the guarantees of formal, LMI-based methods for computational tractability. As a consequence of the local approach, special care has to be taken to maintain a consistent state space basis for the reduced-order LPV model.

The first proposed method achieves consistency by making use of modal state space representations to construct a parameter-varying basis for the reduced-order model which is suitable for interpolation. With its parameter-varying basis, this method can be used to perform frequency range decompositions of LPV systems and is further also applicable to systems that involve unstable dynamics. The main technical contribution of this method is a “matching function” to associate dynamic modes of individual LTI models representing the same LPV system for different parameter-values, i.e., operating conditions. The method is demonstrated on the high-fidelity industry-grade model of an unmanned aircraft and shown to yield a model with lower order than the state-of-the-art method while similar accuracy is maintained. These results have been pre-published in the research paper by Theis, Takarics, Pfifer, Balas & Werner [2015c].

A parameter-varying oblique projection is proposed as the second novel method. It achieves consistency by constructing a Petrov-Galerkin approximant with a constant basis. The novelty of the method and a major theoretical contribution is the use of a parameter-varying basis for the test space in this approximation which provides additional freedom over using a parameter-independent basis. Basis and test space are calculated from observability and controllability information so that the resulting projection resembles state-of-the-art balancing and truncation model order reduction. The method is demonstrated on another high-fidelity model of an unmanned aircraft and also on the aerodynamics model of a wind turbine, proving applicability for models with up to thousands of state variables. It has been pre-published in research papers by Theis, Seiler & Werner [2016b, 2018].

Using these model order reduction techniques, it becomes possible to reduce models down to an order where existing controller synthesis methods are applicable. While such synthesis techniques are well documented in the literature, a lack of design guidelines is identified. Another major contribution of the present thesis is therefore a comprehensive compilation and an in-depth review of the loopshaping paradigm for control systems design. A particular emphasis is put on the importance of two-degrees-of-freedom control structures and on providing links between several common design methodologies in order to extract a unified design perspective. The result is a general parameterization for mixed sensitivity design which involves a manageable number of design parameters with a clear physical interpretation and an intuitive relation to design objectives. Further, the transition to concrete design guidelines is provided. For instance, special challenges of aeroservoelastic systems are met by formulating damping requirements in terms of reducing peaks in modal sensitivities.

The final contribution of this thesis are two detailed application-oriented aeroservoelastic control design studies. These design studies not only substantiate the previously formulated
1 Introduction and Motivation

Design guidelines but also provide insight into the challenges and limitations that arise in aeroservoelastic control through in-depth evaluations and robustness analyses. The first study is concerned with active flutter suppression on a small unmanned aircraft and envelope expansion into the naturally unstable flight regime. The complete design process is expounded and the importance of individual components such as actuators and computation devices are highlighted. The second study is concerned with the model of a large flexible aircraft and introduces a flight control design concept for integrating structural damping augmentation with primary flight control. The proposed multivariable controller is compared in detail with a conventional cascaded control architecture which employs dedicated collocated feedback loops for structural mode attenuation. The proposed controller is shown to greatly reduce structural vibrations and also to be able to target modes that are not addressable via the conventional controller. Initial results from these design studies were published in research papers by Theis, Pfifer & Seiler [2016a] and Theis, Pfifer, Balas & Werner [2015a].

1.5 Thesis Structure

Aeroservoelastic systems are either modeled directly as LPV systems or can be approximated very well by LPV models through parameter-dependent linearization of nonlinear models. This approximation is detailed in Chapter 2. The chapter then continues with a brief review of the LPV framework and compiles several fundamental analysis results.

Models of aeroservoelastic systems may be of high order which severely limits the applicability of existing LPV analysis and synthesis tools. Chapter 3 starts with a review of fundamental model order reduction techniques and compiles the state-of-the-art methods for LPV systems. The limitations of these methods, recent approaches to alleviate them, and open questions are discussed and the two novel model order reduction methods are developed.

The purpose of feedback control is, in short, sensitivity redistribution. This perspective is elaborated in Chapter 4 along the lines of a review of classical loopshaping, optimal loopshaping, and mixed sensitivity loopshaping design techniques. Systematic guidelines for the design of controllers through mixed sensitivity loopshaping are then developed with a particular focus on the use of a manageable number of design parameters which have clear relations to the design objectives.

Finally, Chapter 5 provides a number of comprehensive application examples. The novel model order reduction methods are used to obtain low-order approximations of two different models of unmanned aircraft and the aerodynamics model of a wind turbine. The thesis concludes with two in-depth design studies for aeroservoelastic control applications: active flutter suppression on a small unmanned aircraft and a flight control design concept for a large aircraft that integrates structural damping augmentation and primary flight control systems.
2 Linear-Parameter Varying Systems

This chapter briefly reviews the framework of LPV systems and sets the ground for the other chapters of this thesis. A particular emphasis is put on the interpretation of LPV models as continuous linearizations of nonlinear systems with a parametric dependence on external operating conditions. The remainder of the chapter compiles some fundamental analysis results for stability and performance of LPV systems.

LPV systems are dynamic systems whose state space representation involves continuous matrix functions of a time-varying parameter vector that is not known in advance but can be measured at each time instant. Such a time-varying parameter is called scheduling parameter.

**Definition 2.1 (Scheduling parameter).** A scheduling parameter for an LPV systems is an unknown continuous function of time $\rho: \mathbb{R}^+ \mapsto \mathcal{P} \subset \mathbb{R}^{n_{\rho}}$ that is restricted to a compact set of admissible parameter trajectories, defined as

$$\mathcal{T} = \{ \rho \mid \rho(t) \in \mathcal{P} \land \dot{\rho}(t) \in \mathcal{Q} \land t \in \mathbb{R} \}.$$ \hspace{1cm} (2.1)

That is, the scheduling parameter $\rho(t)$ takes values

$$\rho \in \mathcal{P} \subset \mathbb{R}^{n_{\rho}} \quad (2.2)$$

in a specified compact set of admissible parameter values $\mathcal{P}$ and the rate of parameter variation $\dot{\rho}: \mathbb{R}^+ \mapsto \mathcal{Q} \subset \mathbb{R}^{n_{\rho}}$ is bounded by a hypercube

$$\mathcal{Q} = \{ q \in \mathbb{R}^{n_{\rho}} \mid q_i \leq q \leq \bar{q}_i, \quad i = 1, \ldots, n_{\rho} \},$$ \hspace{1cm} (2.3)

where $\underline{q}_i$ and $\bar{q}_i$ denote minimum and maximum rates for each component of the scheduling parameter. The set of admissible parameter trajectories is also called domain of the LPV system and is usually selected based on physical considerations. The set of admissible parameter values is also called parameter domain. ▲

**Definition 2.2 (LPV systems).** A state space representation of a linear parameter-varying system on the domain $\mathcal{T}$ is defined as

$$P_{\rho} : \begin{cases} \dot{x}(t) = A(\rho(t)) \ x(t) + B(\rho(t)) \ u(t) \\ y(t) = C(\rho(t)) \ x(t) + D(\rho(t)) \ u(t) \end{cases},$$ \hspace{1cm} (2.4)
with known continuous functions

\[ A: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x}, \quad B: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_u}, \]
\[ C: \mathcal{P} \mapsto \mathbb{R}^{n_y \times n_x}, \quad D: \mathcal{P} \mapsto \mathbb{R}^{n_y \times n_u}. \]

The function \( x: \mathbb{R}_+ \mapsto \mathbb{R}^{n_x} \) is called the state, \( u: \mathbb{R}_+ \mapsto \mathbb{R}^{n_u} \) is called the input, and \( y: \mathbb{R}_+ \mapsto \mathbb{R}^{n_y} \) is called the output.

The continuity assumption on the scheduling parameter ensures both the existence and uniqueness of the state trajectories of the LPV system [Apkarian et al. 1996, Sec. 8.3]. Further, compactness together with the continuity assumption on the state space matrix functions guarantees boundedness of its state space data [ibid.].

As the actual parameter trajectory \( \rho(t) \) and its derivative \( \dot{\rho}(t) \) are not known a priori, the system has to be treated as uncertain with respect to the scheduling parameter trajectory. This makes LPV systems fundamentally different from linear time-varying (LTV) systems, where time dependence is exactly known. In fact, LPV systems can be seen to encompass a family of LTV systems. For any specific parameter trajectory within the set of admissible trajectories, the LPV system becomes an LTV system. Similarly, for a fixed parameter value, the LPV system becomes an LTI system [Apkarian et al. 1996, Sec. 8.4].

The parameter domain is commonly approximated by a finite dimensional subset \( \{ \rho_k \}_{k=1}^{n_{grid}} \subset \mathcal{P} \), called a grid. In a grid representation, the LPV system is described as a collection of LTI models \( (A_k, B_k, C_k, D_k) := (A(\rho_k), B(\rho_k), C(\rho_k), D(\rho_k)) \) obtained from evaluating the LPV model on the subset \( \{ \rho_k \}_{k=1}^{n_{grid}} \). Such a representation naturally arises, e.g., if a nonlinear system is linearized at multiple operating conditions as detailed in Section 2.1 or if parameterized LTI models form the basis of an LPV model as it is often the case in aerospace. The grid representation is also the most general way of representing an LPV system and requires no further assumptions on the form of parameter dependence. Other classes of LPV models exist and are frequently encountered in the literature, e.g., linear fractional transformation (LFT)-type models whose state space realizations depend rationally on \( \rho \) [e.g., Packard 1994, Apkarian & Gahinet 1995] or polytopic models with affine parameter dependence [e.g., Gahinet et al. 1996]. These models are less general than the gridded models used here, but have additional structure that can be exploited in developing computational algorithms. The dependence on parameters and time is occasionally dropped in this thesis to shorten notation, when no ambiguity is expected.

### 2.1 LPV Models of Nonlinear Dynamic Systems

As the LPV model inevitably is at the core of LPV control, this section revisits the use of Jacobian linearization for obtaining LPV models. A particular emphasis is put on the parameter-dependent coordinate system which results from this modeling approach. This fact was only recently brought to attention by Takarics & Seiler [2015] and appears to be widely ignored in the LPV literature.

Industry and research institutions alike often have dedicated modeling departments or groups that provide their high-fidelity models to the control engineers. This does however
not immediately mean that these models are suitable for control design. Most high-fidelity models, especially obtained from first principal modeling, are derived in the form of nonlinear differential equations

\[ \dot{x}(t) = f(x(t), u(t), \sigma(t)) \]
\[ y(t) = h(x(t), u(t), \sigma(t)), \]

where \( f \) and \( h \) are known continuously differentiable functions, \( x(t) \in \mathbb{R}^{n_x} \) is the state vector, \( u(t) \in \mathbb{R}^{n_u} \) is the input vector, and \( \sigma(t) \in \mathbb{R}^{n_\sigma} \) is a measurable external parameter vector. To make them amenable to control design techniques, these models are then often linearized with respect to a single constant equilibrium operating condition. Doing so results in an LTI model that can be used for linear control design. If the desired operating condition varies with time, then a constant approximation may not be sufficient. Classical remedies include the use of multiple LTI systems which represent different operating conditions for a pointwise design of gain-scheduled controllers [Rugh & Shamma 2000, Leith & Leithead 2000b].

The LPV framework was introduced by Shamma [1988] and Shamma & Athans [1990, 1991] in order to formalize classical gain-scheduling techniques which were (and are) used with great success in many applications [cf. Stein 1980]. Classical gain-scheduling considers isolated operating points [cf., e.g., Rugh 1990, Nichols et al. 1993, Lawrence & Rugh 1995]. Instead, the LPV framework focuses on stability and performance guarantees for a specified range of admissible operating conditions and the transition between different operating points. Models for LPV control have to be obtained through properly rewriting the nonlinear differential equations or by overbounding the original dynamic system with a more general representation. One popular technique is function substitution which refers to replacing nonlinearities with artificial parameters. The result is an LPV system which includes the original nonlinear system at the cost of potentially very high conservatism. Some formalisms exist, e.g., nonlinear state space transformations for so-called output nonlinear systems [Shamma & Cloutier 1993, Packard & Kantner 1996]. Aside from these formal approaches, models are often constructed ad hoc by rewriting the differential equation such that nonlinearities are absorbed in parameters. Detailed application examples of function substitution are provided, e.g., by Marcos & Balas [2004], Tan et al. [2000], Saupe [2013, Sec. 4.2], and Pfifer [2013, Sec. 4.1], but it has to be concluded that the procedure is often very problem specific and restrictive [cf. Leith & Leithead 2000b].

As these techniques are not suitable for many engineering applications, another popular technique for obtaining LPV models is Jacobian linearization. The nonlinear model is linearized with respect to a number of operating points in the parameter domain and the resulting parameterized LTI models are treated as an LPV model in dependence on the operating condition. Such a linearization naturally leads to a gridded representation although polytopic or LFT-type LPV models are also often constructed based on such an initial grid representation to which rational or affine functions are fitted [e.g., Pfifer & Hecker 2011]. Jacobian linearization is easily automated and feasible even in case no analytic expression of the model is available, e.g., when data is represented by lookup tables as it is usually the case in aerospace. Further, the process can easily be adapted and
repeated in case the model changes, e.g., during a design cycle. These are all important features for application and it therefore is not surprising that Jacobian linearization is quite common to obtain LPV models [e.g., Apkarian et al. 1995a, Apkarian et al. 1996, Sec. 14, Fialho et al. 1997, Apkarian & Adams 1998, Barker & Balas 2000, Balas 2002a, Chen 2010, Cha. 5, Abbas et al. 2014, Theis et al. 2014, 2015b]. The deficits of Jacobian linearization for LPV modeling, on the other hand, are well known and have been pointed out in detail, e.g., by Leith & Leithead [1998, 1999, 2000b] and Rugh & Shamma [2000]. Essentially, all guarantees that the LPV framework intends to provide are lost for the nonlinear system, because the resulting LPV model is merely an approximation and not an equivalent representation of this system. Still, the method continues to be the most prevalent way of obtaining models for LPV control and is detailed in the following.

A standard assumption in gain-scheduling and LPV control [e.g., Shamma & Cloutier 1993, Lawrence & Rugh 1995, Rugh & Shamma 2000, Takarics & Seiler 2015] is the existence of a continuum of (forced) equilibrium points \((\bar{x}, \bar{u}, \bar{y})\) for the nonlinear system \(f(x, u, \sigma) = 0\) and further that these points can be parameterized by a vector \(p \in \mathcal{P} \subset \mathbb{R}^n\). Thus, for each value \(p \in \mathcal{P}\)

\[
0 = f(\bar{x}(p), \bar{u}(p), \sigma(p))
\]

\[
\bar{y}(p) = h(\bar{x}(p), \bar{u}(p), \sigma(p)).
\]

(2.6)

It is convenient to associate all possible equilibria \(\bar{x}\) with potential operating conditions for a control system and \(\bar{y}, \bar{u}\) with the corresponding trim output and input. As the manifold of equilibria depends continuously on the parameter, this fixed parameter can be replaced with a time-varying parameter. Doing so results in a scheduling parameter as introduced in Definition 2.1. This scheduling parameter naturally includes (and often coincides with) the exogenous parameter \(\sigma\). A first-order Taylor approximation of the nonlinear dynamic system \((2.5)\) with respect to \((\bar{x}, \bar{u}, \bar{y})\) can be obtained as

\[
\dot{x} = A(p) (x - \bar{x}) + B(p) (u - \bar{u})
\]

\[
y = \bar{y}(p) + C(p) (x - \bar{x}) + D(p) (u - \bar{u}),
\]

(2.7)

where \(A(p) := \frac{\partial f}{\partial x}|_{p}\), \(B(p) := \frac{\partial f}{\partial u}|_{p}\), \(C(p) := \frac{\partial h}{\partial x}|_{p}\), and \(D(p) := \frac{\partial h}{\partial u}|_{p}\), with the subscript \(p\) denoting a function evaluation at \(x = \bar{x}(p), u = \bar{u}(p),\) and \(\sigma(p)\). The state equation \((2.7)\) is still in the original, i.e., absolute, state space coordinates of the original nonlinear dynamic system. This means that \(x\) represents absolute positions in the state space of the original nonlinear system. While each such “linearization” has consistent states, inputs, and outputs, it unfortunately is not actually linear due to the presence of the inhomogenous

\[1\] In case the scheduling parameter also includes internal signals such as state variables of the system, a “quasi LPV” representation results. In practice, actually internal parameters are often treated as being exogenous because of time scale separation or decoupling properties. An example is the airspeed in aircraft that usually changes much slower than the controlled dynamics.
terms related to \( \bar{x}, \bar{u}, \bar{y} \) [Leith & Leithead 1998], i.e.,

\[
\begin{align*}
\dot{x} &= A(p) \bar{x} + B(p) \bar{u} - (A(p) \bar{x}(p) + B(p) \bar{u}(p)) \\
y &= C(p) \bar{x} + D(p) \bar{u} + \bar{y}(p) - (C(p) \bar{x}(p) + D(p) \bar{u}(p)).
\end{align*}
\]

(2.8)

The terms \( d_x(p) \) and \( d_y(p) \), in general, cannot be neglected and may even dominate the dynamics such that the \((A, B, C, D)\) matrices alone are not an adequate approximation for the nonlinear dynamic system.

A (parameter-dependent) change of coordinates

\[
\begin{bmatrix}
x_\delta \\ u_\delta \\ y_\delta
\end{bmatrix} :=
\begin{bmatrix}
x \\ u \\ y
\end{bmatrix} -
\begin{bmatrix}
x(p) \\ u(p) \\ y(p)
\end{bmatrix}
\]

(2.9)

transforms the system to deviation coordinates with respect to the equilibrium manifold and leads to a system

\[
\begin{align*}
\dot{x}_\delta &= A(p) x_\delta + B(p) u_\delta \\
y_\delta &= C(p) x_\delta + D(p) u_\delta.
\end{align*}
\]

(2.10)

In Equation (2.10), the state derivative is still in absolute coordinates, while all other variables are in deviation coordinates. For a constant value \( p \), the derivative \( \dot{x} = \dot{x}_\delta \), but as the referenced equilibrium point depends continuously on the time-varying parameter \( p \), this is no longer true. In particular, the state derivative in deviation coordinates is

\[
\frac{d}{dt} x_\delta = \dot{x}_\delta - \frac{d}{dt} \bar{x}(\rho(t))
\]

(2.11)

such that the correct linearization with respect to the time-varying equilibrium becomes

\[
\begin{align*}
\dot{x}_\delta &= A(p) x_\delta + B(p) u_\delta - \frac{d}{dt} \bar{x}(\rho) \\
y_\delta &= C(p) x_\delta + D(p) u_\delta.
\end{align*}
\]

(2.12)

Compared to the standard form (2.4) of an LPV system, an additional term \( \frac{d}{dt} \bar{x}(\rho) \) is present and acts as a disturbance on the state equation. This term, however, vanishes when the system is evaluated for a specific constant operating condition, which is the essential difference to the inhomogenous form (2.8). Figure 2.1 and Example 2.1 illustrate how this additional term excites the dynamics in the deviation coordinates when the trim condition \( \bar{x}(\rho) \) is varied.

The additional term is usually neglected when LPV models are obtained through linearization. Doing so translates to the assumption that the variation of the equilibrium, which is not necessarily the same as the variation of the scheduling parameter, is sufficiently slow. It appears to be customary in the literature to not mention this assumption. In fact, there appears to be no account at all of this issue other than a remark by Packard & Kantner [1996] and the recent paper by Takarics & Seiler [2015].
Example 2.1. A nonlinear mass-spring-damper system is used to illustrate the relevance of the parameter-dependent deviation coordinate system. The model is taken from Annoni [2016] and represents the interconnection of 50 blocks with mass \( m = 1 \) kg, each connected both to their neighboring blocks and the initial system by a linear damper with damping constant \( d = 1 \) Ns/m and a nonlinear spring with stiffness \( k(q) = k_1 + k_2 q^2 \), \( k_1 = 0.5 \) N/m, \( k_2 = 1 \) N/m³. An illustration is given in Figure 2.2.

An external force \( \rho \) and a controlled force \( u \) act on the 50th block. The force \( \rho \) is a preload that results in different equilibrium positions. It is taken as the exogenous scheduling parameter for the system. The equations of motion for the \( i \)th block in terms of its displacement \( q_i \) from the equilibrium are

\[
m \ddot{q}_i = \begin{cases} -F_1 - F_{1,2}, & i = 1 \\ -F_i - F_{i,i-1} - F_{i,i+1}, & i = 2, \ldots, 49 \\ -F_{50,49} - F_{50} + \rho + u & i = 50 \end{cases}
\]  

(2.13)

The force \( F_{i,j} = d (\dot{q}_i - \dot{q}_j) + k_1 (q_i - q_j) + k_2 (q_i - q_j)^3 \) is caused by the relative motion of neighboring blocks and \( F_i = d \ddot{q}_i + k(q_i) \dot{q}_i \) is due to the connection with the initial
system. The state vector is $[q_1, \ldots, q_{50}, \dot{q}_1, \ldots, \dot{q}_{50}]^T$ and the output $y$ is the displacement $q_{50}$. The parameter range is restricted to $\mathcal{P} = [0, 2]$ and the system is linearized on a grid $\{\rho_k\}_{k=1}^3 = \{0, 1, 2\}$. Figure 2.3 shows a simulation of the original nonlinear system and LPV approximations.

While the LPV model (2.12) which retains the parameter-dependent coordinate transformation is related to the nonlinear system through the inverse transformation $x \approx \bar{x}(\rho) + \delta x$, this is no longer true when the equilibrium variation $\frac{d}{dt}\bar{x}(\rho(t))$ is neglected. The resulting model in this case is not able to capture the behavior of the nonlinear dynamic system sufficiently. In particular, the output of the LPV model in this case simply equals the trim output $\bar{y}$ for $t < 25s$ where no external input $u$ is applied.

Another way of obtaining LPV models was introduced by Leith & Leithead [1998, 1999, 2000a] and termed velocity-based linearization. They differentiate the state space equations (2.5) of the nonlinear dynamic system with respect to time and perform a change of variables $x_D := \dot{x}$ to obtain the quasi-linear form

$$
\begin{align*}
\dot{x} &= x_D \\
\dot{x}_D &= \frac{\partial f}{\partial x}(x, u, \sigma) x_D + \frac{\partial f}{\partial u}(x, u, \sigma) \dot{u} + \frac{\partial f}{\partial \sigma}(x, u, \sigma) \dot{\sigma} \\
\dot{y} &= \frac{\partial h}{\partial x}(x, u, \sigma) x_D + \frac{\partial h}{\partial u}(x, u, \sigma) \dot{u} + \frac{\partial h}{\partial \sigma}(x, u, \sigma) \dot{\sigma}.
\end{align*}
$$

(2.14)

The system (2.14) is still an exact representation of the original nonlinear system (2.5), but it can also be interpreted as an LPV system with scheduling parameters $(x, u, \sigma)$. For
any realistic application, this scheduling parameter needs to be approximated by a finite grid of operating conditions. A natural choice for such an approximation is the equilibrium manifold (2.6), which results in an LPV system

\[
\begin{align*}
\dot{x}_D &= A(\rho) x_D + B(\rho) \dot{u} + \frac{\partial f}{\partial \sigma} \frac{\partial \sigma}{\partial \rho} \dot{\rho} \\
\dot{y} &= C(\rho) x_D + D(\rho) \dot{u} + \frac{\partial h}{\partial \sigma} \frac{\partial \sigma}{\partial \rho} \dot{\rho}.
\end{align*}
\] (2.15)

This model closely resembles the representation (2.12) obtained from Jacobian linearization and in fact has the exact same state space matrices \((A, B, C, D)\). A valid interpretation is that of a model for differential (infinitesimal small) deviations \(y_\delta \approx \dot{y} \, dt\) and \(u_\delta \approx \dot{u} \, dt\). Given that both modeling approaches result in the same gridded state space representation, conversely, any gridded LPV model can be interpreted as either (2.12) or (2.15).²

For control systems design, the important implication of using models which were obtained through linearization is “that linear controllers...are designed to operate on the perturbations of the plant’s inputs and outputs about the equilibrium points” [Kaminer et al. 1995]. This fact has long been recognized in the gain-scheduling literature and much effort was put into deriving conditions to assure that a controller designed for the linearizations (i.e., deviation variables) also performs reasonably well on the nonlinear plant (i.e., absolute variables) [e.g., Shamma 1988, Kaminer et al. 1995, Lawrence & Rugh 1995, Leith & Leithead 1998, Mehendale & Grigoriadis 2006]. In particular, one important condition for this to be achieved is called local linear equivalence and requires that the linearization of the controller about any constant operating point matches the linearly designed controller for that operating point [Lawrence & Rugh 1995, Kaminer et al. 1995, Khalil 2001, Section 12.5]. This condition is naturally fulfilled for all LPV controllers that are scheduled on external signals, as the linearization of the plant and the controller coincide with the evaluation of the respective LPV systems at that particular operating condition.

²There is also an immediate connection to the so-called velocity or \(D\) implementation of gain-scheduled controllers, which implements a controller designed to operate on \(y_\delta\) and \(u_\delta\) such that it instead operates on \(\dot{y}\) and \(\dot{u}\) [Kaminer et al. 1995, cf. Lawrence 2001, Mehendale 2004, Ch. 4, Mehendale & Grigoriadis 2004, 2006.]
2.2 Preliminary Concepts

Analysis and synthesis techniques for LPV systems require the solution of convex optimization problems that are described mathematically by LMIs [e.g., Boyd et al. 1994, Cha. 2]. An LMI is an expression of the form

\[ F(p) := F_0 + \sum_{i=1}^{n} p_i F_i \prec 0. \]  

(2.16)

The vector \( p \in \mathbb{R}^n \) is called the vector of decision variables for the given symmetric matrices \( F_i = F_i^T \in \mathbb{R}^{m \times m}, i = 1, \ldots, n \). The LMI (2.16) is a convex constraint on \( p \), i.e., the set \( \{ p \mid F(p) \prec 0 \} \) is convex. The compare symbol \( \prec \) is used to denote that \( F(p) \) is symmetric negative definite. Similarly, the symbols \( \preceq \), \( \succ \), and \( \succeq \) are used to denote the following relations:

\[ F(p) \prec 0 \quad \overset{\text{def}}{\iff} \quad F(p) = F^T(p) \quad \text{and} \quad u^T F(p) u < 0 \quad \forall u \in \mathbb{R}^m \setminus \{0\} \]

\[ F(p) \preceq 0 \quad \overset{\text{def}}{\iff} \quad F(p) = F^T(p) \quad \text{and} \quad u^T F(p) u \leq 0 \quad \forall u \in \mathbb{R}^m \]

\[ F(p) \succ 0 \quad \overset{\text{def}}{\iff} \quad F(p) = F^T(p) \quad \text{and} \quad u^T F(p) u > 0 \quad \forall u \in \mathbb{R}^m \setminus \{0\} \]

\[ F(p) \succeq 0 \quad \overset{\text{def}}{\iff} \quad F(p) = F^T(p) \quad \text{and} \quad u^T F(p) u \geq 0 \quad \forall u \in \mathbb{R}^m \]

Further, in accordance with common abuse of notation,

\[ F(p) \prec G(p) \quad \overset{\text{def}}{\iff} \quad F(p) - G(p) \prec 0 \quad \iff \quad G(p) - F(p) \succ 0. \]

2.2.1 Fundamental Properties

The parameter variation in LPV systems can cause instability even if the system is stable for fixed parameters [e.g., Apkarian et al. 1996, Sec. 8.6]. Assessing stability for LPV systems thus requires the use of Lyapunov stability theory [e.g., Khalil 2001, Cha. 4]. A sufficient condition for global exponential stability is formulated in Theorem 2.1.

**Theorem 2.1** (Parameter-dependent Stability [Wu 1995, Lemma 3.2.1, p. 59, Wood 1995, Definition 2.4.6, p. 14]). An autonomous LPV system \( \dot{x} = A(\rho)x \) defined on the domain \( T \) is globally exponentially stable if there exists a continuously differentiable symmetric positive definite matrix function \( X: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x} \) such that \( \forall (p,q) \in \mathcal{P} \times \mathcal{Q} \)

\[ A^T(p) X(p) + X(p) A(p) + \sum_{i=1}^{n_\rho} \frac{\partial X}{\partial \rho_i} q_i < 0. \]  

(2.17)

**Proof.** Define \( a := \inf_{p \in \mathcal{P}} \lambda_{\min}(X(p)) \) and \( b := \sup_{p \in \mathcal{P}} \lambda_{\max}(X(p)) \). It follows that \( a > 0 \) and \( b < \infty \) by the compactness of \( \mathcal{P} \). Thus, a lower and upper bound on \( X(p) \) is given by

\[ a I \preceq X(p) \preceq b I. \]  

(2.18a)
Further, there exists a sufficiently small positive constant $c$ such that

$$A^T(p) X(p) + X(p) A(p) + \sum_{i=1}^{n_p} \frac{\partial X}{\partial \rho_i} \bigg|_p q_i \preceq -c I.$$  \hspace{1cm} (2.18b)

Left and right multiplication of (2.18a) and (2.18b) by $x^T$ and $x$ shows that

$$V(x, p) := x^T(t) X(p) x(t)$$  \hspace{1cm} (2.19a)

satisfies the conditions for $V$ to be a Lyapunov function [cf. Slotine & Li 1991, Theorem 4.1, pp. 107, Khalil 2001, Theorem 4.8, pp. 151]. In particular, when evaluated along any state trajectory that satisfies the differential equation $\dot{x} = A(\rho) x$ of the LPV system and along any parameter trajectory $\rho \in \mathcal{T}$, the time derivative of $V$ is

$$\dot{V} = x^T(t) \left( A^T(p) X(p) + X(p) A(p) + \sum_{i=1}^{n_p} \frac{\partial X}{\partial \rho_i} \bigg|_p q_i \right) x(t).$$  \hspace{1cm} (2.19b)

Hence, the inequalities (2.18a) and (2.18b) imply

$$a \|x(t)\|^2 \leq V(x(t), \rho(t)) \leq b \|x(t)\|^2 \hspace{1cm} \forall t,$$  \hspace{1cm} (2.20a)

$$\dot{V}(x(t), \rho(t), \dot{\rho}(t)) \leq -c \|x(t)\|^2 \hspace{1cm} \forall t.$$  \hspace{1cm} (2.20b)

Multiplying (2.20b) by $e^{c \tau}$ and integrating over the time interval $[t, t+T]$ yields

$$e^{c \tau} V(x(\tau), \rho(\tau)) \bigg|_{t}^{t+T} - \frac{c}{b} \int_{t}^{t+T} e^{c \tau} V(x(\tau), \rho(\tau)) \, d\tau \leq -c \int_{t}^{t+T} e^{c \tau} \|x(\tau)\|^2 \, d\tau.$$  

Replacing $V$ in the integral with the upper bound (2.20a) results in

$$e^{c \tau} V(x(\tau), \rho(\tau)) \bigg|_{t}^{t+T} \leq \frac{c}{b} \int_{t}^{t+T} e^{c \tau} b \|x(\tau)\|^2 \, d\tau - c \int_{t}^{t+T} e^{c \tau} \|x(\tau)\|^2 \, d\tau.$$  

Expanding the term

$$e^{c \tau} V(x(\tau), \rho(\tau)) \bigg|_{t}^{t+T} = e^{c(t+T)} V(x(t+T), \rho(t+T)) - e^{c t} V(x(t), \rho(t)) \leq 0$$

then implies

$$V(x(t+T), \rho(t+T)) \leq e^{-c T} V(x(t), \rho(t)).$$  \hspace{1cm} (2.21)

Finally, using (2.20a), it follows

$$\|x(t+T)\|^2 \leq \frac{b}{a} e^{-c T} \|x(t)\|^2.$$  \hspace{1cm} (2.22)

Inequality (2.22) is the property known as exponential stability [e.g., Khalil 2001, Definition 4.5, p. 150].
2.2 Preliminary Concepts

Theorem 2.1 is a sufficient (parameter-dependent) stability condition for LPV systems that also incorporates bounds on the parameter rates in order to find a certificate for stability. An LPV system whose stability can be certified with a parameter-independent matrix $X$ is called quadratically stable [Becker et al. 1993, Boyd et al. 1994, Sec. 5.1]. In this case, Theorem 2.1 holds for arbitrary fast parameter rates $\dot{\rho}$. For LTI systems, quadratic stability is both necessary and sufficient for exponential stability [e.g., Khalil 2001, Theorem 4.6, pp. 136].

Two other important properties of LPV systems are observability and controllability.

Definition 2.3 (Controllability and Observability). An LPV system with state space representation

$$\dot{x} = A(\rho)x + B(\rho)u, \quad y = C(\rho)x + D(\rho)u$$

and domain $\mathcal{T}$ is controllable if there exists a continuously differentiable symmetric positive definite matrix function $X_c: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x}$ such that

$$\forall (p,q) \in \mathcal{P} \times \mathcal{Q},\quad A(p)X_c(p) + X_c(p)A^T(p) - \sum_{i=1}^{n_\rho} \frac{\partial X_c}{\partial \rho_i} \bigg|_p q_i + B(p)B^T(p) \prec 0. \quad (2.23a)$$

The matrix function $X_c$ is called a generalized controllability Gramian of the system.

The LPV system is called observable if there exists a continuously differentiable symmetric positive definite matrix function $X_o: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x}$ such that

$$\forall (p,q) \in \mathcal{P} \times \mathcal{Q},\quad A^T(p)X_o(p) + X_o(p)A(p) + \sum_{i=1}^{n_\rho} \frac{\partial X_o}{\partial \rho_i} \bigg|_p q_i + C(p)C^T(p) \prec 0. \quad (2.23b)$$

The matrix function $X_o$ is called a generalized observability Gramian of the system. ▲

The interpretation of controllability and observability is the same as for LTI systems. In particular, all state variables of the state space representation are affected by inputs acting on the plant and contribute to the output of the plant [Kailath 1980, Sec. 2.3, Wood 1995, Sec. 2.4]. If the inequalities in Definition 2.3 can be fulfilled with parameter-independent generalized Gramians $X_c$ and $X_o$, the system is called quadratically controllable and quadratically observable.

2.2.2 Induced $L_2$-norm

The induced $L_2$-norm is used for the characterization of performance in this thesis. It provides an upper bound on the energy of a system’s response when excited by an external finite energy signal.

Definition 2.4. (Induced $L_2$-norm of LPV Systems [Wu 1995, Def. 3.3.1, p. 63]) The induced $L_2$-norm of an LPV system $P_\rho$ with input $u$, output $y$, and zero initial conditions which is defined on the domain $\mathcal{T}$ is

$$\|P_\rho\|_2 = \sup_{\rho \in \mathcal{T}} \|y\|_2 = \sup_{\rho \in \mathcal{T}} \sqrt{\int_0^\infty y^T(\tau) y(\tau) \, d\tau} = \sup_{\rho \in \mathcal{T}} \sqrt{\int_0^\infty u^T(\tau) u(\tau) \, d\tau}. \quad (2.24)$$
It measures the maximum gain of the LPV system $P_p$, i.e., the largest amplification of finite energy input signals over all admissible parameter trajectories.

For LTI systems, the induced $L_2$-norm coincides with the $H_\infty$-norm.

**Definition 2.5 ($H_\infty$-norm, [e.g., Khalil 2001, Theorem 5.4, p. 210]).** The $H_\infty$-norm of an LTI dynamic system $P(s)$ with input $u$ and output $y$, and zero initial conditions is

$$\|P(s)\| = \sup_{u \in L_2(\{0\})} \frac{\|y\|_2}{\|u\|_2} = \sup_{\omega} \sigma_{\text{max}}(P(j\omega))$$ \hspace{1cm} (2.25a)

where $\sigma_{\text{max}}(\cdot)$ denotes the largest singular value.

For input signals with an infinite $L_2$-norm, the induced $L_2$-norm can further be defined as the ratio of the signal-rms-values [Werner 2011], i.e.,

$$\|P\| = \sup_{u \in L_2(\{0\}), \rho \in A} \frac{\|y\|_{\text{rms}}}{\|u\|_{\text{rms}}} = \sup_{u \in L_2(\{0\}), \rho \in A} \lim_{T \to \infty} \frac{1}{T} \int_0^T y^T(\tau) y(\tau) \, d\tau$$ \hspace{1cm} (2.26)

This slight abuse of nomenclature is justified to include practically relevant infinite energy signals such as sinusoids.

Just as for stability, assessing the induced $L_2$-norm of an LPV systems requires to find a certificate. The certificate in this case is a storage function [Willems 1972a, b]. A sufficient condition in terms of an LMI is formulated in Theorem 2.2.

**Theorem 2.2 (Induced $L_2$-norm Bound for LPV Systems [Wu 1995, Theorem 3.3.1 and Corollary 3.3.1, pp. 64]).** An LPV system with a state space representation $\dot{x} = A(\rho) x + B(\rho) u$, $y = C(\rho) x + D(\rho) u$ and domain $T$ has an induced $L_2$-norm less than or equal to $\gamma$ and is exponentially stable if there exists a continuously differentiable symmetric positive definite matrix function $X: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x}$ such that $\forall (p,q) \in \mathcal{P} \times \mathcal{Q}$

$$
\begin{bmatrix}
A^T(p) X(p) + X(p) A(p) + \sum_{i=1}^{n_x} \frac{\partial X}{\partial \rho_i} q_i & X(p) B(p) \\
B^T(p) X(p) & -I
\end{bmatrix} + \gamma^{-2} \begin{bmatrix}
C^T(p) \\
D^T(p)
\end{bmatrix} \begin{bmatrix}
C(p) & D(p)
\end{bmatrix} < 0.
$$

(2.27)

**Proof.** Define $a := \inf_{p \in \mathcal{P}} \lambda_{\text{min}}(X(p))$ and $b := \sup_{p \in \mathcal{P}} \lambda_{\text{max}}(X(p))$. It follows that $a > 0$ and $b < \infty$ by the compactness of $\mathcal{P}$. Thus, a lower and upper bound on $X(p)$ is given by

$$a I \preceq X(p) \preceq b I.$$ \hspace{1cm} (2.28a)

Left and right multiplication of (2.27) with $[x^T \hspace{1cm} u^T]$ and $[x]$ respectively, yields the inequality

$$
(A(p) x + B(p) u)^T X(p) x + x^T X(p) (A(p) x + B(p) u) + x^T \left(\sum_{i=1}^{n_x} \frac{\partial X}{\partial \rho_i} q_i\right) x
$$

$$< u^T u - \gamma^{-2} (C(p) x + D(p) u)^T (C(p) x + D(p) u).$$ \hspace{1cm} (2.28b)
2.2 Preliminary Concepts

Left and right multiplication of (2.28a) with by $x^T$ and $x$, respectively, then shows that

$$V(x, p) := x^T X(p) x$$  \hspace{1cm} (2.29a)

is a storage function with supply rate $\|u(t)\|^2 - \gamma^{-2}\|y(t)\|^2$ [cf. Khalil 2001, Theorem 5.1, p. 202 and Theorem 5.5, p. 211]. In particular, when evaluated along any state trajectory $x$ that satisfies the differential equation $\dot{x} = A(\rho) x + B(\rho) u$ of the LPV system and along any parameter trajectory $\rho \in \mathcal{T}$, the time derivative of $V$ is

$$\dot{V} = (A(p) x + B(p) u)^T X(p) x + x^T X(p) (A(p) x + B(p) u) + x^T \left( \sum_{i=1}^{n_p} \frac{\partial X}{\partial \rho_i} \bigg|_{p=\rho} q_i \right) x.$$  \hspace{1cm} (2.29b)

Hence, the inequalities (2.28a) and (2.28b) imply

$$a \|x(t)\|^2 \leq V(x(t), \rho(t)) \leq b \|x(t)\|^2 \quad \forall t,$$  \hspace{1cm} (2.30a)

$$\dot{V}(x(t), \rho(t), \dot{\rho}(t)) < \|u(t)\|^2 - \gamma^{-2}\|y(t)\|^2 \quad \forall t.$$  \hspace{1cm} (2.30b)

Integrating $\dot{V}$ over the time interval $[0, t]$ with $x(0) = 0$ yields

$$V(x(t), \rho(t)) < \int_0^t \|u(\tau)\|^2\,d\tau - \gamma^{-2} \int_0^t \|y(\tau)\|^2\,d\tau.$$  \hspace{1cm} (2.31)

Using that $V(x(t), \rho(t)) \geq 0$, this implies

$$\gamma^{-2} \int_0^t \|y(\tau)\|^2\,d\tau < \int_0^t \|u(\tau)\|^2\,d\tau.$$  \hspace{1cm} (2.32)

Taking the square roots and the limit $t \to \infty$ then yields

$$\sqrt{\int_0^\infty \|y(\tau)\|^2\,d\tau} \leq \gamma \sqrt{\int_0^\infty \|u(\tau)\|^2\,d\tau}.$$  \hspace{1cm} (2.33)

Consequently, $\gamma$ is a bound on the induced $L_2$-norm of Definition 2.4. Further, the LMI (2.27) also immediately implies

$$A^T(p) X(p) + X(p) A(p) + \sum_{i=1}^{n_p} \frac{\partial X}{\partial \rho_i} \bigg|_{p=\rho} q_i + \gamma^{-2} C^T(p) C(p) \succ 0,$$  \hspace{1cm} (2.34)

which by Theorem 2.1 establishes exponential stability for the unforced system.

For a constant matrix $X$, the *Bounded-Real Lemma* [e. g., Boyd et al. 1994, Sec. 2.7.3], necessary and sufficient for LTI systems, is recovered.
3 Model Order Reduction for LPV Systems

This chapter starts by providing an overview of model order reduction techniques for LPV systems and their limitations. Two novel methods to reduce the number of state variables of LPV systems in grid representation are then developed. One is rooted in the paradigm of local reduction and interpolation but provides a way to achieve state space consistency by exploiting modal forms. It can be combined with any available LTI model order reduction technique. The second method performs global reduction but uses local information and a partially parameter-varying projection. This projection has a clear interpretation in terms of a varying test space in the Petrov-Galerkin approximation of dynamic systems and is shown to avoid additional rate dependence. This extension of the projection framework relates very naturally to balanced truncation and related approaches which employ Gramian-based information to quantify the importance of subspaces for model order reduction. The two methods can be used to obtain low-order models for LPV analysis and synthesis. They thus help to increase the applicability of LPV control techniques to models that are otherwise out of the scope due to computational complexity. Parts of this chapter were pre-published as research papers by Theis, Takarics, Pfifer, Balas & Werner [2015c] and Theis, Seiler & Werner [2016b, 2018].

3.1 Preliminary Model Order Reduction Concepts

The problem of LPV model order reduction consists in finding an approximation

\[
P_{\rho,\text{red}}: \begin{cases} \dot{z}(t) = A_{\text{red}}(\rho(t)) z(t) + B_{\text{red}}(\rho(t)) u(t) \\ y(t) = C_{\text{red}}(\rho(t)) z(t) + D_{\text{red}}(\rho(t)) u(t) \end{cases} \tag{3.1}
\]

for the dynamic system

\[
P_{\rho}: \begin{cases} \dot{x}(t) = A(\rho(t)) x(t) + B(\rho(t)) u(t) \\ y(t) = C(\rho(t)) x(t) + D(\rho(t)) u(t) \end{cases} \tag{3.2}
\]

The reduced state vector \( z(t) \in \mathbb{R}^{n_z} \) should be of much lower dimension than \( x(t) \in \mathbb{R}^{n_x} \), while the input-output behavior \( u \to y \) should be as similar as possible to that of the original model. Further, stability of the original model should be preserved in the reduced-order model.
### 3.1.1 State Space Manipulations

This section briefly reviews elementary state space manipulations which are used in model order reduction. For an LTI change of state space coordinates \( \xi = T x \) with a nonsingular transformation matrix \( T \in \mathbb{R}^{n_x \times n_x} \), the LPV state space representation (3.2) becomes

\[
\begin{align*}
\dot{\xi} &= \bar{A}(\rho) \xi + \bar{B}(\rho) u \\
y &= C(\rho) \xi + D(\rho) u.
\end{align*}
\]  

(3.3)

The important property of the transformation is that, due to time invariance, \( \xi = T x \Leftrightarrow \dot{\xi} = T \dot{x} \). The use of an LPV transformation to change the state space coordinates to \( \xi = T(\rho) x \), on the other hand, leads to

\[
\begin{align*}
\dot{\xi} &= T(\rho) \bar{A}(\rho) T^{-1}(\rho) \xi + \sum_{i=1}^{n_p} \frac{\partial T(\rho)}{\partial \rho_i} T^{-1}(\rho) \dot{\rho}_i \\
y &= C(\rho) T^{-1}(\rho) \xi + D(\rho) u.
\end{align*}
\]  

(3.4)

This additional rate dependence is a realization artifact, attributed to expressing the state space model in parameter-dependent coordinates: Clearly, starting from the rate dependent model (3.4), the transformation \( x = T^{-1}(\rho) \xi \) leads back to the rate-independent realization (3.2). In fact, any affine rate dependence can be removed through a suitable transformation [Wood 1995, p. 144]: If \( \bar{A}(\rho, \dot{\rho}) = \bar{A}_0(\rho) + \sum_{i=1}^{n_p} \bar{A}_i(\rho) \dot{\rho}_i \) and a parameter-dependent transformation \( \bar{T}(\rho) \) is performed, the resulting state space matrix becomes

\[
\bar{A}(\rho, \dot{\rho}) = \bar{T}(\rho) \bar{A}_0(\rho) \bar{T}^{-1}(\rho) + \sum_{i=1}^{n_p} \left( \bar{T}(\rho) \bar{A}_i(\rho) \bar{T}^{-1}(\rho) + \frac{\partial \bar{T}(\rho)}{\partial \rho_i} \bar{T}^{-1}(\rho) \right) \dot{\rho}_i.
\]

Dependence on the rate \( \dot{\rho}_i \) is therefore eliminated if the transformation \( \bar{T}(\rho) \) is selected to satisfy \( \frac{\partial \bar{T}(\rho)}{\partial \rho_i} = -\bar{T}(\rho) \bar{A}_i(\rho) \). Unfortunately, the construction of a transformation that satisfies such a requirement is in practice not easy. The only other possibility is then to treat \( \dot{\rho} \) as a newly introduced independent parameter vector, which effectively doubles the number of scheduling parameters and hence increases model complexity.

The purpose of changing the basis of the state space, either by LTI or LPV transformations, is to bring the dynamic system into a form that permits a partitioning
\[ \dot{\xi}_1 = A_{11} \xi_1 + A_{12} \xi_2 + B_1 u \]
\[ \dot{\xi}_2 = A_{21} \xi_1 + A_{22} \xi_2 + B_2 u \]
\[ y = C_1 \xi_1 + C_2 \xi_2 + D u. \] 

In this form, \( \xi_1 \) represents the subset of state variables that is to be retained in the reduced-order model while \( \xi_2 \) represents state variables that are to be removed. Given the partitioning (3.5), there are essentially two different ways of removing \( \xi_2 \): truncation and residualization.

**Definition 3.1** (Residualization [Liu & Anderson 1989]). Residualization, also known as *singular perturbation approximation*, refers to setting \( \dot{\xi}_2 = 0 \), solving for \( \xi_2 \), and substituting the resulting expression into Equation (3.5). The reduced-order model is
\[ \dot{\xi}_1 = (A_{11} - A_{12} A_{22}^{-1} A_{21}) \xi_1 + (B_1 - A_{12} A_{22}^{-1} B_2) u \]
\[ y = (C_1 - C_2 A_{22}^{-1} A_{21}) \xi_1 + (D - C_2 A_{22}^{-1} B_2) u. \] 

Residualization preserves the steady-state gain \( D - CA^{-1}B \) of the original system and hence retains accuracy “at low frequencies”. Residualized state variables \( \xi_2 \) can be interpreted to immediately attain their steady-state values, which is a good approximation when the corresponding dynamics are very fast compared to those of the state variables \( \xi_1 \) and hence transients are negligible. A classical example for residualization is to remove the state variables that represent pitch rate and angle of attack from an aircraft model that is concerned with the slow altitude forward velocity oscillation known as *phugoid*.

**Definition 3.2** (Truncation [e.g., Moore 1981]). Truncation refers to simply discarding \( \xi_2 \) from the dynamic system (3.5), i.e., the reduced-order model is
\[ \dot{\xi}_1 = A_{11} \xi_1 + B_1 u \]
\[ y = C_1 \xi_1 + D u. \] 

Truncation exactly preserves the feedthrough gain \( D \) and hence the truncated model equals the full-order model at infinite frequency. Model reduction by truncation is therefore preferred when accuracy of the reduced-order model “at high frequencies” is required. The truncated state variables \( \xi_2 \) can be interpreted to remain at the initial value zero for all times, which is a good approximation when the corresponding dynamics are very slow.

1 In fact, both methods can be interpreted as the extreme cases of a *generalized singular perturbation approximation* that seeks to approximate a dynamic system such that its frequency response gain is exactly preserved at a specified frequency \( \sigma_0 \) [Fernando & Nicholson 1982]. Residualization corresponds to \( \sigma_0 = 0 \) and truncation to \( \sigma_0 \to \infty \). This perspective is intimately related to the approximation of dynamic systems by moment matching [Antoulas 2005, Cha. 11].
Model Order Reduction for LPV Systems

compared to those of the state variables $\xi_1$ and hence $\xi_2$ is almost constant during changes in $\xi_1$. A classical example for truncation is to remove the state variables that represent altitude and forward velocity from an aircraft model that is concerned with short period dynamics, i.e., pitch motion.

Another perspective on the problem of removing unwanted state variables is provided by the projection framework [e.g., de Villemagne & Skelton 1987, Saad 2000, Cha. 5].

Definition 3.3 (Projection). A projection is a linear operation $\mathbb{R}^{n_x} \rightarrow \mathcal{V} \subseteq \mathbb{R}^{n_x}$ defined by a matrix

$$\Pi = V (W^T V)^{-1} W^T$$

(3.8)

with $V \in \mathbb{R}^{n_x \times n_z}$, $W \in \mathbb{R}^{n_x \times n_z}$, $n_x > n_z$ and rank($W^T V$) = $n_z$. A matrix $\Pi \in \mathbb{R}^{n_x \times n_x}$ represents a projection if and only if it is idempotent, i.e., $\Pi = \Pi^2$. A projection is completely characterized by its range space, $\text{span}(\Pi) = \text{span}(V)$, and its nullspace, $\text{ker}(\Pi) = \text{span}(\Pi^T)^\perp = \text{span}(W)^\perp$. The range space of a projection

$$\mathcal{V} := \text{span}(\Pi) = \text{span}(V)$$

(3.9)

is called basis space. The subspace orthogonal to a projection’s nullspace

$$\mathcal{W} := \text{ker}(\Pi)^\perp = \text{span}(W)$$

(3.10)

is called test space.

The basic facts associated with Definition 3.3 are easy to prove by replacing $V$ and $W$ with their respective thin QR-factorizations. A vector space is said to be projected by $\Pi$ along the orthogonal complement of the subspace spanned by the columns of $W$ and onto a subspace spanned by the columns of $V$.

Given $V$, $W$, and a point $x \in \mathbb{R}^{n_x}$, the projection $\Pi x$ provides the unique approximation to $x$ in $\mathcal{V}$ with zero error within $\mathcal{W}$. This is illustrated in the following. As the approximation lies in the span of $\mathcal{V}$, it can be written as

$$x_{\text{approx}} := V z$$

(3.11a)

with some coefficient vector $z \in \mathbb{R}^{n_z}$. The basis space $\mathcal{V}$ thus literally forms a basis for the approximation. The component of $x$ that is eliminated by the projection, i.e., the approximation error or residual

$$r := x - x_{\text{approx}} = x - V z$$

(3.11b)

is in the nullspace of $\Pi$. It is hence orthogonal to $W$ which can be expressed as

$$W^T r = 0.$$  

(3.11c)

2A thin QR-factorization is the unique factorization $X = QR$ of a matrix $X \in \mathbb{R}^{n \times m}$ with $n > m$ into a matrix $Q \in \mathbb{R}^{n \times m}$ with orthonormal columns and an upper-triangular matrix $R \in \mathbb{R}^{m \times m}$ with positive diagonal entries [e.g., Golub & Van Loan 2013, Theorem. 5.2.3, p. 248].

\[26\]
3.1 Preliminary Model Order Reduction Concepts

The test space $W$ thus determines the measure of error and “tests” the approximation. Substituting Equation (3.11b) into Equation (3.11c) and solving for $z$ results in

$$z = (W^T V)^{-1} W^T x.$$  
(3.11d)

The inverse in Equation (3.11a) exists according to Definition 3.3 such that the solution is unique. Substituting Equation (3.11d) into Equation (3.11a) finally shows that

$$x_{\text{approx}} = V (W^T V)^{-1} W^T x = \Pi x$$  
(3.11e)

and from Equation (3.11b) it further follows that

$$r = (I - \Pi) x.$$  
(3.11f)

If $\Pi = \Pi^T$, then $\Pi$ is called an orthogonal projection since its nullspace is orthogonal to its range space. In this case, basis and test space coincide, i.e., $W = V$.

**Definition 3.4 (Orthogonal Projection).** An orthogonal projection is a linear operation defined by a matrix

$$\Pi = V (V^T V)^{-1} V^T$$  
(3.12)

with $V \in \mathbb{R}^{n_x \times n_z}$, $n_x > n_z$, and rank$(V^T V) = n_z$. It is completely characterized by its range space $\text{span}(\Pi) = \ker(\Pi)^\perp = \text{span}(V)$.

The approximation error of an orthogonal projection is orthogonal to the approximation, i.e., $V^T (x - V z) = 0$. The unique solution to this equation is the least squares solution $z = (V^T V)^{-1} V^T x$. To make a clear distinction, the general projection of Definition 3.3 is also referred to as an *oblique projection*.

The following theorem simplifies the treatment in this chapter.

**Theorem 3.1 ([de Villemagne & Skelton 1987, Corollary 2.1]).** Any projection can be parameterized by $V \in \mathbb{R}^{n_x \times n_z}$ and a nonsingular square matrix $S \in \mathbb{R}^{n_x \times n_x}$ as

$$\Pi = V \left(V^T S V\right)^{-1} V^T S \overbrace{W^T}^{W^T}.$$  
(3.13)

**Proof.** de Villemagne & Skelton [1987] provide a proof based on generalized inverses. □

Any $W$ constructed in this way is biorthogonal to $V$, i.e., $W^T V = I_{n_z}$. Thus, biorthogonality of $V$ and $W$ can always be assumed without loss of generality.  

Model order reduction requires the approximation of a dynamic system given by a differential equation, rather than an approximation for a single point in the state space.

---

3Biorthogonality can also be achieved by reassigning $W^T \leftarrow (W^T V)^{-1} W^T$ in Definition 3.3.
The goal is thus to find an approximate solution $x_{\text{approx}} := V z$ to the differential equation in (3.2), i.e.

$$\dot{x}_{\text{approx}} \approx A(\rho) x_{\text{approx}} + B(\rho) u.$$  

(3.14)

The residual of this approximation is

$$r := V \dot{z} - (A(\rho) V z + B(\rho) u).$$  

(3.15)

Restricting the residual (3.15) again to be orthogonal to the test space $W$ leads to

$$W^T (V \dot{z} - (A(\rho) V z + B(\rho) u)) = 0.$$  

(3.16)

The unique solution is

$$\dot{z} = (W^T V)^{-1} W^T A(\rho) V z + (W^T V)^{-1} W^T B(\rho) u.$$  

(3.17)

The approximation $x_{\text{approx}} = V z$ where $z$ is the solution to Equation (3.17) is known as Petrov-Galerkin approximation [e.g., Antoulas 2005, Sec. 9.1.2]. Augmenting Equation (3.17) with an output equation $y = C(\rho) x_{\text{approx}} + D(\rho) u$ then yields the reduced-order model of Definition 3.5. In view of Theorem 3.1, the following simpler definition is used.

**Definition 3.5 (Petrov-Galerkin Approximation).** Let an LPV system with state space representation $\dot{x} = A(\rho) x + B(\rho) u$, $y = C(\rho) x + D(\rho) u$, and matrices $V \in \mathbb{R}^{n_x \times n_z}$, $W \in \mathbb{R}^{n_x \times n_z}$ with $W^T V = I_{n_z}$ be given. The reduced-order model from Petrov-Galerkin approximation with basis space $V = \text{span}(V)$ and test space $W = \text{span}(W)$ is

$$\dot{z} = A_{\text{red}}(\rho) V z + W^T B(\rho) u$$

$$y = C_{\text{red}}(\rho) V z + D(\rho) u.$$  

(3.18)

Comparing the approximation (3.18) with the reduced-order model (3.7) which is obtained through transformation and subsequent truncation, it is clear that both are equivalent with $V = T^{-1} \left[ I_{n_z} 0_{n_z \times (n_x-n_z)} \right]^T$, $W = T \left[ I_{n_z} 0_{n_z \times (n_x-n_z)} \right]$. Truncation is thus an orthogonal projection $\Pi_{\text{truncation}} = \left[ \begin{array}{c} I_{n_x} \\ 0 \end{array} \right]$. A coordinate transformation followed by truncation is an oblique projection. Residualization, on the other hand, cannot be expressed as a projection, as is immediately apparent from the modified $D$-matrix of the system (3.6).

For LPV systems, it seems natural to let $V$ and $W$ also depend on the scheduling parameter, just as in the case of parameter-varying state transformations. Repeating the Petrov-Galerkin approximation with parameter-dependent matrices $V(\rho(t))$ and $W(\rho(t))$ yields

$$\dot{x}_{\text{approx}} = \frac{d}{dt} \left( V(\rho) z \right) = V(\rho) \dot{z} + \sum_{i=1}^{n_p} \frac{\partial V(\rho)}{\partial \rho_i} \dot{\rho}_i z.$$  

(3.19)
The residual of this approximation is hence
\[ r := V(\rho) \dot{z} + \sum_{i=1}^{n_\rho} \frac{\partial V(\rho)}{\partial \rho_i} \dot{\rho}_i z - (A(\rho) V(\rho) z + B(\rho) u). \tag{3.20} \]

The reduced-order LPV model obtained by enforcing the orthogonality constraint \( W^T(\rho) r = 0 \) is
\[ \dot{z} = A_{\text{red}}(\rho, \dot{\rho}) z + B_{\text{red}}(\rho) u \tag{3.21} \]
y explicitly depends on the parameter rate \( \dot{\rho} \) in addition to the original parameter \( \rho \).

### 3.1.2 State-of-the-Art Method: LPV Balancing and Truncation

The state-of-the-art method for LPV model order reduction is LPV balancing and truncation [Wood 1995, Ch. 7, Wood et al. 1996]. It is based on transforming the system into a state space representation where state variables can be related to their contribution to the input-output behavior in terms of the induced \( L_2 \)-norm. This is achieved by first calculating generalized Gramians \(^4\) from Definition 2.3, i.e., to find symmetric positive definite matrix functions \( X_c : \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x} \) and \( X_o : \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x} \) that \( \forall (p,q) \in \mathcal{P} \times \mathcal{Q} \) satisfy the LMIs
\[ A(p) X_c(p) + X_c(p) A^T(p) - \sum_{i=1}^{n_\rho} \frac{\partial X_c(p)}{\partial \rho_i} \bigg|_{p} q_i + B(p) B^T(p) \prec 0, \tag{3.22a} \]
\[ A^T(p) X_o(p) + X_o(p) A(p) + \sum_{i=1}^{n_\rho} \frac{\partial X_o(p)}{\partial \rho_i} \bigg|_{p} q_i + C^T(p) C(p) \prec 0. \tag{3.22b} \]

The generalized Gramians establish (parameter-dependent) bounds on the energy in the output from a response to an initial condition and on the input energy required to reach a certain point in the state space. Given a point \( x_0 \), the minimum energy required to steer the system from \( x = 0 \) to \( x = x_0 \) is lower bounded by \( \epsilon_c := x_0^T X_c^{-1} x_0 \) [Wood 1995, Theorem 7.3.3, p. 137]. Further, \( \epsilon_o := x_0^T X_o x_0 \) is an upper bound on the maximum energy of the free response to the initial condition \( x_0 \) [Wood 1995, Theorem 7.3.2, p. 136]. It is important to emphasize that Gramians correspond to a particular state space realization and that any nonsingular state transformation \( \xi = T x \) leads to different Gramians
\[ \tilde{X}_c = T X_c T^T, \quad \tilde{X}_o = (T^{-1})^T X_o T^{-1}. \tag{3.23} \]

\(^4\) Controllability and observability are necessary for the existence of positive definite generalized Gramians, which requires to start from a minimal realization of the LPV system that is not always easy to find.
Under some technical assumptions [detailed by Wood 1995, Sec. 7.8], it is possible to perform a balancing transformation

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = T(\rho) \begin{bmatrix} x \end{bmatrix} \text{ so that } T(\rho) X_c(\rho) T^T(\rho) = (T^{-1}(\rho))^T X_o(\rho) T^{-1}(\rho) = \Sigma^{1/2}(\rho), \quad (3.24)$$

where the matrix $\Sigma(\rho)$ is diagonal and contains the eigenvalues of the product $X_c(\rho) X_o(\rho)$ ordered by decreasing magnitude along its diagonal. Since these values correspond to the ratios $\epsilon_o/\epsilon_c$ for each state variable in the new coordinates, $\Sigma(\rho)$ is a bound for the contribution of each state variable to the input-output behavior in terms of the induced $L_2$-norm. Partitioning the state vector such that $\xi_1$ contains the state variables corresponding to large entries of $\Sigma(\rho)$ therefore collects highly controllable and observable state variables. The state variables in $\xi_2$ then contribute little to the input-output behavior and are removed from the state vector by truncation or residualization. The reduced-order model is guaranteed to be stable and satisfies an a-priori error bound in the induced $L_2$-norm [Wood 1995, Cha. 7].

LPV balancing and truncation is a direct generalization of the LTI counterpart, proposed first by Moore [1981] and further investigated by Pernebo & Silverman [1982]. For LTI systems, the Lyapunov inequalities (3.22) simplify to the Lyapunov equations

$$AX_c + X_c A^T + BB^T = 0,$$

$$A^T X_o + X_o A + C^T C = 0.$$  

(3.25a)  

(3.25b)

In this case, the Gramians are unique and the bounds are tight, i.e., the minimum energy required to steer the system from $x = 0$ to $x = x_0$ is $\epsilon_c := x_0^T X_c^{-1} x_0$ and the maximum energy of the free response to the initial condition $x_0$ is $\epsilon_o := x_0^T X_o x_0$ [Moore 1981]. The ratio $\epsilon_o/\epsilon_c$ thus precisely measures how much each state variable is affected by the input and how much it affects the output.

**Gramians for Unstable Systems**

Gramians and their generalizations exist only for exponentially stable systems, as the existence of Gramians implies the existence of a symmetric positive definite matrix $X$ that $\forall p \in P$ satisfies

$$A^T(p) X(p) + X(p) A(p) - \sum_{i=1}^n \left. \frac{\partial X}{\partial \rho} \right|_{q_i} q_i < 0.$$  

Such a matrix implies the existence of a parameter-independent Lyapunov function $V(x, \rho) = x^T X(\rho) x$ in accordance with Theorem 2.1. The approach is hence not suitable for models that include unstable dynamics. A popular remedy is provided by a coprime factorization, which represents an unstable system as two stable input-output pairs [e.g., Vidyasagar 1985, Meyer 1990, Prempain 2006]. Generalized Gramians for an LPV system can be calculated as proposed by Wood [1995, Sec. 7.5] and detailed, e.g., by Moreno [2015, Sec. 4.2.3] from a contractive coprime factorization via symmetric positive definite
3.1 Preliminary Model Order Reduction Concepts

Matrix functions $Z_1^{-1}: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x}$ and $Z_2^{-1}: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x}$ that satisfy the LMI

$$
\begin{bmatrix}
Z_1^{-1}(p) \hat{A}^T(p) + \hat{A}(p) Z_1^{-1}(p) - B(p) S^{-1}(p) B^T(p) - \sum_{i=1}^{n_x} \frac{\partial Z_1^{-1}}{\partial p_i} q_i & Z_1^{-1}(p) C^T(p) \\
C(p) Z_1^{-1}(p) & -R(p)
\end{bmatrix} < 0,
$$

(3.26a)

$$
\begin{bmatrix}
Z_2^{-1}(p) \hat{A}(p) + \hat{A}^T(p) Z_2^{-1}(p) - C^T(p) R^{-1}(p) C(p) + \sum_{i=1}^{n_x} \frac{\partial Z_2^{-1}}{\partial p_i} q_i & Z_2^{-1}(p) B(p) \\
B^T(p) Z_2^{-1}(p) & -S(p)
\end{bmatrix} < 0,
$$

(3.26b)

with $\hat{A} := A - B S^{-1} D T C$, $\hat{A} := A - B D T R^{-1} C$, and $S := I + D^T D$, $R := I + D D^T$. A generalized observability Gramian is $X_o = Z_1$ and a generalized controllability Gramian is $X_c = (I + Z_2 Z_1)^{-1} Z_2$ [Glover & McFarlane 1989, Wood 1995, Lemma 7.5.9, p. 150]. This approach is again a direct generalization of the corresponding LTI result which uses a normalized coprime factorization and calculates Gramians from the generalized algebraic Riccati equations

$$
\hat{A}^T Z_1 + Z_1 \hat{A} - Z_1 B S^{-1} B^T Z_1 + C^T R^{-1} C = 0,
$$

(3.27a)

$$
\hat{A} Z_2 + Z_2 \hat{A}^T - Z_2 C^T R^{-1} C Z_2 + B S^{-1} B^T = 0.
$$

(3.27b)

Frequency-weighted Gramians

While simultaneous observability and controllability is a useful metric for model order reduction, many engineering problems require the emphasis of a certain frequency region. This is especially true for control-oriented models, where the available bandwidth of a control system provides a natural upper frequency limit on the fidelity requirement. Similarly, a lower frequency limit is often available from the structure of a control system. When, e.g., inner control loops are to be designed, low-frequency dynamics are often negligible, as cascaded outer loops are added later in the design process. Such information is not captured by controllability and observability, but can be incorporated through frequency weighting. Doing so is possible by augmenting the full-order system with stable, minimum phase LTI weighting filters $\Omega_o(s) = C_\Omega_o(s I - A_\Omega_o)^{-1} B_\Omega_o + D_\Omega_o$ and $\Omega_i(s) = C_\Omega_i(s I - A_\Omega_i)^{-1} B_\Omega_i + D_\Omega_i$ which emphasize a certain frequency range as first proposed by Enns [1984] and adapted for LPV systems by Wood [1995, Sec. 7.4].

Frequency-weighted generalized Gramians $X_c$ and $X_o$ that measure controllability and observability in the frequency range of interest can be obtained as symmetric positive definite matrix functions $X_c: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x}$ and $X_o: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x}$ that $\forall (p, q) \in \mathcal{P} \times \mathcal{Q}$
3 Model Order Reduction for LPV Systems

satisfy the LMIs

\[
\begin{bmatrix}
A(p) & B(p) C_{\Omega} \\
0 & A_{\Omega}
\end{bmatrix}
\begin{bmatrix}
X_c(p) \\
X_o(p)
\end{bmatrix}
+ 
\begin{bmatrix}
A(p) & B(p) C_{\Omega} \\
0 & A_{\Omega}
\end{bmatrix}^T
\begin{bmatrix}
X_c(p) \\
X_o(p)
\end{bmatrix} - 
\begin{bmatrix}
\sum_{i=1}^{n_p} \frac{\partial X_c}{\partial p_i} q_i \\
\sum_{i=1}^{n_p} \frac{\partial X_o}{\partial p_i} q_i
\end{bmatrix} < 0,
\]

(3.28a)

\[
\begin{bmatrix}
A(p) & 0 \\
B_{\Omega}, C(p) & A_{\Omega}
\end{bmatrix}^T
\begin{bmatrix}
X_c(p) \\
X_o(p)
\end{bmatrix} + 
\begin{bmatrix}
A(p) & 0 \\
B_{\Omega}, C(p) & A_{\Omega}
\end{bmatrix}^T
\begin{bmatrix}
X_c(p) \\
X_o(p)
\end{bmatrix} - 
\begin{bmatrix}
\sum_{i=1}^{n_p} \frac{\partial X_c}{\partial p_i} q_i \\
\sum_{i=1}^{n_p} \frac{\partial X_o}{\partial p_i} q_i
\end{bmatrix} < 0.
\]

(3.28b)

The * in (3.28) denotes block matrices that correspond to filter states and are of no interest for model order reduction. Stability of the reduced-order model can still be guaranteed as long as one of the weights is selected as the identity matrix, i.e., either \( \Omega_o = I_{n_y} \) or \( \Omega_o = I_{n_u} \), but an error bound does no longer exist. For LTI systems, the inequalities (3.28) simplify to the equations

\[
\begin{bmatrix}
A & B C_{\Omega} \\
0 & A_{\Omega}
\end{bmatrix}
\begin{bmatrix}
X_c \\
X_o
\end{bmatrix}
+ 
\begin{bmatrix}
A & B C_{\Omega} \\
0 & A_{\Omega}
\end{bmatrix}^T
\begin{bmatrix}
X_c \\
X_o
\end{bmatrix} - 
\begin{bmatrix}
\sum_{i=1}^{n_p} \frac{\partial X_c}{\partial p_i} q_i \\
\sum_{i=1}^{n_p} \frac{\partial X_o}{\partial p_i} q_i
\end{bmatrix} = 0,
\]

(3.29a)

\[
\begin{bmatrix}
A & 0 \\
B_{\Omega}, C & A_{\Omega}
\end{bmatrix}^T
\begin{bmatrix}
X_c \\
X_o
\end{bmatrix} + 
\begin{bmatrix}
A & 0 \\
B_{\Omega}, C & A_{\Omega}
\end{bmatrix}^T
\begin{bmatrix}
X_c \\
X_o
\end{bmatrix} - 
\begin{bmatrix}
\sum_{i=1}^{n_p} \frac{\partial X_c}{\partial p_i} q_i \\
\sum_{i=1}^{n_p} \frac{\partial X_o}{\partial p_i} q_i
\end{bmatrix} = 0.
\]

(3.29b)

Computational Solution of the Model Order Reduction Problem

In order to arrive at a computationally tractable formulation for the LMIs (3.22), the positive definite matrix functions \( X_c : \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x} \) and \( X_o : \mathcal{P} \mapsto \mathbb{R}^{n_o \times n_o} \) must be formulated in terms of a predefined set of basis functions\(^5\) as

\[
X_c(p) = \sum_{i=1}^{a} f_i(p) X_{c,i}, \quad X_c,i \in \mathbb{R}^{n_x \times n_x} \quad \text{and} \quad X_o(p) = \sum_{i=1}^{b} g_i(p) X_{o,i}, \quad X_o,i \in \mathbb{R}^{n_o \times n_o}.
\]

(3.30)

A feasibility problem for the LMIs (3.22) then involves \( n_{dec} = (a + b) n_x (n_z + 1)/2 \) decision variables. Standard algorithms for solving LMIs require a number of floating point operations that scales with \( n_{dec}^3 \) [Boyd & Vandenberghe 2004, Sec. 11.8.3]. The problem thus scales on the order of \( n_x^6 \), which shows that larger state dimensions quickly become

\[5\]The same is true for the respective matrices in LMIs (3.28) in case of frequency-weighted Gramians and LMIs (3.26) in case of coprime Gramians.
3.1 Preliminary Model Order Reduction Concepts

prohibitive. Using such parameter-dependent matrix functions further produces parameter-dependent transformations which consequently results in a reduced-order model with additional parameter rate dependence as described in Section 3.1.1. The state-of-the-art method for balancing of LPV systems therefore uses parameter-independent generalized Gramians [Wood et al. 1996]. Restricting the search space in this manner might decrease the quality of the approximation. Further, the computational effort required to solve the parameter-independent LMIs by numerical methods is still considerable and the method is therefore limited to systems with up to about 50 state variables.

Another important problem is that the generalized Gramians are not unique and hence not equally useful for model order reduction. In order to increase the number of small singular values in the product of the Gramians, Wood et al. [1996] suggest to solve the (non-convex) optimization problem

$$\min_{X_c, X_o} \text{trace}(X_c X_o) \quad \text{subject to}$$

$$X_c \succ 0 \quad (3.31a)$$
$$X_o \succ 0 \quad (3.31b)$$
$$A(p) X_c + X_c A^T(p) + B(p) B^T(p) \prec 0 \quad \forall p \in \mathcal{P} \quad (3.31c)$$
$$A^T(p) X_o + X_o A(p) + C^T(p) C(p) \prec 0 \quad \forall p \in \mathcal{P}. \quad (3.31d)$$

From these Gramians, a parameter-independent balancing coordinate transformation is calculated and used to first partition the state vector such that $$z := \xi_1$$ contains the state variables with large contributions to the input-output map and $$\xi_2$$ those with negligible contributions. In a second step, $$\xi_2$$ is then removed from the state vector by truncation, so that the system (3.1) is obtained. The resulting reduced-order model is guaranteed to be stable and satisfies an error bound in the induced $$L_2$$-norm of twice the sum of the truncated singular values [Wood et al. 1996].

Alternatively, these two steps are combined into a single oblique projection by directly considering the Gramians to define the range space and null space of a projection. A popular way of doing this is known as the square root algorithm [Laub et al. 1987]. It requires the lower Cholesky factorizations \(^6\) of $$X_o = L_o L_o^T$$ and $$X_c = L_c L_c^T$$, as well as the singular value decomposition (SVD) \(^7\) of the product

$$L_c^T L_o = U_1 \begin{bmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_2 & N_1 & N_2 \end{bmatrix}^T. \quad (3.32)$$

The singular values are ordered by descending magnitude, such that the diagonal matrix $$\Sigma_1$$ contains the largest $$n_z$$ singular values. The orthogonal matrices $$[u_1, u_2]$$ and $$[n_1, n_2]$$

\(^6\)A lower Cholesky factorization is the unique factorization $$X = LL^T$$ of a symmetric positive definite matrix $$X \in \mathbb{R}^{n \times n}$$ into two lower-triangular matrices $$L \in \mathbb{R}^{n \times n}$$ with positive diagonal entries [e.g., Golub & Van Loan 2013, Theorem. 4.2.7, p. 163].

\(^7\)A singular value decomposition is the factorization $$X = U \Sigma N^T$$ of a matrix $$X \in \mathbb{R}^{n \times m}$$ into two orthogonal matrices $$U \in \mathbb{R}^{n \times n}$$ and $$N \in \mathbb{R}^{m \times m}$$, and a matrix $$\Sigma \in \mathbb{R}^{n \times m}$$ with positive diagonal entries ordered by decreasing magnitude [e.g., Golub & Van Loan 2013, Theorem. 2.4.1, p. 76]. The notation $$N$$ is chosen here to avoid confusion with the matrix $$V$$ used to denote a basis for the range space of a projection.
contain the corresponding left and right singular vectors. The projection that corresponds to balancing and truncation then is [e.g., Laub et al. 1987, Antoulas 2005, Sec. 7.3]

$$
\Pi_{\text{bal}} = L_c U_1 \Sigma_1^{-1/2} \Sigma_2^{-1/2} N_1^T L_o^T W^T,
$$

(3.33)

That is, \( \mathcal{V} = \text{span}(L_c U_1 \Sigma_1^{-1/2}) \) is the basis space and \( \mathcal{W} = \text{span}(L_o N_1 \Sigma_1^{-1/2}) \) is the test space for a Petrov-Galerkin approximation. It is of interest to note that the singular values in the projection (3.33) do not affect these spaces. They are mere scalings of directions within the respective subspaces \( \mathcal{V} \) and \( \mathcal{W} \). Hence, they only affect the internal representation of the reduced-order model but not its input-output properties. It is thus possible to obtain a reduced-order model that is not internally balanced but which has the exact same input-output behavior as the balanced reduced-order model. This possibility was first described by Varga [1991] and termed balancing-free balancing.

### 3.1.3 Related Problems

As detailed in the previous section, there are two key issues with the state-of-the-art method: 1) the computational limitations associated with solving the LMIs in order to determine a transformation and 2) the additional rate dependence caused by parameter-varying transformations which in practice restricts the search space to parameter-independent Gramians.

#### Gramian Approximations

Remedies to avoid the computational complexity of the LMI solution usually make use of a “local” approximation, i.e., they evaluate the LPV system for a fixed parameter value and apply methods for LTI systems. For a frozen parameter \( \rho \equiv \rho_k \), the solution of the optimization problem (3.31) simplifies to solving the two Lyapunov equations

$$
A(\rho_k) X_c + X_c A^T(\rho_k) + B(\rho_k) B^T(\rho_k) = 0,
$$

(3.34a)

$$
A^T(\rho_k) X_o + X_o A(\rho_k) + C^T(\rho_k) C(\rho_k) = 0.
$$

(3.34b)

This is an immediate consequence of the fact that \( Y - X \prec 0 \) for all \( X \) and \( Y \) that satisfy \( AX + XA^T + BB^T \prec 0 \) and \( AY + YA^T + BB^T = 0 \) [Dullerud & Paganini 2005, Proposition 4.4, p. 134]. That is, for a single point in the parameter domain it is easy to calculate a projection that represents balancing and truncation. This projection can then be applied to the LPV system with the assumption that it is similar to the (parameter-independent) projection that would balance the state space representation over the whole parameter space. For some applications this approach can be successfully applied [e.g., Balas 2002a]. If the dynamics vary substantially over the parameter space, such a constant projection is usually insufficient. It is then tempting to calculate local approximations of the Gramians at various parameter values and to interpolate between grid points. This interpolation is guaranteed to be smooth for a sufficiently dense grid, since \( X_c(\rho) \)
and $X_o(\rho)$ which satisfy (3.22) are continuous functions of $\rho$ [Wood 1995, Cha. 7]. The problem of additional rate dependence nevertheless remains, as the interpolated Gramian approximations are necessarily parameter-varying and therefore result in a parameter-varying projection. In case either of these local approximations is used, any stability guarantees and error bounds are lost since the local solutions only satisfy the LMI constraints at single points in the parameter domain and not necessarily anywhere else.

For large-scale systems with several thousands of state variables, even the Lyapunov equations (3.34) become intractable. In this case, iterative low-rank approximations of the solutions to (3.34) can be used [e.g., Jaimoukha & Kasenally 1994, Li & White 2002, Benner et al. 2008]. Alternatively, empirical Gramian approximations [Willcox & Peraire 2002, Lall et al. 2002] can be constructed from impulse response data and the integral expressions

\[
X_c = \int_0^\infty e^{At}B^T e^{A^Tt} dt, \tag{3.35a}
\]

\[
X_o = \int_0^\infty e^{A^Tt} C^T C e^{At} dt. \tag{3.35b}
\]

The required matrix of impulse responses $e^{At}B$ in Equation (3.35a) can be obtained by simulating the full-order autonomous system $\dot{x} = Ax$ with initial conditions $x_0 = B_i$, $i = 1, \ldots, n_u$, where $B_i$ denotes the $i^{th}$ column of the matrix $B$. Each of the $n_u$ simulations results in a state trajectory $x_i(t)$, $i = 1, \ldots, n_u$ which is collected in a matrix $X(t) := [x_1(t) \cdots x_{n_u}(t)] = e^{At}B$. As the full-order system is stable, these trajectories decay for a sufficiently long simulation time $T_{\text{sim}}$ and can be calculated in a sampled form by standard numerical integration. Using a fixed step size $\Delta t$, the integral (3.35a) can be approximated as

\[
X_c \approx \frac{T_{\text{sim}}}{\Delta t} \sum_{k=0}^{T_{\text{sim}}/\Delta t} X(k\Delta t) X^T(k\Delta t) \Delta t. \tag{3.36a}
\]

An equivalent representation of Equation (3.36a) is

\[
X_c \approx X_{\text{sample}} X_{\text{sample}}^T \Delta t \text{ with } X_{\text{sample}} = [X(0) X(\Delta t) X(2\Delta t) \cdots X(T_{\text{sim}})] \tag{3.36b}
\]

The number of columns in $X_{\text{sample}}$ (and hence the rank of the estimate of $X_c$) is limited by the number of samples $n_{\text{sample}} := T_{\text{sim}}/\Delta t$. Equation (3.36b) further shows that $\sqrt{\Delta t} X_{\text{sample}}$ provides a factorization that resembles the Cholesky factorization. As the projection (3.33) requires only the Cholesky factor of $X_c$, this means that the multiplication of the two matrices never has to be carried out and that only the $n_u$-by-$n_{\text{sample}}$ matrix $X_{\text{sample}}$ has to be stored, not the $n_u$-by-$n_u$ matrix $X_c$. An approximation for the integral (3.35b) can be obtained using standard duality results from simulating the adjoint system $\dot{x} = A^T x$ with initial conditions corresponding to the $i^{th}$ row of the matrix $C$ [e.g. Antoulas 2005, Sec. 9.1.3]. The computational effort to obtain empirical Gramian approximation scales almost linearly with the state dimension and is hence also applicable for large-scale systems.
Modal State Space Decomposition

The rate dependence introduced by parameter-varying state transformations does not only complicate balanced model order reduction, but also prohibits a modal decomposition. For LTI systems, modal decomposition is a very powerful and widely used tool for model reduction. It allows dynamics below or above the frequency range of interest to be removed without affecting other parts of the system. Therefore, it is most useful when physical insight into the problem is available. Another important application is the decomposition of a system into a stable and an unstable part.

An LTI system \( \dot{x} = \mathbf{A}x + \mathbf{B}u \), \( y = \mathbf{C}x + \mathbf{D}u \) in modal form is described by

\[
\begin{align*}
\dot{x} &= \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \\ & \ddots \\ & & \ddots \\ & & & \ddots \\ & & & & \ddots \\ & & & & & 0 \\ & & & & & A_{n-m} \end{bmatrix} x + \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \\ \vdots \\ B_{n-m} \end{bmatrix} u, \\
y &= \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_m \\ \vdots \\ C_{n-m} \end{bmatrix} x + D u.
\end{align*}
\]

(3.37)

where

\[
A_i = \begin{bmatrix} \text{Re}(\lambda_i) & \text{Im}(\lambda_i) \\ -\text{Im}(\lambda_i) & \text{Re}(\lambda_i) \end{bmatrix} \quad \text{for } i = 1, \ldots, m \\
\begin{bmatrix} \lambda_{i+m} \end{bmatrix} \quad \text{for } i = m + 1, \ldots, n - m
\]

and \( \lambda_i = \overline{\lambda}_{i+m}, \ i = 1, \ldots, m \) denote the complex conjugate eigenvalues of the matrix \( \mathbf{A} \) and \( \lambda_i, \ i = 2m+1, \ldots, n \) denote real eigenvalues. Starting from an arbitrary state space representation with state variables \( \xi \), one possible transformation to transform the system into the modal form (3.37) is

\[
x = \begin{bmatrix} \text{Re}(v_1) & \text{Im}(v_1) & \cdots & \text{Re}(v_m) & \text{Im}(v_m) & v_{2m+1} & v_{2m+2} & \cdots & v_n \end{bmatrix}^{-1} \xi,
\]

where \( v_i \) are the eigenvectors corresponding to the eigenvalues \( \lambda_i \). The basis vectors in the modal coordinates are thus the eigenvectors of the original system.

For LPV systems, eigenvalues and eigenvectors of the parameter-dependent matrix \( A(\rho) \) are also parameter-dependent and hence a transformation into modal coordinates would necessarily be itself parameter-dependent [e.g., Wood 1995, Sec. 7.2]. As detailed in Section 3.1.1, such a parameter-dependent transformation inevitably leads to a rate-dependent term in the transformed model, which may produce large off-diagonal elements for non-zero rates. Consequently, the decoupled structure of Equation (3.37) can, in general, not be attained. The only remedy for this problem is again to apply a parameter-independent transformation. Such a transformation is usually calculated such that a modal
form is attained at a single grid point in the parameter domain. The modal coupling at
the other grid points is then accepted when state variables are removed from the LPV
system [e.g., Moreno et al. 2014, Sec. 4.3.2]. This can severely diminish the usefulness of
a modal form, where achieving decoupling is the primary goal.

3.1.4 Evaluation of Reduced-Order Models

The objective of model order reduction is to obtain reduced-order models that “resemble”
the full-order model. The natural metric to measure the quality of the approximation is
again the induced $L_2$-norm, introduced in Definition 2.4. The approximation error of a
reduced-order model can be expressed in terms of the dynamic system

$$
(P - P_{red}) \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} A(\rho) & 0 \\ 0 & A_{red}(\rho) \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} B(\rho) \\ B_{red}(\rho) \end{bmatrix} u \\
\begin{bmatrix} C(\rho) \\ -C_{red}(\rho) \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + (D(\rho) - D_{red}(\rho)) u.
$$

(3.38)

An upper bound for the induced $L_2$-norm of this system can then in principle be calculated
using Theorem 2.2. As this involves the solution of LMIs with the complete full-order model
appearing in the problem, the calculation is unfortunately very likely to be intractable.
Often, the only possibility to evaluate the quality of the approximation is thus either
to perform nonlinear simulations or to resort to local analysis. In the latter, the LPV
system is seen as the collection of parameterized frequency responses corresponding to
linearizations of a nonlinear system. Evaluation of the $H_\infty$-norm of the error system
(3.38) at multiple parameter values then at least provides an indication of the approximation
quality.

Aeroservoelastic systems often exhibit several very lightly damped modes such that the
$H_\infty$-norm is very sensitive to small frequency shifts. It is therefore suggested by Moreno
et al. [2014] to consider different metrics. In particular, the $\nu$-gap metric of Vinnicombe
[1993] appears to be a useful measure. It is defined as [Zhou & Doyle 1998, Cha. 17]

$$
\delta_\nu(P_1(s), P_2(s)) = \left\| (I + P_2(j \omega) P_2^*(j \omega))^{-1/2} (P_1(j \omega) - P_2(j \omega)) (I + P_1^*(j \omega) P_1(j \omega))^{-1/2} \right\|.
$$

(3.39)

The $\nu$-gap metric takes values between zero and one, where zero is attained for two
identical systems. A system $P_1$ that is close to another system $P_2$ in the $\nu$-gap metric is
stabilized or destabilized by the same feedback controller that stabilizes or destabilizes
$P_2$. The $\nu$-gap metric thus to a certain degree measures similarity under feedback. It thus
captures the likelihood that a feedback controller designed on the reduced-order model will
also perform adequately on the full-order model. When calculated frequency by frequency,
the measure is referred to simply as $\nu$-gap in this thesis. The term $\nu$-gap metric refers to
the maximum over all frequencies as in Equation (3.39).
3 Model Order Reduction for LPV Systems

3.2 Approximation by Modal Interpolation

This section introduces a novel model order reduction method for LPV systems. The proposed approach is to reduce an LPV model locally, i.e., at each grid point individually. Doing so alleviates the restrictions detailed in Section 3.1.2 by making use of well developed and numerically efficient techniques for LTI systems. The result is a set of parameterized, individual models with different, i.e., inconsistent state space bases. To make these models amenable to interpolation and hence to obtain a reduced-order LPV model, a consistent state space basis is required. The main challenge therefore is to transform all individual models into such a consistent basis.

There is a variety of approaches concerned with this problem available in the literature. For example, De Caigny et al. [2011] propose to construct a state space representation with minimum variation in the individual matrix entries. Ferreres [2011] and Adegas et al. [2013] use a companion form, i.e., they interpolate coefficients of polynomials. Roos [2009] and Poussot-Vassal & Roos [2012] use modal forms and the angle between eigenvectors as a measure to pair modes across different local models. Very recently, the use of modal forms was also suggested by Gőzse et al. [2016] and Luspay et al. [2018] with a hyperbolic metric to associate modes of different local models and by Al-Jiboory et al. [2017] with an $H_2$-norm criterion. Still, no consensus on the general applicability of these approaches was reached.

This section proposes to use a modal representation as a state space basis and to associate modes across different models using subspace angles as well as the frequency and damping ratio of the modes. An application example for the method proposed in this section is given Section 5.1.1, where the number of state variables in the high-fidelity model of an unmanned aeroservoelastic aircraft is reduced from 148 to 15 and further a comparison with the state-of-the-art LPV balanced truncation techniques is provided.

3.2.1 Local Model Reduction

In the first step of the proposed method, the LPV system is evaluated for fixed parameter values at each grid point to form a set of LTI systems. These LTI models are then individually reduced using any available LTI model-order reduction technique such that the number of state variables in all reduced-order models is the same and, in particular, such that the number of complex and real poles is the same across all models. The resulting set of individual models, by construction, approximates the full-order LPV model at each grid point. That is, the input-output behavior of all local models is consistent. The internal representation of the dynamics, i.e., the basis of the state space, however, may vary at different grid points as a consequence of the local reduction. The local reductions can be thought of as a global parameter-varying projection, implicitly defined at the grid points in the parameter domain. As detailed in Section 3.1.1, it therefore introduces an (unknown) rate-dependent term in the reduced-order LPV model which is defined through the local approximations. Such a reduction is intuitively expected to yield better local results than a parameter-independent projection calculated at a single grid point or even from the LMI approach. This advantage, however, comes at the potentially very high price...
of an unknown approximation error caused by neglecting rate dependence in the resulting reduced-order model. It is clear that this approximation error can be arbitrary large and that hence an LPV model which is simply defined through interpolation of individual reduced-order LTI models is usually a very bad approximation of the original LPV model.

The important contribution of the proposed method is a way to represent the local models in a state space basis such that interpolation can be performed without introducing large errors. This can be thought of as minimizing the parameter variation in the basis. For single-input single-output (SISO) systems, one such consistent representation would be in terms of polynomial coefficients, resulting in a companion state space form [Kailath 1980, Sec. 2.1.2, cf. Tóth et al. 2012]. Such a representation is however numerically ill-conditioned and therefore not recommended, in particular for high-order systems. It is therefore proposed to use a modal form to recover an approximately consistent state space basis. As all individual models represent the same physical system, they should have similar modes. Further, the dynamic properties of these modes such as natural frequencies and damping ratios should be continuous such that they can be related across individual models throughout the parameter domain. The result is then a reduced-order LPV model with a parameter-dependent state space basis which nevertheless represents consistent physical behavior.

### 3.2.2 Modal Matching

In order to construct such a modal state space basis, each local reduced-order model is first transformed to the modal form (3.37). Although the state space matrix \( A \) in Equation (3.37) is in Jordan real form, the dynamic system is not in a canonical representation. In particular, modes can be permuted and further any block diagonal state transform that commutes with the Jordan blocks (i.e., any block-diagonal matrix with blocks of the form \([ \begin{smallmatrix} a & -b \\ b & a \end{smallmatrix} \]) preserves the matrix \( A \) while the state space basis is altered. The modal form given in Equation (3.37) is thus not unique and is therefore not suitable for interpolation. This problem is addressed by a transformation to what will be referred to as a mode-wise canonical form. This transformation takes each of the oscillatory modal subsystems \((A_i, B_i, C_i), i = 1, \ldots, m\), and represents it in a canonical second-order representation

\[
A_i = \begin{bmatrix} 0 & 1 \\ -\omega_i^2 & -2\omega_i \zeta_i \end{bmatrix}, \quad B_i = \begin{bmatrix} 0 & \cdots & \cdots & * \\ 1 & \cdots & \cdots & * \end{bmatrix}, \quad C_i = \begin{bmatrix} * & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & * \end{bmatrix}, \quad (3.40)
\]

where \( \omega_i \) is the natural frequency of the mode, \( \zeta_i \) is the corresponding damping ratio and \( * \) denotes possible non-zero entries. The two states associated with a single mode thus correspond to generalized displacement and generalized velocity and are normalized in magnitude and sign by the first column of the input matrix. For the non-oscillatory modes \( i = m+1, \ldots, n-m \) corresponding to real poles, the representation simply is \( A_i = [\lambda_i] \).

\*Completely recovering consistency appears to be, in general, impossible since the individual reduction operations are surjections, i.e. non-invertible.
The state space representation of a model in this mode-wise canonical form is

\[
\begin{bmatrix}
\dot{\xi}_1 \\
\dot{\xi}_2 \\
\vdots \\
\dot{\xi}_{n-m}
\end{bmatrix}
= \begin{bmatrix}
A_1 & 0 & 0 & \cdots & 0 \\
0 & A_2 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & A_{n-m}
\end{bmatrix}
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_{n-m}
\end{bmatrix}
+ \begin{bmatrix}
B_{i1} \\
B_{i2} \\
\vdots \\
B_{in-m}
\end{bmatrix}
\begin{bmatrix}
u
\end{bmatrix}
\]

Starting from the state space representation (3.37), it can be obtained using a permuted version of the multivariable controller canonical form construction scheme introduced by Kailath [1980, Sec. 6.4.6]. An explicit representation for this transformation is

\[
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_{n-m}
\end{bmatrix} = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix}
\begin{bmatrix}
A_{i1} & B_{i1} & \cdots & 0 \\
0 & A_{i2} & B_{i2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & A_{i,n-m}
\end{bmatrix}
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_{n-m}
\end{bmatrix}
+ D \begin{bmatrix}
u
\end{bmatrix}
\]

for complex modes and

\[
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_{n-m}
\end{bmatrix} = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix}
\begin{bmatrix}
A_{i1} & B_{i1} & \cdots & 0 \\
0 & A_{i2} & B_{i2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & A_{i,n-m}
\end{bmatrix}
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_{n-m}
\end{bmatrix}
+ D \begin{bmatrix}
u
\end{bmatrix}
\]

for real modes. Still, modes in different models may be ordered differently. This problem is addressed by matching modes at neighboring grid points and permuting the modes such that they appear in the same order in all models. This matching is achieved by comparing the dynamic properties of all modes at one grid point to the dynamic properties of all modes at a neighboring grid point. The key assumption of this modal matching approach is that modes which represent the same dynamic properties of the system at different grid points have similar characteristics on a sufficiently dense grid. The following four criteria are used as a measure for these characteristics:

1. the damped natural frequency \( \tilde{\omega}_i = |\lambda_i| = \sqrt{\omega_i^2 - \zeta_i^2} \)
2. the damping ratio \( \zeta_i = \arctan \frac{\text{Re}(\lambda_i)}{\text{Im}(\lambda_i)} \)
3. the input direction \( \text{span}(B_i^T) \)
4. the output direction \( \text{span}(C_i) \)
3.2 Approximation by Modal Interpolation

These criteria are combined into a matching function $M_k$ which measures the degree of similarity between a mode $(A_{k,i}, B_{k,i}, C_{k,i})$ at the grid point $k$ and a mode $(A_{k+1,j}, B_{k+1,j}, C_{k+1,j})$ at the neighboring grid point $k + 1$. The matching function for each grid point takes the form of a matrix $M_k \in \mathbb{R}^{n-m \times n-m}$ and is defined as

$$
[M_k]_{i,j} = \min \left( \frac{\bar{\omega}_{k,i}}{\bar{\omega}_{k+1,j}}, \frac{\bar{\omega}_{k+1,j}}{\bar{\omega}_{k,i}} \right)^{\alpha} + \min \left( \frac{\zeta_{k,i}}{\zeta_{k+1,j}}, \frac{\zeta_{k+1,j}}{\zeta_{k,i}} \right)^{\alpha} + \left( \frac{\| \langle B_{k,i}, B_{k+1,j} \rangle_F \|}{\| B_{k,i} \|_F \| B_{k+1,j} \|_F} \right)^{\alpha} + \left( \frac{\| \langle C_{k,i}, C_{k+1,j} \rangle_F \|}{\| C_{k,i} \|_F \| C_{k+1,j} \|_F} \right)^{\alpha},
$$

(3.42)

where $\langle X, Y \rangle_F = \text{trace} (X^T Y) = \text{trace} (Y^T X)$ is the Frobenius scalar product, $\| X \|_F = \sqrt{\langle X, X \rangle_F}$ is the Frobenius norm, and $\alpha \geq 1$ is a free parameter to adjust the matching function. The first two terms compare natural frequency and damping ratio, the last two compare the alignment of input and output directions, i.e., how the modes are affected by the input and affected the output. Each of the four terms equals one if the two compared modes are identical and is less than one if they differ. The parameter $\alpha$ determines how sensitive the matching function is to variations of the characteristics; a larger value can help to separate close-by modes. For the application considered in Section 5.1.1, $\alpha = 2$ yielded consistently satisfying results.

Each row of the matching function corresponds to a mode at grid point $k$ while each column represents a mode at grid point $k + 1$. Large values $[M_k]_{i,j}$ are then an indicator for the $i^{th}$ mode at grid point $k$ to correspond to the $j^{th}$ mode at grid point $k + 1$. Consequently, a permutation index can be calculated from determining the column of the matching function with the largest matching value for each row. The modes in the model at grid point $k + 1$ are then permuted according to this permutation index such that matching modes appear in the same order as in the model at grid point $k$. This procedure is repeated for all grid points $k = 1, \ldots, n_{\text{grid}} - 1$.

3.2.3 Stability Considerations and Limitations

Even when all individual models are described in the mode-wise canonical form (3.41) and the modes are correctly ordered, models at different grid points may still be described in (slightly) different state space bases for essentially two reasons: First, the mode shapes change, hence the basis for the mode-wise canonical state space description in fact also varies across the parameter space. Second, variations in the plant’s $B$ matrix may lead to a different normalization of the state variables in the mode-wise canonical form. These consequences appear to be inevitable but pose no real threat: From physical considerations, the assumption that the mode-shapes vary continuously with the parameters is reasonable and in fact motivated the use of a modal coordinate system as the common basis. The same reasoning also applies to the normalization. The effect of the input on a mode should depend continuously on the parameter and hence the variation in the normalization should be small for a sufficiently dense grid.

The major restriction of the matching algorithm is that all local models are required to have the same order and the same number of complex and real poles (i.e., modes). It is
therefore not possible to apply the method to models with “mode-veering” phenomena, i.e., where a complex mode splits into two real modes or vice versa. Further, reduction methods such as balancing and truncation or residualization can sometimes place poles at odd locations in the local reduction step, such that pairing becomes difficult or even impossible. The method therefore is expected to work best when the local reduction is also performed based on modal decompositions that exactly preserve the modal characteristics.

As a consequence of the local approach, no error bounds can be established for the resulting reduced-order LPV model and stability, in general, cannot be guaranteed. There exists, however, a weak stability result: If all individually reduced-order models are stable, then also the LPV model is stable, provided that the parameter variation is sufficiently slow as was shown by Desoer [1969].

### 3.2.4 Numerical Example

The nonlinear mass-spring-damper system of Example 2.1 (p. 14) is used to illustrate the proposed approach and provide a comparison to the state-of-the-art LMI-based method. In order to provide a meaningful comparison, the full-order LPV model obtained through point-wise Jacobian linearization is used as a reference, not the original nonlinear dynamic system. Further, only the state space representation \((A, B, C, D)\) is used, without explicitly bookkeeping the varying operating point \((\bar{x}, \bar{y})\). The function `lpvbalreal` of the LPVTools toolbox [Balas et al. 2015] is used to solve the optimization problem (3.31) and to calculate a reduced-order model through balancing and truncation. Doing so takes more than 2.5 hours on a 64-bit desktop computer with 3.4 GHz 8-core CPU and 8 GB RAM. For the modal interpolation procedure, LTI balancing and truncation at each grid point is used in the local reduction step, employing the Matlab routine `balreal`. For an order of three, this results in one real and one complex mode in each model, such that the modal matching procedure can be applied.\(^9\) The proposed modal interpolation algorithm only takes 0.3 seconds and results in a good approximation of the LPV model as shown in Figure 3.1.

Specifically, Figure 3.1a shows that the local error at each grid point is lower than for the LMI method. Calculating the maximum absolute error at the grid points confirms this observation and results in 0.0024 for the proposed procedure compared to 0.0083 for the LMI method. The time-domain simulation results also agree quite well as shown in Figure 3.1b, although the modal interpolation approximant is not able to match the accuracy of the LMI method. Next, an upper bound \(\bar{\gamma}\) on the error in the induced \(L_2\)-norm is calculated using Equation (3.38) and the LPVTools function `lpvnorm`. This calculation takes about 55 minutes and is performed with a parameter-independent storage function, such that the result is valid for arbitrary fast parameter variation. The obtained bound for the approximation through balancing and truncation is 0.036. The modal interpolation procedure results in a far worse error bound of only 0.96. This shows that the local evaluation alone is not a sufficient measure for the quality of the approximation of LPV

\(^9\)It must be noted, however, that a model of, e.g., order 4 would be impossible to obtain for the present example, as this would lead to a different number of complex and real modes at each grid point.
models. The large (potential) approximation error is expected to be caused by neglecting the rate term associated with the parameter-dependent basis as described previously. To further investigate this effect, the error bound calculation is repeated with rate bounds of \( \pm 1N^2/s \) and an affine storage function \( V(x, \rho) = x^T (X_0 + \rho X_1) x \). This calculation takes over 11 hours and results in bounds of 0.015 for balancing and truncation and 0.26 for the modal interpolation. The error bound for the modal interpolation model is hence still much larger than for the reduced-order model from balancing and truncation, but the gap between the two is significantly reduced. Table 3.1 summarizes these results.

![Figure 3.1: Comparison of original model (100 state variables) and reduced-order model from LMI method (3 state variables) and modal interpolation (3 state variables).](image)

**Table 3.1: Comparison of reduced-order models.**

<table>
<thead>
<tr>
<th></th>
<th>LPV Balancing</th>
<th>Modal Interpolation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computation Time†</td>
<td>9170 s</td>
<td>0.3 s</td>
</tr>
<tr>
<td>Local ( H_{\infty} )-norm error</td>
<td>0.0083</td>
<td>0.0024</td>
</tr>
<tr>
<td>( L_2 )-norm error bound</td>
<td>unbounded rates</td>
<td>bounded rate</td>
</tr>
<tr>
<td></td>
<td>0.0364</td>
<td>0.0149</td>
</tr>
</tbody>
</table>

† on a 64 bit desktop PC with 3.4 GHz 8-core CPU and 8 GB RAM
3.3 Parameter-Varying Oblique Projection

In this section, a Petrov-Galerkin approximation with a parameter-independent matrix to specify the basis space, but a parameter-dependent matrix to define the test space is proposed. Doing so achieves additional freedom over parameter-independent projections and at the same time avoids rate dependence in the reduced-order model. That is, in contrast to the modal interpolation method of Section 3.2, no rate-dependent error is introduced and the reduced-order LPV model is in a completely consistent state space representation. This extension of the projection framework lends itself very naturally to balanced truncation and related approaches that employ Gramian-based information to quantify the importance of subspaces. Suitable spaces for approximate balancing of the LPV system are calculated from local approximations that are numerically well tractable. The novel method can therefore be used in case the LMI solution fails, either due to an overly restricted search space or due to computational complexity. Application examples of this method are presented in Sections 5.1.2 where the high-fidelity model of the longitudinal dynamics of a small unmanned aeroservoelastic aircraft is reduced from 48 to 12 state variables and in Section 5.2 where a reduction from 20502 to 6 state variables is achieved for a model of the aerodynamics of a wind turbine.

3.3.1 Parameter-Varying Petrov-Galerkin Conditions

It was shown in Section 3.1.1 that a parameter-varying projection leads to rate dependence in the reduced-order model. Specifically, this additional dependence is caused by $V(\rho)$ in the approximation, i.e.,

\[
x_{\text{approx}} = \frac{d}{dt} (V(\rho) z) = V(\rho) \dot{z} + \sum_{i=1}^{n_x} \frac{\partial V(\rho)}{\partial \rho_i} \dot{\rho}_i z.
\]

(3.43)

This is intuitively understood as a result of representing the dynamics of the reduced-order model in a parameter-dependent, and hence time-varying, basis. This rate dependence has, however, nothing to do with the choice of a test space for the approximation. Hence, rate dependence of the reduced-order model can be avoided by restricting $V$ to be constant, regardless of whether $W(\rho)$ is parameter-varying. Such a projection is formulated in Proposition 3.2.

**Proposition 3.2.** Let $V \in \mathbb{R}^{n_x \times n_z}$ be a given constant matrix with rank($V$) = $n_z < n_x$ and let $S: \mathbb{R}^{n_c} \mapsto \mathbb{R}^{n_x \times n_x}$ be a given non-singular matrix function. Then,

\[
\Pi(\rho) = V \left( V^T S(\rho) V \right)^{-1} V^T S(\rho) \frac{W(\rho)}{W(\rho)}
\]

(3.44)

is a parameter-varying oblique projection. The dynamic LPV system obtained by Petrov-Galerkin approximation with constant basis space $V := \text{span}(V)$ and parameter-varying test space $W(\rho) := \text{span}(W(\rho))$ does not depend on the rate of parameter variation.

\[\blacksquare\]
Proof. It is readily verified that $W^T(\rho) V = I_{n_z} \forall \rho$ and that hence $\Pi(\rho) = \Pi^2(\rho)$ is an oblique projection. Further, since $V$ is constant, the reduced-order state space equations (3.21) simplify to

$$\dot{z} = W^T(\rho) A(\rho) V z + W^T(\rho) B(\rho) u$$
$$y = C(\rho) V z + D(\rho) u.$$  \hspace{1cm} (3.45)

The range space of the projection (3.44) in Proposition 3.2 is constant, but the nullspace of the projection, i.e., its direction, varies across the parameter space. The results are the seemingly paradox relations $V z \approx x$ and $z \approx W^T(\rho) x$, that become clearer once the approximation

$$\dot{x}_{\text{approx}} = V W^T(\rho) A(\rho) x_{\text{approx}} + V W^T(\rho) B(\rho) u$$  \hspace{1cm} (3.46)

is considered. It is not the approximate state vector, but its derivative that is calculated according to a parameter-varying condition described by $W(\rho)$. Hence, the state update is performed depending on the current parameter value, while the approximate state vector itself always remains in $V$. The latter is an immediate consequence of idempotency of the projection, i.e., $x_{\text{approx}} = V W^T(\rho) x_{\text{approx}} \forall \rho$.

The benefit of the projection in Proposition 3.2 over an arbitrary parameter-varying projection is that it avoids rate dependence in the reduced-order model. It still has additional freedom in constructing a reduced-order model compared to a completely parameter-independent projection. This additional freedom is the choice of a parameter-varying nullspace, i.e., the Petrov-Galerkin condition (3.16) can be changed over the parameter-space as illustrated by the following example.

Example 3.1 (Dominant Subspace Approximation). Assume that a state space realization $\dot{x} = A(\rho) x + B(\rho) u$, $y = C(\rho) x + D(\rho) u$ of an LPV model on a domain $\mathcal{T}$ is given. Let $X_c$ be a parameter-independent generalized controllability Gramian satisfying the LMI (3.22a) and let $X_o(\rho)$ be a generalized observability Gramian, i.e., a solution to the LMI (3.22b). Let $U_c \Sigma_c U_c^T = X_c$ and $U_o(\rho) \Sigma_o(\rho) U_o^T(\rho)$ denote the SVD of the Gramians where the diagonal matrices $\Sigma_c$ contain the eigenvalues in decreasing magnitude and $U_c$ contain the corresponding normalized eigenvectors.

A possible choice for the basis space and test space of a Petrov-Galerkin approximation in accordance with Proposition 3.2 is

$$\mathcal{V} = \text{span} \left( U_c \begin{bmatrix} I_{n_z} \\ 0_{n_z \times n_{nz}} \end{bmatrix} \right), \quad W_\rho = \text{span} \left( U_o(\rho) \begin{bmatrix} I_{n_z} \\ 0_{n_z \times n_{nz}} \end{bmatrix} \right).$$  \hspace{1cm} (3.47)

That is, the basis space is the dominant controllable subspace and the test space is the
dominant observable subspace [cf., e.g., Shi & Shi 2005].\(^{10}\) The proposed parameter-varying oblique projection is advantageous in this case, as it can take into account varying observability of certain directions in the state space over the parameter domain. \(\triangle\)

### 3.3.2 Basis Space and Test Space Construction

The possibility to construct a reduced-order model through Petrov-Galerkin approximation with parameter-independent basis space but parameter-varying test space is very general and completely independent of any specific choice for these spaces. The choice of the dominant controllable subspace and the dominant observable subspace as in Example 3.1 might prove useful for some applications, but, in general, individually considering observability and controllability is not a reliable way to obtain accurate reduced-order models. In particular, observability and controllability depend on the specific state space realization and can be manipulated according to Equation (3.23). Only the product of the Gramians is invariant with respect to a change of coordinates, which is precisely the motivation for balancing. Unfortunately, the product of the Gramians even in Example 3.1 is parameter-dependent and hence cannot be directly used to define basis space and test space for a projection in accordance with Proposition 3.2. Only completely parameter-independent generalized Gramians would avoid this problem and result in the parameter-independent standard projection. Thus, the possible advantages of a varying test space would not be exploited. This section addresses this problem and establishes a method to calculate a parameter-independent basis space and a parameter-varying test space that approximate balancing and truncation.

First, the projection (3.33) for balancing and truncation is reformulated in the following proposition.

**Proposition 3.3.** Let the controllability and observability Gramians \(X_o = L_o L_o^T\) and \(X_c = L_c L_c^T\), and the SVD of the product \(L_c^T L_o = [U_1, u_2] \begin{bmatrix} \Sigma_1 & \Sigma_2 \end{bmatrix} [N_1, N_2]^T\) be given. Let \(Q\) denote a basis for \(\text{span}(L_c U_1)\). A projection that achieves balancing and truncation is

\[
\Pi_{\text{bal}} = Q \frac{(Q^T X_o Q)^{-1} Q^T X_o}{\sqrt{\frac{1}{W^T}}}. \tag{3.48}
\]

**Proof.** Let \(QR\) denote the thin QR-factorization of \(L_c U_1\). Replacing \(Q\) by \(L_c U_1 R^{-1}\) and \(X_o\) by \(L_o L_o^T\) in (3.48) yields

\[
\Pi_{\text{bal}} = L_c U_1 (U_1^T L_o^T L_o L_c U_1)^{-1} U_1^T L_c L_o L_c^T.
\]

\(^{10}\)A similar choice is also commonly encountered in model order reduction for LTI systems based on moment matching [e.g. Antoulas 2005, Sec. 11.3.1], where the basis space and test space are \(V = [B \ A B \ A^2 B \cdots A^{n-1} B]\) and \(W = [C \ C A \ C A^2 \cdots C A^{n-1}]\) which can be recognized to form approximations of the dominant controllable and observable subspaces.
It follows from the given SVD that \( U_T^T L_c^T L_o = \Sigma_1 N_1^T \). Substitution of this expression results in
\[
\Pi_{\text{bal}} = L_c U_1 (\Sigma_1 N_1^T N_1 \Sigma_1)^{-1} \Sigma_1 N_1^T L_o^T.
\]
Equivalence to Equation (3.33) is finally established using \((\Sigma_1 N_1^T N_1 \Sigma_1)^{-1} \Sigma_1 = \Sigma_1^{-1} \). \(\blacksquare\)

It should be noted that Proposition 3.3 results in balancing-free balancing in the sense of Varga [1991]. That is, the resulting reduced-order system is not internally balanced but has the exact same input-output behavior as the balanced reduced-order model.

Equating the form of an admissible parameter-varying oblique projection in Proposition 3.2 with the projection for balancing and truncation in Proposition 3.3, it is clear that the matrix \( X_o \) is allowed to be parameter-dependent. Hence, a parameter-dependent generalized observability Gramian could directly be used. Nevertheless, a constant basis \( Q \in \mathbb{R}^{n_x \times n_z} \) for the basis space is required to avoid rate dependence in the reduced-order model. In order to reproduce balancing and truncation, \( Q \) should be a basis for \( \text{span}(L_c(\rho) U_1(\rho)) \). In general, this can only be achieved by a parameter-varying matrix \( Q(\rho) \). The remaining critical question is thus in what sense a constant subspace \( \text{span}(\bar{Q}) \) should approximate the parameter-varying subspace \( \text{span}(Q(\rho)) \).

This question also arises in parametric model order reduction and a common approach, first introduced by Panzer et al. [2010], is referred to as concatenation of the basis [Benner et al. 2013, Sec. 4.1]. For a given grid of parameter values \( \{\rho_k\}_{k=1}^{n_{\text{grid}}} \), function evaluations at these grid points are stacked into a matrix and an SVD is performed, i.e.,
\[
\begin{bmatrix}
Q(\rho_1) & \cdots & Q(\rho_{n_{\text{grid}}})
\end{bmatrix} = \begin{bmatrix} \bar{Q} & \bar{Q}_\perp \end{bmatrix} \begin{bmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_2 & \Sigma_2 \end{bmatrix} \begin{bmatrix} \bar{U} & \bar{U}_\perp \end{bmatrix}^T.
\]
(3.49)

If \( \Sigma_2 = 0 \), the matrix \( \bar{Q} \) forms a constant basis for the union of the individual subspaces, i.e., \( \text{span}(\bar{Q}) \supset \text{span}(Q(\rho_k)) \), \( k = 1, \ldots, n_{\text{grid}} \). Taking \( \bar{Q} \) as the basis space hence only removes redundant directions and is expected to yield an accurate approximation of the overall system, if the individual \( \{Q(\rho_k)\}_{k=1}^{n_{\text{grid}}} \) provide accurate local approximations. One problem even in this case is that, in general, \( Q \in \mathbb{R}^{n_x \times m} \) where \( m > n_z \). Thus, the number of state variables in the reduced-order model is greater than initially desired and might even be close to (or equal to) the full order if \( Q \) varies substantially. In this case, \( \Sigma_1 \) can be selected to only contain the largest \( n_z \) singular values such that \( \bar{Q} \in \mathbb{R}^{n_x \times n_z} \). This obviously leads to an approximation and results in an additional error in the reduced-order model. It is shown in the following proposition that this approximation is nevertheless optimal in a geometrical sense.

**Proposition 3.4.** Let \( Q : \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_z} \) with \( Q(\rho) \) \( Q(p) = I_{n_z} \quad \forall p \in \mathcal{P} \) be a given parameter-dependent matrix function defined on the compact set \( \mathcal{P} \). Let further
\[
\begin{bmatrix} Q(\rho_1) & \cdots & Q(\rho_{n_{\text{grid}}}) \end{bmatrix}^T \quad \text{denote the collection of function evaluations of } Q \text{ on the grid of parameter values } \{\rho_k\}_{k=1}^{n_{\text{grid}}}.
\]
Let
\[
\begin{bmatrix} Q(\rho_1) & \cdots & Q(\rho_{n_{\text{grid}}}) \end{bmatrix} = \begin{bmatrix} \bar{Q} & \bar{Q}_\perp \end{bmatrix} \begin{bmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_2 & \Sigma_2 \end{bmatrix} \begin{bmatrix} \bar{U} & \bar{U}_\perp \end{bmatrix}^T.
\]
(3.50)
be the SVD such that \( \bar{\Sigma}_1 \) contains the largest \( n_z \) singular values and \( \bar{\Sigma}_2 \) the remaining \( n_x - n_z \) singular values. Then, the constant matrix \( \bar{Q} \in \mathbb{R}^{n_x \times n_z} \) is a basis for an optimal approximation \( \text{span}(\bar{Q}) \) to \( \text{span}(Q(\rho)) \) in the sense that its orthogonal complement \( \bar{Q}_\perp \in \mathbb{R}^{n_x \times (n_x - n_z)} \) minimizes

\[
\min_{\bar{Q}_\perp} \sum_{k=1}^{n_{\text{grid}}} \| Q^T(\rho_k) \bar{Q}_\perp \|^2_F \quad \text{s. t.} \quad \bar{Q}_\perp^T \bar{Q}_\perp = I_{n_x - n_z}.
\]

(3.51)

**Proof.** Since the equality \( \| X \|_F^2 + \| Y \|_F^2 = \| [X \ Y] \|_F^2 \) holds for all matrices \( X \) and \( Y \) of compatible dimensions, (3.51) can be rewritten as

\[
\min_{\bar{Q}_\perp} \left\| \begin{bmatrix} Q^T(\rho_1) \\ \vdots \\ Q^T(\rho_{n_{\text{grid}}}) \end{bmatrix} \bar{Q}_\perp \right\|^2_F \quad \text{s. t.} \quad \bar{Q}_\perp^T \bar{Q}_\perp = I_{n_x - n_z}.
\]

The minimizer \( \bar{Q}_\perp \) is found from the SVD

\[
\begin{bmatrix} Q^T(\rho_1) \\ \vdots \\ Q^T(\rho_{n_{\text{grid}}}) \end{bmatrix} = [U \ \bar{U}_\perp] \begin{bmatrix} \bar{\Sigma}_1 \\ \bar{\Sigma}_2 \end{bmatrix} \bar{Q} \bar{Q}_\perp^T
\]

(3.52)

which is simply the transpose of Equation (3.50). The optimal cost is \( \| \bar{U}_\perp \bar{\Sigma}_2 \|^2_F \).

Proposition 3.4 states that a constant subspace \( \text{span}(\bar{Q}) \) is an optimal approximation (in the Frobenius norm) for the parameter-varying subspace \( \text{span}(Q(\rho)) \) if its orthogonal complement \( \text{span}(\bar{Q}_\perp) = \text{span}(\bar{Q})^\perp \) approximates the parameter-varying nullspace \( \ker(Q^T(\rho)) \). For a single grid point \( \rho_k \), the matrix \( Q^T(\rho_k) \bar{Q}_\perp \) is rank deficient and the optimal “approximation” that attains the exact minimum cost 0 is simply a basis for the nullspace \( \ker(Q^T(\rho_k)) \). Obviously, a constant basis cannot, in general, achieve a zero norm for multiple grid points at once if \( \text{span}(Q(\rho)) \) is not constant. The cost function in (3.51) thus measures the average “nullspace violation” over all grid points.

Combining Propositions 3.2, 3.3, and 3.4, it becomes possible to construct an approximation to balancing and truncation which makes use of parameter-dependent Gramians but nevertheless avoids rate dependence in the reduced-order model. This is summarized as Algorithm 1. For convenience, Matlab notation is used, e.g., \( Q(1:n_x,1:n_z) \) denotes the first \( n_x \times n_z \) elements of the matrix \( Q \). A * denotes quantities of no interest. The algorithm requires a grid of parameter values \( \{ \rho_k \}_{k=1}^{n_{\text{grid}}} \) and a gridded LPV system \( P_\rho \) with state space representation \( \{ A_{\rho_k}, B_{\rho_k}, C_{\rho_k}, D_{\rho_k} \}_{k=1}^{n_{\text{grid}}} \). The first step is to calculate generalized Gramians that satisfy the LMIs (3.22). Analog to the optimization problem (3.31), the singular values of the product of the Gramians can be pushed down by solving an optimization problem with the objective \( \text{trace}(Y) \) and the additional constraint
Algorithm 1 Approximate Balancing and Truncation by Parameter-Varying Oblique Projection

Input: LPV system \( \{P_k\}_{k=1}^{n_{\text{grid}}} \) with state space data \( \{A_{p_k}, B_{p_k}, C_{p_k}, D_{p_k}\}_{k=1}^{n_{\text{grid}}} \), desired order \( n_c \)

Output: LPV system \( \{P_{\text{red},k}\}_{k=1}^{n_{\text{grid}}} \) with state space data \( \{A_{\text{red},p_k}, B_{\text{red},p_k}, C_{\text{red},p_k}, D_{\text{red},p_k}\}_{k=1}^{n_{\text{grid}}} \)

\[
(X_{c,p_k})_{k=1}^{n_{\text{grid}}}, (X_{o,p_k})_{k=1}^{n_{\text{grid}}} \leftarrow \text{generalizedGramians}(P_k)_{k=1}^{n_{\text{grid}}} \]

\%

for \( k = 1 \) to \( n_{\text{grid}} \) do

\( L_{c,p_k} \leftarrow \text{chol}(X_{c,p_k}) \)

\( L_{o,p_k} \leftarrow \text{chol}(X_{o,p_k}) \)

\( (U, \ast, \ast) \leftarrow \text{svd}(L_{c,p_k} L_{o,p_k}) \)

\( (U_1, \ast, \ast) \leftarrow \text{svd}(L_{c,p_k} U(1:n_x,1:n_z)) \)

\( Q(1:n_x,1:n_z (k-1):n_z k) \leftarrow U_1(1:n_x,1:n_z) \)

\%

end for

\( (Q, \ast, \ast) \leftarrow \text{svd}(Q) \)

\( V \leftarrow Q(1:n_x,1:n_z) \)

\%

for \( k = 1 \) to \( n_{\text{grid}} \) do

\( (q, r) \leftarrow \text{qr}(L_{c,p_k} V) \)

\( W_{p_k} \leftarrow L_{o,p_k} q (r^T)^{-1} \)

\%

end for

\( W = X_o(p) V (V^T X_o(p) V)^{-1} \)

\%

The calculation of the Gramians, however, remains a bottleneck in terms of computational complexity as the LMIs (3.22) become intractable for many systems even of moderate state dimension as detailed in Section 3.1.2. As discussed in Section 3.1.3, local approximations of the Gramians are much easier to calculate. In particular, reliable and efficient algorithms such as Hammarling’s [1982] algorithm (e.g., implemented in

\[\text{Algorithm 1 Approximate Balancing and Truncation by Parameter-Varying Oblique Projection}\]

\[\text{Input: LPV system } \{P_k\}_{k=1}^{n_{\text{grid}}} \text{ with state space data } \{A_{p_k}, B_{p_k}, C_{p_k}, D_{p_k}\}_{k=1}^{n_{\text{grid}}}, \text{desired order } n_c\]

\[\text{Output: LPV system } \{P_{\text{red},k}\}_{k=1}^{n_{\text{grid}}} \text{ with state space data } \{A_{\text{red},p_k}, B_{\text{red},p_k}, C_{\text{red},p_k}, D_{\text{red},p_k}\}_{k=1}^{n_{\text{grid}}}\]

\[\{(X_{c,p_k})_{k=1}^{n_{\text{grid}}}, (X_{o,p_k})_{k=1}^{n_{\text{grid}}} \leftarrow \text{generalizedGramians}(P_k)_{k=1}^{n_{\text{grid}}})\]

\%

for \( k = 1 \) to \( n_{\text{grid}} \) do

\[L_{c,p_k} \leftarrow \text{chol}(X_{c,p_k})\]

\[L_{o,p_k} \leftarrow \text{chol}(X_{o,p_k})\]

\[U, \ast, \ast \leftarrow \text{svd}(L_{c,p_k} L_{o,p_k})\]

\[U_1, \ast, \ast \leftarrow \text{svd}(L_{c,p_k} U(1:n_x,1:n_z))\]

\[Q(1:n_x,1:n_z (k-1):n_z k) \leftarrow U_1(1:n_x,1:n_z)\]

\%

end for

\[(Q, \ast, \ast) \leftarrow \text{svd}(Q)\]

\[V \leftarrow Q(1:n_x,1:n_z)\]

\%

for \( k = 1 \) to \( n_{\text{grid}} \) do

\[q, r \leftarrow \text{qr}(L_{c,p_k} V)\]

\[W_{p_k} \leftarrow L_{o,p_k} q (r^T)^{-1}\]

\%

end for

\[W = X_o(p) V (V^T X_o(p) V)^{-1}\]

\%

\[\text{There also exist approaches with more sophisticated interpolation schemes for the projection matrices.}\]

\[\text{In particular, Amsallem & Farhat [2011] propose interpolation on the tangent space to the manifold of orthogonal matrices to preserve orthogonality of } V \text{ and } W \text{ in parametric model reduction. In the context of LPV systems, it seems more natural to adopt the viewpoint of subsuming the projection matrices in the reduced-order model.}\]
Matlab’s routine `lyapchol` or the Bartels-Stewart [1972] algorithm (e.g., implemented in Matlab’s routine `sylvester`) can be used to calculate local Gramians with a number of floating point operations which scales on the order of $n^3$ [Bartels & Stewart 1972]. This computational cost is significantly less than the cost of the LMI approach which scales on the order of $n^6$ [Boyd & Vandenberghe 2004, Sec. 11.8.3]. As Proposition 3.4 already makes use of local information to determine the basis space, it seems natural to consider a procedure that is completely based on local information in a second approximation step. This approach is summarized as Algorithm 2.

**Algorithm 2** Approximate Balancing and Truncation by Parameter-Varying Oblique Projection using Local Information

**Input:** LPV system \( \{P_k\}_{k=1}^{n_{\text{grid}}} \) with state space data \( \{A_{\rho_k}, B_{\rho_k}, C_{\rho_k}, D_{\rho_k}\}_{k=1}^{n_{\text{grid}}} \), desired order \( n_z \)

**Output:** LPV system \( \{P_{\text{red},k}\}_{k=1}^{n_{\text{grid}}} \) with state space data \( \{A_{\text{red},\rho_k}, B_{\text{red},\rho_k}, C_{\text{red},\rho_k}, D_{\text{red},\rho_k}\}_{k=1}^{n_{\text{grid}}} \)

```plaintext
for k = 1 to n_{\text{grid}} do
    L_{c,\rho_k} ← lyapchol(A_{\rho_k}, B_{\rho_k})^T % Cholesky factor of local Gramian
    L_{o,\rho_k} ← lyapchol((A_{\rho_k}^T, C_{\rho_k})^T)
    (U, ⋆, ⋆) ← svd(L_{c,\rho_k} L_{o,\rho_k}) % singular value decomposition
    (U_1, ⋆, ⋆) ← svd(L_{c,\rho_k} U(1:n_x,1:n_z))
    Q(1:n_x, 1+n_z : k-1 : n_z, k) ← U_1(1:n_x, 1:n_z) % concatenate basis
end for

(Q, ⋆, ⋆) ← svd(Q)
V ← Q(1:n_x, 1:n_z) % define test space

for k = 1 to n_{\text{grid}} do
    (q, r) ← qr(L_{o,\rho_k} V)
    W_{\rho_k} ← L_{o,\rho_k} q (r)^{-1} % W(\rho) = X_o(\rho) V (V^T X_o(\rho) V)^{-1} % define test space
    A_{\text{red},\rho_k} ← W_{\rho_k}^T A_{\rho_k} V % calculate reduced-order model
    B_{\text{red},\rho_k} ← W_{\rho_k}^T B_{\rho_k}
    C_{\text{red},\rho_k} ← C_{\rho_k} V
    D_{\text{red},\rho_k} ← D_{\rho_k}
end for
```

The first step in this case consists in calculating approximations to generalized Gramians from the Cholesky factors \( \{L_{c,\rho_k}\}_{k=1}^{n_{\text{grid}}} \) and \( \{L_{o,\rho_k}\}_{k=1}^{n_{\text{grid}}} \) of local Gramians at each grid point \( \rho_k \). It should be noted that the complete Gramians are never required and that thus the algorithm of Hammarling [1982] can be used. The remaining steps are identical to those in Algorithm 1. It should be clear that the Gramians in this local approximation can also be calculated by any of the other methods described in Section 3.1.3, i.e., frequency-weighted Gramians, low-rank approximations, and empirical Gramians can be used.

### 3.3.3 Stability Considerations and Limitations

First, a result from de Villemagne & Skelton [1987] is adapted to show that a reduced-order model preserves stability when the generalized observability Gramian in accordance with Definition 2.3 is used to define the parameter-varying oblique projection.
Theorem 3.5. A reduced-order model obtained through Algorithm 1 is exponentially stable.

Proof. Multiplying the LMI (2.23b) for the original system from the left by $V^T$ and from the right by $V$ results in

$$V^T A \left( \frac{\partial V}{\partial p_i} \right) X_o(p) V + \sum_{i=1}^{n_p} \left. \frac{\partial V^T X_o(p)}{\partial p_i} \right|_p q_i + V^T C \left( \frac{\partial C}{\partial p_i} \right) C \left( \frac{\partial C}{\partial p_i} \right) V \prec 0. \quad (3.53)$$

Using $A_{\text{red}}(p) = (V^T X_o(p))^{-1} V^T X_o(p) A V$ and $C_{\text{red}}(p) = C(p) V$, it can be shown that (3.53) is equivalent to

$$A_{\text{red}}^T(p) (V^T X_o(p)) V + (V^T X_o(p)) V A_{\text{red}}(p) + \sum_{i=1}^{n_p} \left. \frac{\partial V^T X_o(p)}{\partial p_i} \right|_p q_i + C_{\text{red}}^T(p) C_{\text{red}}(p) \prec 0. \quad (3.54)$$

Since $X_o(p)$ is symmetric positive definite, so is $X_{o,\text{red}}(p) := V^T X_o(p) V$. Making further use of the fact that $C_{\text{red}}(p) C_{\text{red}}(p) \succeq 0$, it follows that

$$A_{\text{red}}^T(p) X_{o,\text{red}}(p) + X_{o,\text{red}}(p) A_{\text{red}}(p) + \sum_{i=1}^{n_p} \left. \frac{\partial X_{o,\text{red}}(p)}{\partial p_i} \right|_p q_i \prec 0 \quad (3.55)$$

which by Theorem 2.1 establishes exponential stability of the reduced-order model.

If the observability Gramian is calculated from point-wise solutions as in Algorithm 2, stability guarantees for the reduced-order model are lost, since the interpolated matrix functions do not necessarily satisfy the LMIs (3.22) due to the rate term. It is nevertheless possible to guarantee stability of the reduced-order model for “slowly” varying parameters when the Gramians are obtained as solutions to the Lyapunov equations (3.34). Following the same steps as in the proof of Theorem 3.5 for the Lyapunov equation (3.34b), it follows that at every grid point

$$A_{\text{red}}^T(p) (V^T X_o(p)) V + (V^T X_o(p)) V A_{\text{red}}(p) + C_{\text{red}}^T(p) C_{\text{red}}(p) = 0. \quad (3.56)$$

Consequently $A_{\text{red}}$ at each grid point has all its eigenvalues in the left half plane. This property guarantees stability for sufficiently slow parameter variation as was shown by Desoer [1969]. In case frequency-weighted Gramians, low-rank approximations, or empirical Gramians are used, even this weaker stability property cannot be guaranteed anymore. It is however worth noting that an a posteriori stability test can be performed using Theorem 2.1. Since this only involves the reduced-order model, the calculation is likely to be tractable, even if it is not tractable for the full-order model.

The computational cost of Algorithm 1 is the same as the current state-of-the-art method and scales on the order of $n_u^6$ [Boyd & Vandenberghe 2004]. The cost of Algorithm 2 is dominated by solving the two Lyapunov equations and calculating the SVD at all grid points. The number of floating point operations thus scales on the order of $n_u^3$ [Bartels & Stewart 1972]. Ultimately, this also limits the state dimension of problems that can be addressed by this algorithm. However, this computational cost is significantly less than the cost of the state-of-the-art method. If low-rank approximations or empirical Gramian approximations are used, the computational cost is further reduced.
3.3.4 Numerical Example

The nonlinear mass-spring-damper system of Example 2.1 (p. 14) is again used to compare the proposed approach with the state-of-the-art LMI-based method. The same reduced-order model from balancing and truncation as in Section 3.2.4 is used such that a cross-comparison of the two novel methods is also possible. The proposed Algorithm 2 which uses local information takes $0.1$ seconds compared to the $2.5$ hours required for the LMI solution. It nevertheless results in an almost identical reduced-order model as confirmed by Figure 3.2.

![Frequency response at frozen parameters and simulation results](image)

(a) Frequency response at frozen parameters $\rho = 0, 1, 2$ with relative error $|P - P_{\text{red}}|/|P|$ shown as dotted lines.

(b) Simulation.

Figure 3.2: Comparison of original LPV model (--- 100 state variables) and reduced-order model from LMI method (- - - 3 state variables) and proposed algorithm (--- 3 state variables).

Figure 3.2a shows that the local error at the grid points is of comparable magnitude for both methods. A slightly lower absolute error of 0.043 compared to 0.083 is achieved by the proposed method. This reflects the additional freedom of the parameter-varying oblique projection over the constant projection of the LMI method. The time-domain simulation results shown in Figure 3.2b are completely indistinguishable. Most importantly, the upper bound on the induced $L_2$-norm of the error system (3.38) is of comparable magnitude, both for bounded and unbounded parameter variation rates. Table 3.2 summarizes these results.

<table>
<thead>
<tr>
<th>Table 3.2: Comparison of reduced-order models.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPV Balancing</td>
</tr>
<tr>
<td>Computation Time$^1$</td>
</tr>
<tr>
<td>Local $H_\infty$-norm error</td>
</tr>
<tr>
<td>$L_2$-norm error bound</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

$^1$on a 64-bit desktop PC with 3.4 GHz 8-core CPU and 8 GB RAM
4 Control Design by Loopshaping

This chapter reviews methods for control system design rooted in the paradigm of loopshaping to achieve desirable sensitivity properties. Its goal is to provide systematic design guidelines and a problem formulation with a manageable number of comprehensible design parameters. A particular emphasis is put on the importance of a suitable two-degrees-of-freedom control structure to separate feedback requirements such as robustness and disturbance rejection capabilities from tracking requirements. Further, links between several common design methods are provided in order to create a unified and intuitive design perspective. The techniques of this chapter prove particularly useful for aeroservoelastic control applications where an important objective is to provide damping augmentation and a high level of robustness with respect to various sources of uncertainty needs to be achieved. Two detailed design studies are presented in Section 5.3 and 5.4.

4.1 Control System Structure

A basic feedback control loop is depicted in Figure 4.1. It consists of a plant and a compensator through which an error signal \( e \) is processed to generate a control signal \( u \). This error signal is the difference between reference input \( r \) and the measured output \( y \) which is corrupted by noise \( n \). Further, a disturbance \( d \) is present. The control system consists only of a single compensator and is hence referred to as a single-degree-of-freedom control system. The only “degree of freedom” is the choice of how the error is mapped into a control signal.

![Figure 4.1: Schematic of a single-degree-of-freedom control loop.](image)

\[ r \rightarrow e \rightarrow u \rightarrow \text{Compensator} \rightarrow \text{Plant} \rightarrow y \]

\[ d \rightarrow \text{Disturbance Model} \rightarrow \text{Plant} \rightarrow y \]

\[ n \]

\[ e = r - y \]

Compatible dimensions are assumed without loss of generality. In case only the first \( n_r \) measured output signals are to be controlled, i.e., \( r(t) \in \mathbb{R}^{n_r} \) is the reference for a subset of the outputs \( y(t) \in \mathbb{R}^{n_y} \) with \( n_r < n_y \), the signal \( r \) is elongated as \( y_{\text{ref}} := \begin{bmatrix} 1_{n_r} \\ 0 \end{bmatrix} r \) and the error is calculated as \( e = y_{\text{ref}} - y \).
A more general control loop is depicted in Figure 4.2. The control system there consists of two different subsystems. The reference signal is processed by a feedforward compensator and the measured signal is processed by a feedback compensator. While the first loop only operates on the difference of the measurement and the reference signal, the second loop processes both signals independently. The second loop is hence referred to as a two-degrees-of-freedom control loop. It is superior to a single-degree-of-freedom control loop which can be recovered as a special case if the feedforward and feedback compensator are selected to be equal.

![Figure 4.2: Schematic of a two-degrees-of-freedom control loop.](image)

A yet more complex control loop is shown in Figure 4.3. It consists of five different subsystems. The reference signal is still passed through a feedforward compensator, but the output of this feedforward compensator is compared to the output of the feedback compensator to form a generalized error which is then processed by a cascade compensator to calculate a control signal. Additional contributions to the control signal are further directly calculated from the reference command and the measurement signal, bypassing the cascade compensator. The naming convention for these subsystems is inconsistent and ambiguous in the literature and actually not important. The important point is that the control loop still has only two degrees of freedom as there are only two independent signals: the reference command and the measurement.

![Figure 4.3: A more complex two-degrees-of-freedom control loop.](image)
4.1 Control System Structure

Hence, any five subsystems can be combined into the structure of Figure 4.2 without altering any properties of the control loop [e.g., Horowitz 1963, § 6.1, Lunze 1988, Sec. 2.3]. The following example illustrates this.

**Example 4.1 (Two-Degrees-of-Freedom PID Control).** To illustrate the use of different two-degrees-of-freedom control loop architectures, a standard proportional-integral-derivative (PID) controller [e.g., Åström & Hägglund 1995] is considered. In short, the proportional part is used to specify the bandwidth of the control loop, the integral part guarantees steady-state accuracy and the derivative part adds phase lead where required. The PID controller is described as the series interconnection of a PI stage $C_{PI}(s)$ and a lead compensator stage $C_D(s)$ as

$$C_{PI}(s) = \frac{k_p s + k_i}{s} \quad \text{and} \quad C_D(s) = \sqrt{\frac{1 + b^2}{1 + \frac{\bar{\omega}}{b}} \frac{s + \bar{\omega}/b}{s + \bar{\omega}}},$$

with positive given constants $k_p$, $k_i$, $\bar{\omega}$, and $b$. The phase lead provided by $C_D(s)$ is critical for robustness and to achieve a sufficient bandwidth for satisfactory disturbance rejection, but it is usually adverse for following step changes in the reference signal: the well-known derivative kick causes overshoot. Assuming a controller was designed to satisfy requirements on robustness and disturbance rejection (e.g., using classical loopshaping as in Section 4.2.1), there are several ways of mitigating this effect by using a two-degrees-of-freedom control structure. Four completely equivalent possibilities are shown in Figure 4.4.

![Diagram](image-url)

**Figure 4.4:** Some equivalent two-degrees-of-freedom PID controllers.
The loop shown in Figure 4.4a uses the structure of Figure 4.2. The controller is implemented such that the feedback properties are determined by the full PID controller, while the feedforward compensator lacks the derivative part. The configuration in Figure 4.4b is sometimes referred to as using a prefilter. The lead compensator dynamics are canceled in the forward path by first processing the reference signal through a lag compensator. In Figure 4.4c, the additional feedforward compensator is used to subtract the control signal that corresponds to the derivative kick. In Figure 4.4d, the controller is implemented as a cascade in which only the output measurement is processed through the lead compensator and the resulting control signal bypasses the PI-part. All of these loops represent the exact same two-degrees-of-freedom controller and yield identical input-output maps.

In this example, the only purpose of the two-degrees-of-freedom structure is to avoid the derivative kick. Similar phenomena are also common with more complex high-order controllers, although they are far less obvious in that case. The feedforward path of a properly designed two-degrees-of-freedom controller will therefore typically lack the lead action of the feedback path or adequately compensate for it. Two-degrees-of-freedom controllers can also be used for a variety of different purposes. For example, it is possible to achieve different feedforward and feedback bandwidths or to apply specific signal modifications such as notching out frequencies in the feedforward path.

In this thesis, the structure of Figure 4.2 is used for all control systems. The mathematical model to describe this structure is shown in Figure 4.5. The plant model is denoted \( P \) (\( n_y \) outputs, \( n_u \) inputs) and a disturbance model \( P_d \) (\( n_y \) outputs, \( n_d \) inputs) is included. The disturbance model can encompass output disturbances (\( P_d = I_{n_y} \)) but also provides the possibility that disturbances are filtered through part of the plant dynamics. In this case, a common realization of \( P \) and \( P_d \) is realistic, with the limiting case of load disturbances that occur directly at the plant input (\( P_d = P \)). Feedforward compensator \( C_{FF} \) (\( n_u \) outputs, \( n_r \) inputs) and feedback compensator \( C_{FB} \) (\( n_u \) outputs, \( n_y \) inputs) are subsumed in the controller \( K \), i.e.,

\[
\begin{align*}
    u &= \begin{bmatrix} C_{FF} & C_{FB} \end{bmatrix} \begin{bmatrix} r \\ -y \end{bmatrix},
\end{align*}
\]

Figure 4.5: Block diagram of two-degrees-of-freedom control loop.
One particular advantage of this structure is that feedback properties solely depend on $C_{FB}$ and that the single-degree-of-freedom case is trivially recovered by setting $C_{FF} = C_{FB} = C$. This simplifies notation while the discussion of feedback properties remains valid both for two-degrees-of-freedom and single-degree-of-freedom control systems.

The input-output maps for a control loop with such a controller are

$$
y = \frac{R}{1 + PC_{FB}} r - (I + PC_{FB})^{-1} PC_{FB} n + (I + PC_{FB})^{-1} P_{d} d \tag{4.2a}
$$

$$
u = \frac{S_{o} P_{d}}{1 + PC_{FB}} r - (I + PC_{FB})^{-1} PC_{FB} n - (I + PC_{FB})^{-1} C_{FB} P_{d} d \tag{4.2b}
$$

How controllers are designed to specifically adjust these input-output maps is discussed in the subsequent sections. First, the terminology of sensitivity functions is introduced. The elementary sensitivity functions of a feedback loop are output sensitivity $S_{o} = (I + PC_{FB})^{-1}$ and input sensitivity $S_{i} = (I + CP_{FB})^{-1}$. Further, the complementary output sensitivity $T_{o} = (I + PC_{FB})^{-1} PC_{FB}$ and complementary input sensitivity $T_{i} = (I + CP_{FB})^{-1} CP_{FB}$ are defined based on the relations $S_{o} + T_{o} = I$ and $S_{i} + T_{i} = I$. If SISO loops are considered, $S_{i} = S_{o} =: S$ and $T_{i} = T_{o} =: T$. The concatenation $S_{o} P_{d}$ of the disturbance model with the output sensitivity is referred to as disturbance sensitivity. If $P_{d} = P$, i.e., load disturbances at the plant input are assumed, the additional relation $S_{o} P = PS_{i}$ holds and $S_{i} CP_{FB} P$ is the same as $T_{i}$. All of these sensitivity functions depend solely on the feedback compensator. The concatenation $S_{i} CP_{FB}$ of a compensator (either feedforward or feedback) and the input sensitivity function is referred to as control sensitivity. In a two-degrees-of-freedom setting, reference tracking is governed by a reference transmission function $R = S_{o} PC_{FF}$ such that a tracking error function $S_{r} = \left[ \begin{array}{c} r \\ n \end{array} \right] - R$ can also be defined.\(^2\) For a single-degree-of-freedom control loop, Equations (4.2) simplify to

$$
y = \frac{T_{o}}{S_{i} CP_{FB}} r - (I + CP_{FB})^{-1} CP_{FB} n \tag{4.3a}
$$

$$
u = \frac{S_{r} P_{d}}{S_{i} CP_{FB}} r - (I + CP_{FB})^{-1} CP_{FB} n \tag{4.3b}
$$

In this case, the relation $S_{i} CP_{FB} = CS_{o}$ for the control sensitivity holds and tracking is governed by the complementary sensitivity function.

\(^2\)Again, it is without loss of generality assumed that the first $n_r$ outputs are to be tracked.
4.2 Loopshaping Design

This section reviews loopshaping as a design perspective for control systems. Loopshaping is a model-based control design paradigm which seeks to achieve desirable closed-loop sensitivity functions, i.e., input-output maps (4.2) and (4.3). A controller is designed such that these sensitivity functions are “shaped” in ways that correspond to desired properties of a control system. Such properties are, e.g., specific disturbance attenuation levels ($S_o P_d$), tracking capabilities ($R$ or $T_o$), and a limited frequency range of control activity ($S_i C_{FB}$ and $S_i C_{FF}$). These properties are deduced from design specifications, e.g., as detailed in Example 4.2. Models of the plant, sensors, and actuators are only approximations of the reality and hence inaccurate. Robustness with respect to these uncertainties is therefore another key requirement for any control system. Loopshaping addresses this problem in a very general way and provides robustness with respect to generic (frequency domain) uncertainty descriptions (such as multiplicative uncertainties through $T_o$ and $T_i$).

Example 4.2. To briefly illustrate how classical design specifications are translated into sensitivity functions, a single-degree-of-freedom control loop for a SISO system is considered. The desired response to a reference signal is given in terms of a second-order system $T_{ideal} = \frac{1}{s^2 + 4s + 1}$ (damping $\zeta = 0.7$, natural frequency $\omega_n = 1$ rad/s). Such a requirement corresponds to a step response with a 90% rise time of 2.6s, a 1% settling time of 6.5s, an overshoot of less than 5%, and zero steady-state error (Figure 4.6a). Using the relationship $S + T = 1$, the desired sensitivity function for this response is $S_{ideal} = 1 - T_{ideal} = \frac{s}{s^3 + 1.4s + 1}$. It is shown in Figure 4.6b. The desired sensitivity hence has wash-out characteristic with a 20 dB/dec slope, crosses the 0 dB line at a frequency of 0.7 rad/s and has a peak magnitude of slightly more than 2 dB. It is important here to note that the “bandwidth” (i.e., 0 dB crossing) of the desired sensitivity function is below the bandwidth of the desired response model (i.e., $-3$ dB frequency) such that care must be taken when specifications are given in terms of bandwidths. Assuming an available control bandwidth of 5 rad/s, the controller should roll off beyond that frequency to avoid high-frequency control action. As $S_{ideal} \approx 1$ for frequencies above 5 rad/s, $C S \approx C$. Thus, a roll-off in $C$ can equivalently be expressed as a roll-off in $C S$. Finally, robust stability should be achieved. A possible assumption is that the plant model is known only with an accuracy of about $\pm 50\%$ up to a frequency of 3 rad/s and that due to unmodeled high-frequency dynamics the accuracy beyond that frequency degrades with $100\%$/dec. Such an uncertainty can be described as a multiplicative perturbation $M(s) = \Delta (1/3 s + 1)$, where $\Delta \in H_\infty$ is a stable uncertain dynamic system with $\|\Delta\| < 1$ [e.g., Zhou et al. 1995, Cha. 9]. The $\pm 50\%$ accuracy in the operating range can also be interpreted as a gain margin requirement of 6 dB, although the multiplicative uncertainty is more general and also includes possible phase shifts as detailed in Section 4.4. Using the small-gain theorem [e.g., Zhou et al. 1995, Theorem 9.1, p. 212], the perturbed closed loop is robustly stable if and only if $\|T M\| < 1$. Therefore, the magnitude of $T$ has to be less than the magnitude of $M^{-1}$ at all frequencies (Figure 4.6c).
Example 4.2 illustrates that a variety of important control objectives can be expressed in terms of sensitivity functions and, in particular, as the requirement to reduce the sensitivity $S$ (compared to $S = 1$ for an open-loop system). A fundamental limitation on sensitivity reduction is provided by Bode’s sensitivity integral [e.g., Skogestad & Postlethwaite 2005, Theorem. 5.1, pp. 168]. For SISO systems\(^4\), the integral is

$$
\int_0^\infty \ln |S(j\omega)| \, d\omega = \pi \sum_{i=1}^{n_p} \text{Re}(p_i), \quad (4.4)
$$

where $\{p_i\}_{i=1}^{n_p}$ are the unstable poles of the loop transfer function $L = PC$. As emphasized by Stein [2003], the integral (4.4) must actually be satisfied over a finite frequency range for any practical control system, where the upper limit is imposed by available actuator bandwidth and possibly by non-minimum phase zero dynamics of the plant [e.g., Skogestad & Postlethwaite 2005, Theorem. 5.2, pp. 170]. Therefore, the quintessential consequence of Equation (4.4) is that sensitivity is a preserved quantity. Feedback necessarily increases sensitivity (i.e., $|S| > 1$) at some frequencies if it reduces sensitivity at others: the waterbed effect. Thus, the sensitivity is not actually reduced, but rather redistributed or shaped by the control system. Equation (4.4) further highlights the important fact that unstable systems are inherently more difficult to control and that the difficulty increases with faster unstable dynamics [Stein 2003].

The $H_\infty$-norm as given in Definition 2.5 and, respectively, the induced $L_2$-norm (Definition 2.4), are used throughout the chapter as a performance criterion. These norms are useful in this context essentially because they are induced norms and hence they are submultiplicative, i.e., $\|WG\| \leq \|W\| \|G\|$. An important consequence is that $\|WG\| < \gamma$ implies $\sigma_{\text{max}}(G) < \gamma \sigma_{\text{max}}(W^{-1})$ (or for SISO systems $|G| < \gamma |W^{-1}|$) at every frequency. That is, the gain of $G$ is upper-bounded by the gain of $\gamma W^{-1}$ so that for a known $W$ the

\(^4\)For multi-input multi-output (MIMO) systems, several generalizations with the same essence exist [e.g., Zhou et al. 1995, Sec. 6.2, Skogestad & Postlethwaite 2005, Sec. 6.2.3].

---

(a) Step response requirement. (b) Complementary sensitivity (---) and equivalent sensitivity (——) requirement. (c) Bound (—) on complementary sensitivity (——) for multiplicative uncertainty.

Figure 4.6: Examples of translating design requirements to sensitivity functions.
“shape” of $G$ over frequency is known. This property was already used in Example 4.2 to determine the bound on the complementary sensitivity for robust stability.

The description in the following subsections focuses mostly on feedback properties and hence on the feedback part $C_{FB}$ of a two-degrees-of-freedom control system. In the selected structure (Figure 4.5), feedback properties are solely determined by $C_{FB}$. Further, $C$ and $C_{FB}$ are completely equivalent such that all results for single-degree-of-freedom controllers directly carry over to two-degrees-of-freedom controllers.

### 4.2.1 Classical Loopshaping

Classical loopshaping design for SISO systems [e.g. Horowitz 1963, §§ 5.5–5.11, Doyle et al. 1990, Sec. 7–8] is based on “shaping” the loop transfer $L = PC$. The (open-)loop transfer $L$ is approximately the inverse of the (closed-loop) sensitivity $S = \frac{1}{1+L}$ if $|L| \gg 1$ and approximately equals the complementary sensitivity $T = \frac{1}{1+L}$ if $|L| \ll 1$. Hence, the compensator $C$ is selected such that the loop transfer has high gain at low frequencies and low gain at high frequencies. For stability, the loop transfer function must further have sufficient phase margin around the crossover frequency, i.e., the frequency $\omega_c$ where $|L(j\omega_c)| = 1$. This requirement translates to a slope of around $-20$ dB/dec in the crossover region for stable, minimum phase systems due to Bode’s gain-phase relationship [e.g., Skogestad & Postlethwaite 2005, p. 19]

$$\angle L(j\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{d \ln |L(j\omega)|}{d \ln \omega} \ln \left| \frac{\omega + \omega_0}{\omega - \omega_0} \right| \frac{1}{\omega} \frac{d \omega}{d \ln \omega} \approx \frac{\pi}{2} \left. \frac{d \ln |L(j\omega)|}{d \ln \omega} \right|_{\omega=\omega_0}. \quad (4.5)$$

Thus, an ideal loop resembles an integrator $\frac{1}{s}$. Loopshaping vividly illustrates inversion-based compensator design as $C = \frac{1}{s} P^{-1}$ immediately results in the desired $\frac{1}{s}$-shape. Even if the complete plant model cannot be inverted, it is often possible to cancel poles and zeros around the crossover frequency such that the $\frac{1}{s}$-shape is attained in a frequency range around crossover. This inversion can, however, result in poor disturbance rejection properties. In fact, a design for disturbance attenuation is less straightforward with loopshaping as the disturbance sensitivity $SP_d = \frac{P_d}{1+L}$ is not immediately related to the loop transfer. For uniform disturbance attenuation across all frequencies, $L \approx P_d$ is desirable [Skogestad & Postlethwaite 2005, Sec. 2.6.4]. Similarly, disturbance rejection up to a frequency $\omega_c$ can be achieved through a loopshape $L \approx \frac{1}{s} P_d$ and hence with a compensator $C = \frac{1}{s} P^{-1} P_d$ [ibid.].

Single control loops are, in general, easily designed through a graphical interface such as Matlab’s sisotool. The designer has direct control over the compensator and can make adjustments to satisfy magnitude and bandwidth limitations. Robustness is also intuitively incorporated through the classical gain and phase margins, which can be verified during design. A useful design strategy is to sequentially modify “compensator stages” as proposed by Packard [2011]. These stages are constructed such that they address only one of the three frequency regions, i.e., low-frequency region (below $0.3 \omega_c$), crossover region (between $0.3 \omega_c$ and $3 \omega_c$), and high-frequency region (beyond $3 \omega_c$), without interfering with each other. A typical design sequence consists of the following steps: First, the
bandwidth of the loop transfer is set by selecting the gain $k = \omega_c / |P(j\omega_c)|$ as an initial compensator. Second, integral action is added in the low-frequency region through an integral stage $\frac{3\omega_c}{s + 3\omega_c}$, which does not affect the crossover frequency. A low-pass filter $\frac{3\omega_c}{s + 3\omega_c}$ can also be added such that the compensator rolls off in the high-frequency regime, again without altering the crossover frequency. Finally, the slope approximately one decade around crossover is adjusted to about $-20\,\text{dB/dec}$ to provide sufficient phase margin. This then comes at the expense of a lower gain in the low-frequency regime and a higher gain in the high-frequency regime and therefore can be seen to trade off performance in favor for robustness.

### 4.2.2 Optimal Loopshaping

For MIMO systems, classical loopshaping is of limited use as the simple magnitude-phase relation (4.5) does not hold anymore. Thus, it becomes difficult to achieve stability. To address this issue but retain the intuitive and comprehensible design procedure of loopshaping, McFarlane & Glover [1990, 1992] extended ideas originally formulated by Doyle & Stein [1981] and proposed an “$H_\infty$ loopshaping design procedure” which guarantees stability through an $H_\infty$-norm optimization problem. The procedure consists of a design phase, in which pre- and post-compensators $C_1$ and $C_2$ are selected in accordance with classical loopshaping guidelines to achieve a desirable loop transfer $C_1PC_2$. This loop transfer is defined as a new “shaped plant” $\hat{P} := C_1PC_2$ and a robust stabilization problem

$$
\min_{C_\infty} \left\| \begin{bmatrix} \hat{S}_o & \hat{S}_o \hat{P} \\ \hat{C}_\infty \hat{S}_o \hat{C}_\infty^{-1} & \hat{C}_\infty \hat{S}_o \hat{C}_\infty^{-1} \end{bmatrix} \right\|_1
$$

is solved. This problem corresponds to shaping the sensitivity functions of the interconnection shown in Figure 4.7a. The actual controller $C := C_2C_\infty C_1$ is obtained as a combination of the compensators $C_1$ and $C_2$ and the $H_\infty$-controller $C_\infty$, resulting in the interconnection of Figure 4.7b. Formally, the controller $C_\infty$ optimizes robustness in the gap metric, but the practically relevant corollary of the procedure is that the singular values of the resulting loop transfer $PC$ remain close to the specified loop transfer $C_1PC_2$ [McFarlane & Glover 1992, Theorems 3.1 & 3.3]. Specifically, the decrease in gain in the low-frequency regime is lower bounded and the increase in gain in the high-frequency range is upper bounded. The trade-off between robustness and performance, exemplified by the design of the lead compensator in classical loopshaping, is hence formalized and solved in an optimal way. All other steps of classical loopshaping such as providing high gain in the low-frequency range, roll-off in the high-frequency range, and setting the crossover bandwidth are still done “by hand” through choice of the compensators $C_1$ and $C_2$. Often $C_2$ is selected as a PI stage to include integral action and to set the desired bandwidth while $C_1$ is selected as a low-pass filter stage to include roll-off.
Textbooks such as Zhou et al. [1995], Zhou & Doyle [1998], and Skogestad & Postlethwaite [2005] refer to the McFarlane-Glover design strategy as the de facto standard for robust control design. Comprehensive guidelines for selecting pre- and post-compensators are provided by Hyde [1991, Cha. 4], Zhou & Doyle [1998, Cha. 16], and Skogestad & Postlethwaite [2005, Sec. 9.4]. The one-shot design of two-degrees-of-freedom controllers is also possible by including a reference model which should be matched [Limebeer et al. 1993]. Alternatively, the freedom in implementing different parts of the controller such that a two-degrees-of-freedom structure results can be just as useful. The robust stabilization controller (which usually provides phase lead) can be implemented in the feedback path as shown in Figure 4.7c to avoid overshoot, while $C_2$ includes integral action and is implemented as a cascade compensator [e.g. Hyde 1991, Sec. 4.2.3, Skogestad & Postlethwaite 2005, Sec. 9.4.3].

Nevertheless, the method has its limitations. In particular, the disturbance model necessarily equals the plant model, i.e., disturbances are assumed to always occur at the plant input. While such load disturbances are often a reasonable assumption, the design of cascaded control systems where the “plant model” includes lower-level control systems can be complicated as load disturbances in this case correspond to disturbances within the control system. Further, the relation between the loop transfer singular values and multivariable design objectives is, in general, not necessarily obvious. In particular, it can be difficult to achieve different bandwidths for different outputs. Even more difficulties are encountered when the plant is non-square, which includes the case in which a subset of the measured outputs is to be tracked.
4.2 Loopshaping Design

4.2.3 Mixed Sensitivity Loopshaping

The second prevalent control design approach that makes use of $\mathcal{H}_\infty$-norm minimization is referred to as mixed sensitivity loopshaping. It directly specifies performance requirements for a combination of closed-loop sensitivity functions in terms of weighting filters. A controller is found from the optimization problem

$$
\min_K \|G_{CL}(K)\|, \tag{4.7}
$$

where $G_{CL}$ is a specified generalized closed-loop interconnection that includes weighting filters as well as fictitious inputs and outputs to define the involved sensitivity functions [e.g., Zhou et al. 1995, Sec. 5.7]. The name mixed sensitivity loopshaping is used ambiguously throughout the literature, often synonymously with a weighted $S_o/CS_o$ or $S_o/CS_o/T_o$ minimization [Kwakernaak 1993]. A weighted interconnection that involves only $S_o$ and $CS_o$ is sometimes called two-block design. Such a setup assumes that disturbances only occur at the output of the plant. The two-block design is known to result in an inversion controller which potentially lacks robustness and sufficient disturbance attenuation capabilities [Sefton & Glover 1990, Cao & Hori 1997]. Christen & Geering [1997] therefore proposed an alternative two-block design using $PS_t$ and $T_t$. Such a setup assumes that disturbances occur only at the plant input. It usually requires high-order weights to enforce a roll-off in the controller\(^4\) and it can be difficult to achieve satisfactory tracking behavior, in particular with ill-conditioned plants for which $S_o$ and $S_i$ can be vastly different. If an interconnection which represents the four sensitivity functions $S_o$, $S_oP_o$, $CS_o$, and $T_i$ is used, the resulting control problem is sometimes referred to as a four-block design [Sefton & Glover 1990]. Such a design is similar to the McFarlane-Glover design technique discussed in Section 4.2.2, as is subsequently detailed. Different weighting schemes are compared, e.g., by Christen [1996, Cha. 7] and Saupe [2013, Cha. 5]. Skogestad & Postlethwaite [2005] also coined the term “signal-based $H_\infty$ control”, moving “the focus of attention...to the size of signals and away from the size and bandwidth of selected closed-loop transfer functions” [ibid., Sec. 9.3.6]. This very general formulation has the benefit of capturing a variety of control formulations within a single conceptual framework. Unfortunately, it can also obscure the underlying sensitivity minimization and make it all too easy to set up seemingly meaningful, but ill-posed, optimization problems.\(^5\)

In this thesis, the term mixed sensitivity is used for any norm minimization problem that involves multiple sensitivity functions. In particular, it is proposed to use the setup illustrated in Figure 4.8 which corresponds to the input-output maps

$$
\begin{bmatrix}
  z_1 \\
  z_2
\end{bmatrix} =
\begin{bmatrix}
  W_e & 0 & D_e^{-1} & 0 & S_o & S_o P_d \end{bmatrix} \begin{bmatrix}
  C S_o & C S_o P_d \end{bmatrix} \begin{bmatrix}
  D_e & 0 & D_d \end{bmatrix} \begin{bmatrix}
  w_1 \\
  w_2
\end{bmatrix}. \tag{4.8}
$$

The outputs $z_1, z_2$, as well as the inputs $w_1, w_2$ are “fictitious” in the sense that they are necessary to define the input-output map and represent any signal injected or measured at

\(^4\)As $T_i$ inherits the plant model’s roll-off, the weight must have an order of at least the relative degree of the plant [cf. Saupe 2013, Sec. 5.5].

\(^5\)In fact, much of the criticism expressed about the use of state variables as target quantities in control design by Horowitz & Shaked [1975] could be repeated in opposition to this “signal-based” perspective.
that particular point within the loop. The formulation (4.8) involves diagonal weighting filters $W_e$ and $W_u$, as well as diagonal scaling matrices $D_e$ (maximum allowable errors), $D_u$ (maximum allowable inputs), and $D_d$ (maximum expected disturbances). These scalings are usually subsumed in scaled models $\hat{P} = D_e^{-1}PD_e$ and $\hat{P}_d = D_e^{-1}P_dD_e$, leading to the two equivalent representations shown in Figure 4.8. Scalings are critical to ensure that all magnitudes are of comparable size and cross-couplings are of equal importance as highlighted by Skogestad & Postlethwaite [2005, Sec. 1.4].

Figure 4.8: Generalized closed-loop interconnection for single-degree-of-freedom design.

Figure 4.8 also reveals that the scalings can be seen as part of the weights, contrary to the perspective of Skogestad & Postlethwaite [2005] who treat the scalings as part of the model. Doing so facilitates the choice of design parameters. In particular, any static parts of the weights are absorbed into the scalings and standard choices for the frequency-dependent parts $W_e$ and $W_u$ become valid for a wide range of control problems. A default choice for the entries of the diagonal matrix $W_e$ are filters with approximately integral behavior up to a desired bandwidth $\omega_{b,i}$, $i = 1, \ldots, n$, and gain 0.5 beyond that frequency for controlled outputs. The rationale is to decrease (disturbance) sensitivity up to $\omega_b$ and to limit sensitivity degradation beyond that frequency to no more than a factor of two. For additional feedback quantities such as velocities in position control, constant gains of 0.5 are used. Similarly apparent choices also exist for a variety of specific design requirements for disturbance rejection or tracking. If, e.g., ramps are to be tracked, then a second-order weight is used such that the sensitivity function has a $+40\,\text{dB/dec}$ slope in the low-frequency range. If sinusoidal signals need to be tracked or rejected, a resonance
4.2 Loopshaping Design

can be added to the weighting filter, which then translates to a sensitivity reduction at that particular frequency. More general, the internal model principle [Francis & Wonham 1975, 1976] applies and the weights should include a model of the signal which is to be rejected or tracked. The entries for the diagonal weight \( W_u \) can be selected as constant unit gains to limit controller magnitude. If a controller roll-off is desired, filters with unit gain up to the available bandwidth \( \omega_{a,i} \), \( i = 1, \ldots, n_u \) and approximately differentiating behavior beyond that frequency can be used. Another useful choice are inverse band-pass filters which lead to a controller with band-pass characteristic. It is also possible to include resonances to achieve notch-filter-like characteristics. The weighting filters thus mainly reflect design specifications rather than being treated as tuning knobs.

To further motivate the particular structure of Equation (4.8) with the scalings appearing explicitly as weights, the sensitivity functions of the two loops shown in Figure 4.8 are compared. The sensitivity functions of the actual closed loop (Figure 4.8a) are related to those with scaled models \( \tilde{P} = D_e^{-1} P D_u \) and \( \tilde{P}_d = D_e^{-1} P_d D_d \), and a controller \( \tilde{C} = D_u^{-1} C D_c \) (Figure 4.8b) by

\[
\begin{align*}
S_o &= D_e \tilde{S}_o D_e^{-1} \\
\tilde{S}_o P_d &= D_e \tilde{S}_o \tilde{P}_d D_d^{-1} \\
\tilde{C} S_o &= D_u \tilde{C} \tilde{S}_o D_e^{-1} \\
\tilde{C} S_o P_d &= D_u \tilde{C} \tilde{S}_o \tilde{P}_d D_d^{-1}.
\end{align*}
\] (4.9)

Mixing the sensitivity functions in a sensible way requires some care as each is expressed in different units. In particular, \( S_o P_d \) is in \([\text{units of } y]/[\text{units of } d]\) and \( CS_o \) is in \([\text{units of } u]/[\text{units of } y]\), while \( S_o \) is unitless and necessarily unity at high frequencies. A weight \( W_e \) with high-frequency gain above unity is therefore not a meaningful specification for \( S_o \), as it would demand the impossible: a sensitivity of less than unity at high frequencies.

The diagonal elements of \( S_o \) are further independent of scalings, such that this fundamental property dictates the choice of \( W_e \). The scaling \( D_e \) then has to be selected such that cross-couplings are of similar magnitude and \( D_d \) has to be selected such that \( W_e \) also acts sensibly on \( S_o P_d \). In practice, this often means that the plant and disturbance model are scaled to unit gain in the frequency range of interest. Slightly larger gains then emphasize the respective sensitivity functions relative to those with slightly lower gains. For instance, \( D_d \) can be increased to emphasize disturbance rejection (through \( S_o P_d \)) or \( D_u \) can be decreased to lower control activity (through \( CS_o \)). Tuning by means of altering the scalings can thus be related to parameters with an appealing physical interpretation and good initial guesses are usually possible.

It is also instructive to revisit the McFarlane-Glover loopshaping design and to compare the sensitivity functions of the resulting closed-loop (Figure 4.7b) to those that are minimized during the robust stabilization synthesis step (Figure 4.7a). Assuming a square...
A low-pass input weight would relax requirements on weight $w$. The latter clearly is in opposition to the aim of restricting control activity by means of a where $|w|$.

Weights with inverses appearing as left and right factors further means that their frequency capabilities. The only reasonable frequency-dependent input weight would hence represent relax requirements on a signal-based interpretation of argument that a common input weight appears to be of limited use.

Disturbance model restrictive at first. However, any frequency dependence of disturbances can appear in the includes no frequency-dependent input weights (acting on $CS$) according to the McFarlane-Glover design.

It remains to discuss the limitations of the proposed interconnection. Equation (4.6) accordingly shows that the generalized closed-loop interconnection for the McFarlane-Glover technique constitutes a four-block mixed sensitivity formulation.

Rewriting Equation (4.6) accordingly shows that the generalized closed-loop interconnection for the McFarlane-Glover technique constitutes a four-block mixed sensitivity formulation.

$$S_o = C_1^{-1} \tilde{S}_o C_1$$  
$$S_o P = C_1^{-1} \tilde{S}_o \hat{P} C_2^{-1}$$  
$$CS_o = C_2 C_\infty \tilde{S}_o C_1$$  
$$T_i = C_2 \tilde{T}_i C_2^{-1}. $$  

(4.10)

By design $\hat{P} = C_1 PC_2$ has unit gain at the desired closed-loop bandwidth, i.e., at the frequency where $\tilde{S}_o$ is intended to have approximately unit gain. Thus, the problem of different units is inherently avoided in the McFarlane-Glover design.

The structure of the weights with inverses appearing as left and right factors further means that their frequency content has little influence on $S_o$ and $T_o$ in the norm minimization problem. The relevant sensitivity functions for performance in this formulation are $S_o P$ and $C S_o$. The standard choice of $C_1$ as consisting of low-pass filters and $C_2$ as consisting of PI stages means that $C S_o$ is effectively shaped in the high-frequency regime through $C_1^{-1}$ while $S_o P$ is shaped in the low-frequency regime through $C_2$. These requirements are consistent with the choices for $W_o$ and $W_u$ in the proposed mixed sensitivity formulation.

It remains to discuss the limitations of the proposed interconnection. Equation (4.8) includes no frequency-dependent input weights (acting on $w_1$ and $w_2$), which appears restrictive at first. However, any frequency dependence of disturbances can appear in the disturbance model $P_d$. Thus, the only restriction that actually remains is on $w_1$, i.e., on the common right-factor weight for both $S_o$ and $C S_o$. As these two sensitivity functions are to be shaped in different frequency regions ($S_o$ for low frequencies, $C S_o$ for high frequencies), a common input weight appears to be of limited use. In particular, the “signal-based” argument that $w_1$ represents reference signals with low-frequency content is misleading here. A low-pass input weight would relax requirements on $S_o$ beyond the tracking bandwidth where $|S| \approx 1$ anyway, but also relax requirements on $C S_o$ in the high-frequency regime. The latter clearly is in opposition to the aim of restricting control activity by means of a weight $W_u$ which increases with frequency. A similar argument can be developed if instead a signal-based interpretation of $w_1$ as noise is made. Then, a high-pass input weight would relax requirements on $S_o$ in the low-frequency range, inevitably diminishing tracking capabilities. The only reasonable frequency-dependent input weight would hence represent

---

7 The gain of the resulting robust stabilization controller $C_\infty$ is thus also immediately relatable to the achieved norm which in return establishes that the desired loopshape is preserved.

8 These choices are nonrestrictive and it is readily verified that the same is true if the compensators are switched, i.e., if $C_2$ consists of low-pass filters (such that $C_2^{-1}$ has differentiating behavior at high frequencies) and $C_1$ consists of PI stages.

9 An exception to this statement is the use of an input weight in $S_o/K S_o$ minimization to achieve partial pole placement by Tsay et al. [1992], Kwakernaak [1993], and Saupe [2013, Sec. 5.6], although this is a rather specific application and has no immediate connection to loopshaping.
specific frequency contents of disturbances, which is exactly what the second fictitious input $w_2$ and the disturbance model do. The proposed parameterization of the weights and scalings is therefore not actually restrictive in most cases, but has the advantage of providing a low number of intuitive tuning knobs. The interconnection can be regarded as a four-block design where the ratio of $W_e$ and $W_d$ determines a trade-off between the two “ideal” compensators $C = \frac{w_c}{s} P^{-1}$ (for tracking) and $C = \frac{w_c}{s} P^{-1} P_d$ (for disturbance rejection), which were briefly motivated in Section 4.2.1.

A real limitation of the proposed mixed sensitivity interconnection, however, is its restriction to single-degree-of-freedom control. Horowitz [1963, § 5.22] was among the first to emphasize the fundamental inadequacy of a single-degree-of-freedom architecture for tracking control. This inadequacy is also apparent in the single-degree-of-freedom problem formulation (4.8). In particular, disturbance rejection (governed by $S_o P_d$) cannot be adjusted independently of tracking requirements (governed by $T_o$), due to the fundamental limitation $S_o + T_o = I$. While less of a problem if $P_d$ is small when $S$ is large and vice versa, this limitation becomes critical if $P_d$ and $S_o$ are simultaneously large, e.g., around resonances. The modified generalized closed-loop interconnection which extends the proposed setup to two-degrees-of-freedom controllers is illustrated in Figure 4.9.

![Generalized closed-loop interconnection for two-degrees-of-freedom design.](image)

**Figure 4.9:** Generalized closed-loop interconnection for two-degrees-of-freedom design.
It is without loss of generality assumed that the first $n_r$ outputs are to be controlled (tracked) and that the additional $n_y - n_r \geq 0$ outputs of the plant are auxiliary feedback quantities. Figure 4.9 then corresponds to the input-output map

$$G_{CL}(K) = \begin{bmatrix} W_e D_e^{-1} & 0 \\ 0 & W_u D_u^{-1} \end{bmatrix} \begin{bmatrix} S_o & S_o P_d \\ S_o C_{FB} & C_{FB} S_o P_d \end{bmatrix} \begin{bmatrix} [I_{n_r}] - R_s \\ S_i \end{bmatrix} \begin{bmatrix} D_e & 0 & 0 \\ 0 & D_d & 0 \\ 0 & 0 & D_e \end{bmatrix}. \quad (4.12)$$

The first four sensitivity functions (defined through $w_1$, $w_2$, $z_1$, $z_2$) are unchanged in comparison to the single-degree-of-freedom setup (4.8), but it is important to recall that $S_o$ no longer relates to the tracking error. Instead, the tracking error is now governed by the error function $[I_{n_r} 0] - R_s$ (with $R_s = S_o P C_{FF}$), introduced through the new input $w_3$ which represents the point in the loop where the reference signal is injected. The scaling $D_e := [I_{n_r} 0]$ for this input is simply the part of $D_e$ that corresponds to the controlled outputs. Thus, the bandwidth requirement for tracking is still set by the weight $W_e$. The important consequence of the two-degrees-of-freedom structure is that the feedback sensitivity $S_o$ is different from the error function $[I_{n_r} 0] - R_s$, depending on $P_d$. If, e.g., $P_d$ has integral behavior and the weight $W_e$ has integral behavior, the result is a $+40 \text{ dB/dec}$ slope at low frequencies in $S_o$, but a $+20 \text{ dB/dec}$ slope at low frequencies in $[I_{n_r} 0] - R_s$. Thus, the two-degrees-of-freedom control system is able to reject ramp output disturbances, while at the same time excessive overshoot is avoided for step response commands. Different bandwidth requirements for disturbance rejection and tracking can be achieved by modifying the disturbance model or $D_d$ accordingly, as long as the feedback bandwidth is intended to be larger than the tracking bandwidth. This is the case for many application where changes in the reference signal are slow compared to disturbances.

**Modal Performance Representation**

It is sometimes useful to add a non-physical performance output $y_p$ to the model and hence to include sensitivity functions that relate to this fictitious output in the generalized closed-loop interconnection. The augmented plant model is then described by a partitioned system

$$\begin{bmatrix} P \\ P_{p} \end{bmatrix} : \begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \\ y_p = E x. \end{cases} \quad (4.13)$$

The generalized closed-loop interconnection for a single-degree-of-freedom design with such an augmented plant model is shown in Figure 4.10, assuming for simplicity that disturbances occur at the plant input. The corresponding input-output map is

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} W_e D_e^{-1} & 0 & 0 \\ 0 & W_u D_u^{-1} & 0 \\ 0 & 0 & W_p \end{bmatrix} \begin{bmatrix} S_o & S_o P \\ C S_o & -T \end{bmatrix} \begin{bmatrix} D_e & 0 & 0 \\ 0 & D_d & 0 \\ 0 & 0 & D_e \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}. \quad (4.14)$$
Using arbitrary disturbance models leads to a partitioned model \( [P \quad P_d] \) and changes the second column in Equation (4.14) accordingly. The application within a two-degrees-of-freedom structure leads to an interconnection that involves nine sensitivity functions.

One particularly useful augmentation is to penalize velocities in order to improve damping as suggested by Saupe [2013, Sec. 5.3]. Even more specific, Hanel [2001, Sec. 8.4.7] introduced the idea of expressing structural damping requirements in terms of modal velocities, following a “signal-based” approach with additional inputs and outputs for each structural mode in the plant model. This idea can also be incorporated into the mixed sensitivity loopshaping framework. Sensitivity functions of modal velocities appear as sharp peaks and have band-pass characteristic, i.e., zero steady-state and feedthrough gains. They are therefore an ideal description of damping requirements in the framework of induced \( L_2 \)-norm optimization: The objective simply translates to flattening the peaks and other specifications remain largely unaffected due to the narrow frequency band. The modal sensitivity functions hence emphasize, or “expose”, the specific dynamic modes within the minimization problem. An example is shown in Figure 4.11.

To define modal velocity outputs for an LTI model, the additional output matrix \( E \) in (4.13) is calculated from the state space coordinate transformation to modal coordinates. Specifically, let \( \lambda_i, \quad i = 1, \ldots, 2m \) denote the conjugate complex eigenvalues of the matrix \( A \), ordered such that \( [\lambda_1 \lambda_2 \cdots \lambda_m] = [\bar{\lambda}_{m+1} \bar{\lambda}_{m+2} \cdots \bar{\lambda}_{2m}] \). Further, let \( \lambda_i \),
\( i = 2m + 1, \ldots, n \) denote the real eigenvalues of the matrix \( A \) and \( v_i, i = 1, \ldots, n \) denote the normalized eigenvectors to the eigenvalues. The state transformation
\[
\tilde{x} = \left\{ \text{Re}(v_1) \quad \text{Im}(v_1) \quad \cdots \quad \text{Re}(v_m) \quad \text{Im}(v_m) \quad v_{2m+1} \quad v_{2m+2} \quad \cdots \quad v_n \right\}^{-1} x, \quad (4.15)
\]
leads to a decoupled state space representation
\[
\begin{align*}
\dot{\tilde{x}} &= \begin{bmatrix} \dot{\tilde{x}}_1 \\
\vdots \\
\dot{\tilde{x}}_{n-m} 
\end{bmatrix} = \begin{bmatrix} 0 & \cdots & \lambda_i & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & \lambda_{n-m} 
\end{bmatrix} \begin{bmatrix} \tilde{B}_1 \\
\vdots \\
\tilde{B}_{n-m} 
\end{bmatrix} u \\
y &= \begin{bmatrix} \tilde{C}_1 & \cdots & \tilde{C}_{n-m} \end{bmatrix} \begin{bmatrix} \dot{\tilde{x}}_1 \\
\vdots \\
\dot{\tilde{x}}_{n-m} 
\end{bmatrix} + D u
\end{align*}
\]
where \( \tilde{A}_i = \begin{bmatrix} \text{Re}(\lambda_i) & \text{Im}(\lambda_i) \\
-\text{Im}(\lambda_i) & \text{Re}(\lambda_i) \end{bmatrix} \), for \( i = 1, \ldots, m \) and \( \tilde{A}_i = \lambda_{m+i} \) for \( i = m + 1, \ldots, n - m \).

In this form, each of the decoupled subsystems \( (\tilde{A}_i, \tilde{B}_i, \tilde{C}_i) \) describes the dynamics of a single dynamic mode, which can be transformed to a canonical second-order system, i.e.,
\[
\tilde{A}_i = \begin{bmatrix} -\omega_{0,i} & \zeta_i \\
-\omega_{d,i} & -\omega_{0,i} \zeta_i \end{bmatrix} \Phi_i \begin{bmatrix} 0 & 1 \\
-\omega_{0,i}^2 & -2\omega_{0,i} \zeta_i \end{bmatrix} = \tilde{A}_i \quad i = 1, \ldots, m
\]
where \( \omega_{0,i} \) is the natural frequency, \( \zeta_i \) is the damping ratio, and \( \omega_{d,i} = \sqrt{\omega_{0,i}^2 (1 - \zeta_i^2)} \) is the damped frequency of the \( i \)th mode. As in Section 3.2, the scheme of Kailath [1980, Sec. 6.4.6] is used and hence an explicit representation for the second transformation is
\[
\begin{bmatrix} \xi_i \\
\dot{\xi}_i 
\end{bmatrix} = \begin{bmatrix} [0 \ 1] \tilde{B}_i \tilde{A}_i \tilde{B}_i \tilde{A}_i^{-1} \\
[0 \ 1] \tilde{B}_i \tilde{A}_i \tilde{B}_i \tilde{A}_i^{-1} \end{bmatrix} \begin{bmatrix} \tilde{x}_i \\
\dot{\tilde{x}}_i \end{bmatrix} \quad i = 1, \ldots, m. \quad (4.16)
\]
In this form, the two states associated with the \( i \)th mode correspond to modal displacement \( \xi \) and modal velocity \( \dot{\xi} \). The output matrix \( E \) therefore can be obtained as the concatenation of the transformations (4.15) and (4.16) with a selector matrix that extracts the velocity states, i.e.,
\[
E = \begin{bmatrix} [0 \ 1] \tilde{\Phi}_1 & 0_{1,n-2m} \\
\vdots \\
[0 \ 1] \tilde{\Phi}_m & 0_{1,n-2m} \end{bmatrix} \tilde{\Phi} \quad (4.17)
\]
It is of course also possible to only use a subset of the modal velocities as outputs by left multiplication of \( E \) by another selector matrix.
4.3 Synthesis Machinery

Both the McFarlane-Glover method (Section 4.2.2) and the mixed sensitivity method (Section 4.2.3) require to find controllers that minimize the norm of a weighted closed-loop interconnection. This section very briefly reviews some of the relevant synthesis machinery for obtaining such controllers. While Section 4.2 described how meaningful performance specifications can be formulated in terms of sensitivity functions, i.e., how a controller can be designed, the term synthesis refers to the technicality of obtaining a controller that achieves these performance specifications.

Throughout this section, the partitioned state space realization

\[ G: \begin{cases} \dot{x} = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{bmatrix} \begin{bmatrix} x \\ w \\ u \end{bmatrix} \\ z = \begin{bmatrix} x \\ w \end{bmatrix} \end{cases} \]  

(4.18)

for the (open-loop) generalized plant is used. The input-output map \( w \to z \) describes the mixed sensitivity requirements for a closed-loop interconnection with \( u = K v \). Thus, \( v \) denotes the measured signals available to the controller (including references for feedforward) and \( u \) denotes available control inputs. Clearly, the pair \((A, B_2)\) must be stabilizable and \((A, C_2)\) must be detectable, for a stabilizing controller to exist.

A useful special structure can be achieved through loop-shifting and scalings under mild conditions [Safonov & Chiang 1989]. Specifically, it is possible to make \( D_{22} = 0 \), \( D_{21} = [0 \ I] \) and \( D_{12} = [0 \ I]^T \). Further, denote \( C_1^T = [c_1^T, c_2^T] \), \( B_1 = [b_1, b_{12}] \), and \( D_{11} = [d_{11}, d_{111}] = [d_{1111}, d_{1112}] \) such that the partitioning is compatible.

The state space realization (4.18) then takes the form

\[ G: \begin{cases} \dot{x} = \begin{bmatrix} A & B_{11} & B_{12} & B_2 \\ C_{11} & D_{1111} & D_{1112} & 0 \\ C_{12} & D_{1121} & D_{1122} & I \\ C_2 & 0 & I & 0 \end{bmatrix} \begin{bmatrix} x \\ w_1 \\ w_2 \\ u \end{bmatrix} \end{cases} \]  

(4.19)

The partitioning \((z_1, z_2)\) and \((w_1, w_2)\) is induced by the feedthrough components \( w_2 \to v \) and \( u \to z_2 \). The implied rank conditions have clear interpretations in terms of a well-posed control problem: All available measurements of plant outputs are subject to some disturbance (through \( w_2 \to v \)) and all control signals are penalized (through \( u \to z_2 \)).

4.3.1 \( H_\infty \) Controller Synthesis

**Theorem 4.1** (Gahinet & Apkarian [1994, Theorem 4.1] in consideration of the special structure (4.19)). Let \( G \) be an open-loop generalized plant with the special structure (4.19). There exists a linear controller

\[ K: \begin{cases} \dot{x} = A x + B v \\ u = C x + D v \end{cases} \]  

(4.20)
that internally stabilizes the closed-loop interconnection $G_{CL}$ such that $\|G_{CL}\| < \gamma$ if and only if

$$\gamma > \max(\sigma_{\text{max}}(D_{11\bullet}), \sigma_{\text{max}}(D_{11\bullet}^T))$$

and there exist $X_\infty \succ 0$ and $Y_\infty \succ 0$ such that

$$\begin{bmatrix} X_\infty & I \\ I & Y_\infty \end{bmatrix} \prec 0$$

(4.21a)

$$\begin{bmatrix} \Lambda X_\infty - \gamma B_2B_2^T X_\infty C_{11}^T - \gamma I & B_1 - B_2 D_{11\bullet} \\ * & -\gamma I & D_{11\bullet}^T \end{bmatrix} \prec 0$$

(4.21b)

$$\begin{bmatrix} \Lambda Y_\infty - \gamma C_2^T C_2 Y_\infty B_{11} & C_1^T - C_2^T D_{11\bullet}^T \\ * & -\gamma I & D_{11\bullet}^T \end{bmatrix} \prec 0$$

(4.21c)

where $*$ denotes symmetric completion and

$$\Lambda X_\infty := X_\infty (A - B_2 C_{12})^T + (A - B_2 C_{12}) X_\infty,$$

$$\Lambda Y_\infty := Y_\infty (A - B_{12} C_2) + (A - B_{12} C_2)^T Y_\infty.$$

Proof. The proof is given in detail by Gahinet & Apkarian [1994, Theorem 4.1] and is based upon showing that feasible solutions to (4.21) imply the existence of a matrix that satisfies the bounded-real lemma for the closed-loop system. The simplified form given here immediately follows from making use of the fact that $D_{12}$ and $D_{21}$ have full rank and are normalized.

A controller can be constructed by closed formulae from the open-loop plant matrices, $X_\infty$, $Y_\infty$, and $\gamma$, [e.g., Glover & Doyle 1988, Doyle et al. 1989, Gahinet & Apkarian 1994, Gahinet 1996, Zhou et al. 1995, Theorem 16.4, p. 411].

Matlab’s Robust Control Toolbox [Balas et al. 2014] provides the hinfsyn routine to solve the $\mathcal{H}_\infty$ controller synthesis problem. This is done either based on the LMI characterization of Theorem 4.1 and convex optimization or based on a formulation by Glover & Doyle [1988] and Doyle et al. [1989] which uses Riccati equations. In the latter case, an iteration of the value for $\gamma$ through a bisection algorithm is performed until a solution close enough to the optimum is found. The Riccati solution appears to be numerically more benign in most cases and very recent results by Glover & Packard [2017] are likely to further improve reliability. The routine hinfsyn also automatically performs all necessary transformations to achieve the special form (4.19).\(^\text{10}\)

\(^{10}\)The particular structure of the McFarlane-Glover controller can further be exploited to avoid the iteration and immediately obtain a suboptimal controller as implemented in the Matlab routines ncfsyn and loopsyn.
4.3 Synthesis Machinery

4.3.2 LPV Controller Synthesis

A synthesis condition for gridded LPV systems was first published by Wu et al. [1995, 1996] and is detailed by Wu [1995, Cha. 4]. Essentially identical results were also obtained around the same time by Wood [1995, Cha. 8]. The synthesis conditions can be seen as immediate extensions of the $\mathcal{H}_\infty$ controller synthesis conditions of Theorem 4.1.

**Theorem 4.2.** [Wu 1995, Theorem 4.3.2, p. 81] Let $G_\rho$ be an open-loop generalized LPV plant with the special structure (4.19) defined on the domain $T$. There exists an LPV controller

$$K_\rho: \begin{cases} \dot{x} = A(\rho, \dot{\rho}) x + B(\rho) v \\ u = C(\rho) x + D(\rho) v \end{cases} \quad (4.22)$$

that internally stabilizes the closed-loop interconnection $G_{CL, \rho}$ and guarantees $\|G_{CL, \rho}\| < \gamma$ if

$$\gamma > \max_{p \in \mathcal{P}} \max \left( \sigma_{\max}(D_{111}(p)), \sigma_{\max}(D_{111}^T(p)) \right)$$

and if there exist symmetric positive definite matrix functions $X: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x}$ and $Y: \mathcal{P} \mapsto \mathbb{R}^{n_x \times n_x}$ such that for all $(p, q) \in \mathcal{P} \times \mathcal{Q}$

$$\begin{bmatrix} X(p) & \frac{1}{\gamma} I \\ \frac{1}{\gamma} I & Y(p) \end{bmatrix} \prec 0 \quad (4.23a)$$

$$\begin{bmatrix} \Lambda_X(p, q) - B_2(p) B_3^T(p) & X(p) C_{11}^T(p) & \frac{1}{\gamma} (B_1(p) - B_2(p) D_{112}(p)) \\ \ast & -I & \frac{1}{\gamma} D_{111}(p) \\ \ast & \ast & -I \end{bmatrix} \prec 0 \quad (4.23b)$$

$$\begin{bmatrix} \Lambda_Y(p, q) - C_2^T(p) C_2(p) & Y(p) B_{11}(p) & \frac{1}{\gamma} (C_{11}(p) - C_{22}(p) D_{112}(p)) \\ \ast & -I & \frac{1}{\gamma} D_{111}(p) \\ \ast & \ast & -I \end{bmatrix} \prec 0 \quad (4.23c)$$

where $\ast$ denotes symmetric completion and

$$\Lambda_X(p, q) := X(p) (A(p) - B_2(p) C_{12}(p))^T + (A(p) - B_2(p) C_{12}(p)) X(p) - \sum_{i=1}^{n_\mu} \frac{\partial X}{\partial \rho_i} \bigg|_p q_i,$$

$$\Lambda_Y(p, q) := Y(p) (A(p) - B_{12}(p) C_2(p)) + (A(p) - B_{12}(p) C_2(p))^T Y(p) + \sum_{i=1}^{n_\mu} \frac{\partial Y}{\partial \rho_i} \bigg|_p q_i .$$

**Proof.** The proof is provided in extensive length by Wu [1995, Sec. 4.3] and is based upon showing that feasible solutions to the LMIs (4.23) imply the existence of a matrix function that satisfies Theorem 2.2 for the closed-loop system. $\blacksquare$
The parameter rate $q$ appears affinely in the terms $\Lambda_X$ and $\Lambda_Y$, and hence the constraints (4.23) need to be satisfied for the vertices of the set $Q$, i.e., both for minimum and maximum values of the rates at each grid point. In order to arrive at a tractable formulation, the positive definite matrix functions $X : P \mapsto \mathbb{R}^{n_x \times n_x}$ and $Y : P \mapsto \mathbb{R}^{n_x \times n_x}$ must further be formulated in terms of a predefined set of basis functions as

$$
X(\rho) = \sum_{i=1}^{a} f_i(\rho) X_i, \quad X_i \in \mathbb{R}^{n_x \times n_x} \quad \text{and} \quad Y(\rho) = \sum_{i=1}^{b} g_i(\rho) Y_i, \quad Y_i \in \mathbb{R}^{n_x \times n_x}. \quad (4.24)
$$

It is important to emphasize that the conditions used in Theorem 4.2 are only sufficient. The first reason for this is the restriction to a quadratic-in-the-states storage function as a certificate to bound the induced $\mathcal{L}_2$-norm in Theorem 2.2, which forms the basis for Theorem 4.2. In the context of LMIs, this restriction is inevitable. The second reason is related to the conservatism that is introduced by additionally restricting the search space for the parameter-dependent matrices $X(\rho)$ and $Y(\rho)$, which implicitly constitute such a storage function. Thus, selecting more basis functions for $X(\rho)$ and $Y(\rho)$ to enhance the search space usually increases performance. A controller can be constructed by closed formulae from the open-loop plant matrices and the feasible values of $X$, $Y$, and $\gamma$ as described by Wu [1995, p. 82] or Lee [1997, Theorem 4.2.5, pp. 47]. The relationship to the conditions of Theorem 4.1 is established for LTI systems and constant matrices $X = 1/\gamma X_\infty$, $Y = 1/\gamma Y_\infty$ through multiplication of all inequalities (4.23) by $\gamma$.

The freely available LPVTools toolbox [Balas et al. 2015, Hjartarson et al. 2015] for Matlab provides the routine `lpvsyn` that implements the LPV synthesis conditions as a convex optimization problem to minimize $\gamma$. Further, basis functions as stated in Equation (4.24) can be conveniently defined and an interface with many Robust Control toolbox and Control Systems toolbox functions is provided. The routine also performs all necessary transformations to achieve the special form (4.19).\(^1\)

### 4.3.3 Suboptimal Synthesis and Implementation

Both in $\mathcal{H}_\infty$ and LPV control, suboptimal controllers are practically more relevant than actual “optimal” ones. Particular insight into the problem is again obtained by the McFarlane-Glover method that permits an analytical solution which, however, turns out to be singular [McFarlane & Glover 1992]. The same limiting behavior can be observed for most synthesis conditions: Being a convex optimization problem, it is not surprising that the minimum often is located on the boundary of the constraints. Hence, nearly singular matrices $X$ and $Y$ are often obtained and as the controller reconstruction usually involves inverses of these matrices, numerical problems are commonly encountered.

Further, the optimal solution with its flat frequency response characteristic is not necessarily desirable for control problems [cf. Zhou et al. 1995, Sec. 16.9]. Specifically, nearly optimal solutions often result in unnecessarily fast controller dynamics, which can severely complicate implementation. Several remedies have been proposed, e.g., to

\(^1\)Again, the special structure of the McFarlane-Glover design problem permits a simplified solution of the LPV controller synthesis problem as described by Wood [1995, Sec. 8.4–8.5].
incorporate additional (pole region) constraints [Lee 1997, Sec. 4.2.3]. The so far most successful approach, however, appears to be a relaxation of the problem after the achievable performance index has been determined. In a second step, a new synthesis problem is solved with a fixed suboptimal performance index that is 5–20% above the optimal \( \gamma \). This often eliminates fast controller dynamics, while the controllers’ frequency response is usually indistinguishable in the relevant frequency range and almost identical results in time domain are achieved. It is further possible to use this second step to explicitly improve conditioning of the matrices \( X \) and \( Y \) by solving additional optimization problems [e.g., Saule 2013, Sec. 3.4]. A very similar approach is also implemented in **lpvsyn** [Balas et al. 2015].

### 4.3.4 Discretization

Naturally, controllers need to be implemented on a digital computer, i.e., in discrete time. It is in principle possible to use synthesis techniques that directly yield discrete time controllers [e.g., Packard 1994, Apkarian & Gahinet 1995]. The question of whether the design should be carried out in discrete time rather than continuous time therefore naturally arises. A discrete time design would automatically incorporate the delay due to the zero-order hold operation. This delay would, however, usually only be a small part of the overall delay. The remaining part, caused by sensor and actuator components as well as computation units, would still require a model. While the advantages of a discrete-time design thus seem to be limited, insight into the problem would be lost to a certain extent as engineering intuition is commonly higher developed in the continuous-time domain [cf. Apkarian et al. 1996, Sec. 12]. It therefore appears preferable to design a controller in continuous time and then to discretize the resulting controller using any available discretization scheme as discussed in detail by Apkarian et al. [1996, Sec. 12]. In particular, the standard Tustin discretization, also known as trapezoidal discretization or bilinear transformation, is applicable to LPV systems and can be formulated as [ibid., Theorem. 12.2.1, p. 205]

\[
\begin{align*}
    x_{k+1} &= \left( I - \frac{T}{2} \rho k \right)^{-1} \left( I + \frac{T}{2} \rho k \right) x_k + \sqrt{T} \left( I - \frac{T}{2} \rho k \right)^{-1} B_{\rho k} y_k \\
    u_k &= \sqrt{T} C_{\rho k} \left( I - \frac{T}{2} \rho k \right)^{-1} x_k + \left( \frac{T}{2} C_{\rho k} \left( I - \frac{T}{2} \rho k \right)^{-1} B_{\rho k} + D_{\rho k} \right) y_k.
\end{align*}
\]

(4.25)

In Equation (4.25), \( k \) refers to the time step and \( T \) is the sampling time which has to be chosen sufficiently small. Thus, \( x_k \) refers to \( x(t) \big|_{t=kT} \). This form is easy to implement in a lookup table representation, completely analog to the continuous-time state space representation.
4.4 Robustness Margins

Control systems are always subject to uncertainty and the potentially destabilizing effects of feedback control necessitate sufficiently large robustness margins. This section briefly reviews some important concepts to evaluate the robustness of closed-loop systems. It must be highlighted that these concepts rely on linear systems theory and thus have no rigorous theoretical justification for LPV systems. This gap in the literature has been addressed only very recently using integral quadratic constraints, e.g., by Pfifer & Seiler [2015, 2016] and Schug et al. [2017]. As these methods cannot be considered to be mature yet, this section focuses on linear tools that have gained wide acceptance in practice.

The most common metric to quantify robustness for a linear control system is given by the classical gain and phase margins (GM and PM). The former specifies how much gain variation a single loop transfer function can tolerate before instability occurs. The second measures the amount of phase loss that this loop can tolerate. Both margins are independent of each other and common practice in aerospace usually requires at least 6 dB gain margin and 45° phase margin. These numbers are, however, derived from experience with rigid aircraft and certain highly structured control architectures. They are thus not necessarily meaningful for a particular problem.

4.4.1 Disk Margins

In the complex plane, the gain margin measures the distance from the critical point to the Nyquist plot along the real axis. Similarly, the phase margin measures the distance along the unit circle. Considering only these classical margins can therefore easily overlook destabilizing combinations of gain and phase that independently are considered safe. A popular example can be found in Zhou et al. [1995, Sec. 9.6]. It is therefore important to take into account simultaneous gain and phase variations. In aerospace, the “Nichols diamond” is a common tool to achieve this by considering the sum of the distances from the critical point measured in gain and phase direction (Figure 4.12a). A very similar metric is to take the minimum Euclidean distance to the critical point in the complex plane, corresponding to the peak value \( \| S \| \) of the sensitivity function \( S \) (Figure 4.12b).

The sensitivity criterion implies gain and phase margins [Balas & Seiler 2014]

\[
\begin{align*}
\text{GM}_{\text{lower}} &= \frac{\| S \|}{\| S \| + 1}, \quad \text{GM}_{\text{upper}} = \frac{\| S \|}{\| S \| - 1}, \quad \text{PM} = \pm 2 \arcsin \left( \frac{1}{2\| S \|} \right).
\end{align*}
\]

As \( \| S \| \geq 1 \), the largest margins that can be deduced from the sensitivity criterion are \( \text{GM}_{\text{lower}} = -6 \text{ dB} \), \( \text{GM}_{\text{upper}} = \infty \text{ dB} \), and \( \text{PM} = \pm 60^\circ \). Similarly, a bound on the complementary sensitivity \( T \) can be used to derive bounds [ibid.]

\[
\begin{align*}
\text{GM}_{\text{lower}} &= 1 - \frac{1}{\| T \|}, \quad \text{GM}_{\text{upper}} = 1 + \frac{1}{\| T \|}, \quad \text{PM} = \pm 2 \arcsin \left( \frac{1}{2\| T \|} \right).
\end{align*}
\]

As \( \| T \| \geq 1 \), the maximum guaranteeable margins from this criterion are \( \text{GM}_{\text{lower}} = -\infty \text{ dB} \), \( \text{GM}_{\text{upper}} = 6 \text{ dB} \), and \( \text{PM} = \pm 60^\circ \). These limitations of the individual sensitivity functions motivated the introduction of a symmetric disk margin by Blight et al. [1994].
4.4 Robustness Margins

\[ |L(j\omega)| (\text{dB}) \]
\[ \angle (L(j\omega)) (\degree) \]
\[ \text{Re}(L(j\omega)) \]
\[ \text{Im}(L(j\omega)) \]

(a) Nichols plot (—) with “diamond” (——).

(b) Nyquist plot (—) with sensitivity bound (——).

(c) Nyquist plot (—) with disk margin (——).

\[ (4.28) \]

\[ \mu_{\Delta}(M) = \begin{cases} 
\min_{\Delta \in \Delta} \left( \frac{1}{\sigma_{\max}(\Delta)} \right) & \text{if } I - M \Delta \text{ is singular for some } \Delta \in \Delta \\
0 & \text{otherwise.} 
\end{cases} \]

Figure 4.12: Examples of simultaneous gain and phase margins. The independent gain and phase margins are shown in gray.

Definition 4.1 (SISO Disk Margins). The disk margins for a SISO system are defined as

\[ \text{DGM}_{\text{lower}} = \frac{1}{1 + \frac{1}{||S-T||}} \quad \text{DGM}_{\text{upper}} = \frac{1}{1 - \frac{1}{||S-T||}} \quad \text{PM} = \pm 2 \arctan \left( \frac{1}{||S-T||} \right) \]

As \( ||S-T|| \geq 1 \), the disk margins allow the certification of (symmetric) margins up to \( \text{GM} = \pm \infty \) dB and \( \text{PM} = \pm 90^\circ \). They can also be given the interpretation of a circle in the complex plane, whose center is, however, not at the critical point but at \( -\frac{1}{||S-T||^2 + 1} \) \( ||S-T||^2 + 1 \) (Figure 4.12c).

The extension to multivariable control loops makes use of the structured singular value, first introduced by Doyle [1982]. Detailed descriptions of this powerful analysis technique can be found in the tutorial paper by Packard & Doyle [1993] and many robust control text books, e.g., Zhou et al. [1995, Cha. 8], Sanchez-Peña & Sznaier [1998, Cha. 7], and Skogestad & Postlethwaite [2005, Sec. 8.8–8.11].

Definition 4.2 (Structured Singular Value [Packard & Doyle 1993]). Given a complex matrix \( M \) and a block structure \( \Delta \), the structured singular value is defined as

\[ (4.28) \]

\[ \mu_{\Delta}(M) = \begin{cases} 
\min_{\Delta \in \Delta} \left( \sigma_{\max}(\Delta) \right) & \text{if } I - M \Delta \text{ is singular for some } \Delta \in \Delta \\
0 & \text{otherwise.} 
\end{cases} \]
The structured singular value depends both on the given matrix and the structure of the uncertainty. It measures the size of the smallest uncertainty within that structure that causes $I - M \Delta$ to lose rank. In conjunction with the multivariable Nyquist criterion [e.g. Skogestad & Postlethwaite 2005, Theorem 8.1, p. 301, originally formulated by MacFarlane 1970], this can be used to verify stability of an uncertain dynamic system.

**Definition 4.3 (MIMO Disk Margins).** Let $(M, \Delta)$ denote the uncertain interconnection shown in Figure 4.13. That is, $M$ is a known LTI system that includes $C$ and $P$ and establishes the interconnection structure of Figure 4.13a and $\Delta := \text{diag}(\delta_1, \ldots, \delta_{ny+nu})$ is a diagonal uncertainty that collects the individual scalar uncertainties $\delta_i$, $i = 1, \ldots, ny$ acting on all output channels and $\delta_i$, $i = ny+1, \ldots, ny+nu$ acting on all input channels.

![Figure 4.13: Equivalent interconnections for MIMO disk margin calculation.](image)

Denote

$$\frac{1}{r} := \inf_{\omega} \mu_\Delta(M(j\omega)).$$

The multi-input-multi-output disk margin for a MIMO system is then defined as

$$DGM_{\text{lower}} = \frac{1 - r}{1 + r}, \quad DGM_{\text{upper}} = \frac{1 + r}{1 - r}, \quad PM = \pm 2 \arctan (r).$$

Common other uncertainty structures consider only simultaneous individual perturbations at all outputs (multi-output disk margin) or all inputs (multi-input disk margin), i.e., either $\delta_i$, $i = ny+1, \ldots, ny+nu$ or $\delta_i$, $i = 1, \ldots, ny$ are fixed to zero in $\Delta$.

It is important to emphasize that such a multivariable analysis takes into account directionality effects that are not addressable with single-loop analysis techniques. The degradation of multi-loop margins relative to single-loop margins can therefore be a useful indicator of potential issues related to directionality. It is in principle also possible to use specific perturbations, e.g., a perturbation that acts identically on several outputs. This kind of analysis is, however, conceptually closer to general $\mu$-techniques than to the calculation of generic margins.
4.4 Robustness Margins

4.4.2 Robustness in Mixed Sensitivity Design

This chapter closes with a brief discussion of the inherent robustness properties of the mixed sensitivity design which was proposed in Section 4.2.3. As robustness is solely determined by feedback, it suffices to consider the single-degree-of-freedom interconnection of Figure 4.8a which is repeated here as Figure 4.14 for convenience.

![Figure 4.14: Generalized closed-loop for single-degree-of-freedom mixed sensitivity design.](image)

As the generalized closed-loop interconnection is guaranteed to satisfy a norm bound $\|G_{CL}\| < \gamma$, the relation to robustness analysis is immediate: Uncertainties can be included to close the fictitious loops with

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}. \quad (4.29)$$

According to the small-gain theorem [e.g., Zhou et al. 1995, Theorem 9.1, p. 212, originally formulated by Desoer & Vidyasagar 1975, Sec. III.2, pp. 40], the interconnection is guaranteed to be stable as long as $\gamma \|\Delta\| < 1$. Figure 4.15 shows the corresponding generalized closed-loop interconnection, resulting in an “allowable” uncertain plant $P_{\Delta}$. In particular, the original plant model can be perturbed with an additive uncertainty $(P_d D_d \Delta_{22} + D_e \Delta_{12}) W_u D_c^{-1}$. Such uncertainty is representative of unmodeled high frequency dynamics, as well as uncertain non-minimum phase zeros, including a change in the number of non-minimum phase zeros [Zhou et al. 1995, Sec 9.3.4]. Similarly, the allowable inverse multiplicative uncertainty $(P_d D_d \Delta_{21} + D_e \Delta_{11}) W_c D_e^{-1}$ is representative of low-frequency parameter errors as well as uncertain unstable poles including a change in the number of unstable poles [ibid.]. While this demonstrates “general robustness”, not much insight is gained due to the complexity of the involved expressions.

A more insightful interpretation can be provided if instead load disturbances are considered ($P_d = P$ and $D_d = D_u$). In this case, the “allowable uncertainty” shown in Figure 4.16 is obtained. The uncertainty $\Delta$ is allowed to be fully populated, so any structured uncertainty is included. Fixing $\Delta_{11} = \text{diag}(\delta_1, \ldots, \delta_{n_y})$ and $\Delta_{22} = \text{diag}(\delta_{n_y+1}, \ldots, \delta_{n_y+n_u})$ to diagonal uncertainties and $\Delta_{21} = \Delta_{12} = 0$ as is done in the margin tests which were previously discussed is hence permissible. That is, the loop is also guaranteed to be stable.
Figure 4.15: Stable closed-loop with mixed-sensitivity controller and unstructured uncertainty $\Delta = \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix}$ with $\|\Delta\| < 1/\gamma$.

Figure 4.16: Stable closed-loop with mixed-sensitivity controller, $P_d = P$, $D_d = D_u$, and unstructured uncertainty $\Delta = \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix}$ with $\|\Delta\| < 1/\gamma$.

for

$$\gamma \left\| \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_{n_y+n_u} \end{bmatrix} \right\| < 1 \Leftrightarrow \gamma \|\delta_i\| < 1 \quad i = 1, \ldots, n_y+n_u \quad (4.30)$$

In this case, owing to their diagonal structure, $D_u, W_u, D_e$, and $W_e$ all commute with the uncertainties such that the scalings $D_u$ and $D_e$ drop out, i.e.,

$$D_u \Delta_{22} W_u D_u^{-1} = \text{diag}(\delta_{n_y+1}, \ldots, \delta_{n_y+n_u}) W_u,$$

$$D_e \Delta_{11} W_e D_e^{-1} = \text{diag}(\delta_1, \ldots, \delta_{n_u}) W_e.$$

Assume the standard weights $W_u$ and $W_e$ are used as proposed in Section 4.2.3, i.e., $W_u$ has, in each channel, at least unit gain and $W_e$ has, in each channel, at least a gain of 0.5. Then, the closed-loop (shown in Figure 4.17) can tolerate multiplicative perturbations with a size of $1/\gamma$ in each input channel, and inverse multiplicative perturbations of size $1/2\gamma$ in
4.4 Robustness Margins

each output channel. It should be noted that $\gamma > 1$ in this case, as $\gamma \geq \|G_{CL}\| > \|T\| \geq 1$. Using Equations (4.26) and (4.27), the closed-loop system has, for each channel, guaranteed margins

$$GM_{\text{output lower}} = \frac{2\gamma}{2\gamma + 1}, \quad GM_{\text{output upper}} = \frac{2\gamma}{2\gamma - 1}, \quad PM_{\text{output}} = \pm 2 \arcsin \left( \frac{1}{4\gamma} \right)$$ (4.31)

$$GM_{\text{input lower}} = 1 - \frac{1}{\gamma}, \quad GM_{\text{input upper}} = 1 + \frac{1}{\gamma}, \quad PM_{\text{input}} = \pm 2 \arcsin \left( \frac{1}{2\gamma} \right).$$ (4.32)

For example, a performance index $\gamma = 2$ guarantees an upper output gain margin of 2.5 dB, an upper input gain margin of 3.5 dB, and phase margins of $14^\circ$ at the output and $29^\circ$ at the input. These bounds are conservative, such that dedicated robustness tests as discussed in the previous section can provide better results and should always be carried out after a controller was designed. The existence of guaranteed margins nevertheless demonstrates the inherent robustness of the proposed mixed sensitivity design and is in stark contrast, e.g., to LQG control which has no robustness guarantees [Doyle 1978].

Figure 4.17: Stable closed loop with mixed-sensitivity controller, $P_d = P$, $D_d = D_u$, and scalar uncertainties $\delta_i$ with $\|\delta_i\| < 1 \quad i = 1, \ldots, n_y + n_u$.

It should further be noted that the guaranteed input margins are related to the assumption of disturbances occurring at the plant input [cf. Doyle & Stein 1979]. That is, they can only be guaranteed when the disturbance model equals the plant model. For a general disturbance model, it is nevertheless possible to derive guaranteed input margins by shifting the uncertainty to the plant input [e.g. Skogestad & Postlethwaite 2005, p. 295]. Considering again $\Delta_{21} = \Delta_{12} = 0$, structured diagonal uncertainties $\Delta_{11}, \Delta_{22}$, and (for simplicity) $D_d = D_u$, the additive uncertainty

$$P_d D_d \Delta_{22} W_u D_u^{-1} = P_d \text{diag}(\delta_{n_y+1}, \ldots, \delta_{n_y+n_u}) W_u$$

can be expressed as a multiplicative input uncertainty

$$P_d \text{diag}(\delta_{n_y+1}, \ldots, \delta_{n_y+n_u}) W_u.$$

Thus, at every frequency, an additional factor $1/\sigma_{\text{max}}(P_d^TP_d)$ appears which reduces the guaranteed admissible uncertainty. The guarantees of the previously discussed case are, however, asymptotically recovered as $P_d$ is selected to resemble $P$. Thus, the disturbance model should include dominant high-gain dynamics of $P$ such as resonances in order to achieve robustness.


5 Application Examples

This chapter provides detailed application examples for the techniques which were developed in this thesis. First, the applicability and effectiveness of the proposed model order reduction methods of Chapter 3 are demonstrated on two different high-fidelity industry-grade models of unmanned aeroservoelastic aircraft. Further, a large-scale aerodynamics model with over 20000 states is considered, both as a proof-of-concept and to demonstrate the versatility of the proposed oblique projection technique. Finally, two comprehensive design studies for aeroservoelastic control applications are provided. The first study is concerned with active flutter suppression on a small unmanned aircraft and envelope expansion into the naturally unstable flight regime. The second study introduces a flight control design concept for a large aircraft that integrates structural damping augmentation. These design studies substantiate the systematic design guidelines of Chapter 4. Parts of this chapter were published as research papers by Theis, Takarics, Pfifer, Balas & Werner [2015c], Theis, Seiler & Werner [2018], Theis, Pfifer & Seiler [2016a], and Theis, Pfifer, Balas & Werner [2015a].

5.1 Model Order Reduction for Aeroservoelastic Aircraft

Design of flight control systems for aeroservoelastic unmanned aircraft requires accurate, yet low-order models. The modal interpolation method is demonstrated on a high-fidelity longitudinal dynamics model of the Body Freedom Flutter Vehicle. The Body Freedom Flutter Vehicle is a high aspect ratio flying wing with a span of 3m that was deliberately designed by Lockheed Martin and the U.S. Air Force Research Laboratory to flutter at very low airspeed [Burnett et al. 2010, Holm-Hansen et al. 2010]. Lockheed Martin and the Air Force Research Laboratory also developed a larger flying wing aircraft with a span of 8.5m, designated the X56A multi utility technology testbed (MUTT). The X56A is currently used as a research platform by NASA [Ryan et al. 2014]. The oblique projection method is demonstrated on a model of this aircraft. For both of these MIMO systems, the LMI-based state-of-the-art methods fail to produce a reduced-order model.

5.1.1 Body Freedom Flutter Vehicle

The Body Freedom Flutter Vehicle was developed by the U.S. Air Force Research Laboratory and Lockheed Martin Aeronautics Company as a test platform for the investigation of aeroservoelastic effects and demonstration of active aeroservoelastic control [Holm-Hansen et al. 2010]. Five aircraft were built of which four were destroyed during flight testing.
The last remaining Body Freedom Flutter Vehicle, shown in Figure 5.1, has been donated to the University of Minnesota, Minneapolis.

![Body Freedom Flutter Vehicle](image)

**Figure 5.1:** Body Freedom Flutter Vehicle.

The aircraft is a flying wing with a mass of approximately 6 kg and a wing span of 3 m. Its wings are slit to decrease structural stiffness. As a result, flutter occurs at very low airspeed. Due to the absence of a horizontal stabilizer, this flutter is a body freedom flutter that involves wing flapping and torsion in conjunction with a pitch motion of the aircraft. The aircraft has eight control surfaces, two at the trailing edge of the body and six at the trailing edge of the wings. It is further equipped with a variety of sensors. Among these sensors are six accelerometers, two located at each wing tip and two in the center body, fore and aft, respectively. The aircraft is also equipped with a Pitot tube to measure airspeed and with an inertial measurement unit (IMU) to measure rates of rotation and attitude.

Lockheed Martin also developed a mathematical model of the aircraft [Burnett et al. 2010]. That model was originally constructed in MSC/NASTRAN using a stiffness model with 376 degrees of freedom to define the structural dynamics and a DLM unsteady aerodynamics model with 2252 degrees of freedom. A set of 21 linear models parameterized by equivalent airspeed from 40 to 80 knots in steps of 2 knots was then generated using a rational function approximation and assumed shapes for structural deformation [ibid.]. The models have a total of 148 state variables and represent straight level flight at 1000 ft altitude. The first 37 states represent generalized displacements related to 5 rigid-body modes (lateral, plunge, roll, pitch and yaw), 8 flexible modes (symmetric and anti-symmetric bending and torsion) and 24 degrees of freedom associated with local vibrations of the control surfaces. The second 37 states are the corresponding generalized velocities and the remaining 74 states are associated with aerodynamic lags from the rational function approximation of the unsteady aerodynamics model. A thorough description of the aircraft and its mathematical model is provided by Moreno [2015, Cha. 3]. The 21 linear models depend continuously on the airspeed and are hence treated as a single LPV system. No
nonlinear model is available such that the effects of rate variation on the dynamics are unknown in the context of the present work and can only be assumed to be negligible. A reduced-order model can hence only be constructed from local information represented by the individual linear models.

The objective of the model order reduction problem is to obtain a low-order model suitable for the design of active aeroservoelastic control. Figure 5.2 shows the natural frequency and damping ratio of the short period dynamics and the elastic modes as a function of airspeed. Symmetric wing bending becomes unstable beyond 42 knots, followed by symmetric wing torsion beyond 60 knots and anti-symmetric wing torsion beyond 62 knots. Capturing these modes accurately in the reduced-order model is vital for a successful control design.

While there are a total of 13 sensors and 8 control surfaces on the vehicle, the actuator and sensor selection study by Moreno [2015, Sec. 5.3.1–5.3.2] suggests that a subset of inputs and outputs is almost as effective. As feedback measurements, the six accelerometers located by twos fore and aft at the left wing, center body, and right wing are used. The two body flaps and the two outboard flaps are the control inputs. This reduced configuration is shown in Figure 5.3. It is not only easier to handle due to its reduced size but also leaves the remaining control surfaces to provide enough control authority for the primary flight control system.

The actuators for all flaps have a bandwidth of about 125 rad/s. Beyond this frequency, control becomes impossible. Figure 5.2 shows that the unstable modes (first symmetric wing bending, first symmetric wing torsion, and first anti-symmetric wing torsion) all have frequencies between 25 rad/s and 65 rad/s. Thus, the actuators are a factor of two faster than the fastest unstable mode. This gap is usually considered sufficient, but it is nevertheless very small and necessitates high accuracy of the reduced-order model up to the actuator bandwidth. The roll-off of the actuators beyond this frequency then gradually ensures enough gain attenuation to tolerate larger modeling errors. The low frequency range is, on the other hand, less important. A flutter suppression controller is typically designed with
bandpass characteristic to limit interaction with rigid-body motion (cf. Section 5.3.2) and hence is not active at low frequencies. From these considerations, the frequency range of interest is defined as 10–125 rad/s.

Due to the number of state variables, standard LMI-based techniques are computationally intractable on the full-order model. The problem is further particularly challenging due to the three stable-unstable bifurcations that occur with increasing airspeed. Due to instability, Gramian-based methods are difficult to apply and would require a coprime factorization. This stability issue as well as the presence of a large number of modes outside the frequency range of interest would make modal decomposition the method of choice for LTI model order reduction. To make use of the benefits of this approach but still obtain a continuous LPV model, the modal interpolation method of Section 3.2 is used.

**Modal Interpolation Reduction**

The first part of the model order reduction procedure consists of locally reducing the LPV model. The following consecutive steps are performed:

1. **Modal Transformation**: In a first step, the LTI models at the grid points are individually transformed into modal form. This also permits a stable-unstable decomposition for each LTI model, such that the unstable modes are exactly preserved.
   a) **Modal Truncation**: All modes below the frequency range of interest (i.e., 10 rad/s) are removed by truncation. This eliminates rigid-body motion and slow dynamics that are irrelevant for aeroservoelastic control. This step removes 7 state variables and leads to models with 141 state variables each.
   b) **Modal Residualization**: All modes which cannot be effectively controlled since they are outside of the control bandwidth are removed. It is important to take into account the required frequency range for the controller to roll-off, i.e., the model should still be accurate slightly above the actuator bandwidth. Hence, all modes above 200 rad/s are residualized. This step removes 50 state variables and thus leads to models with 91 state variables each.
2. **Balanced Truncation:** The stable parts of each local model are individually transformed into a frequency-weighted balanced realization and state variables are truncated based on their Hankel singular values. This step removes all state variables with negligible contribution to the input/output behavior of the system in the frequency range of interest. After the truncation, the unstable parts are added back to form the reduced-order models. The Robust Control Toolbox’ square root balanced truncation algorithm `balancmr` [Safonov & Chiang 1989] is employed with the frequency weighting shown in Figure 5.4a that defines the region of interest as $10\, \text{rad/s}$ to $200\, \text{rad/s}$. This frequency weighting is necessary to minimize the effect of the errors which where introduced by the modal truncation and residualization in the previous step on the balancing procedure. As a criterion to determine the number of state variables which are retained, a $\nu$-gap of less than 0.2 for all frequencies below $125\, \text{rad/s}$ and less than 0.3 at $200\, \text{rad/s}$ is selected. Maintaining 15 states in all local models leads to the $\nu$-gap plotted in Figure 5.4b.

![Frequency weight for local balanced truncation.](image1)

**Figure 5.4:** Local model reduction.

The result of this procedure is a collection of 21 reduced-order models $(A_{k,\text{red}}, B_{k,\text{red}}, C_{k,\text{red}}, D_{k,\text{red}})$ that approximate the input-output behavior of the original LPV system at each grid point and exactly preserve all unstable modes. The local reduced-order models are then transformed into the mode-wise canonical form and the matching algorithm of Section 3.2.2 is applied to construct local realizations that allow the interpolation of the local models to be performed in a consistent state space basis.
Evaluation of the Reduced Order Model

The interpolated LPV model exactly reproduces the reduced LTI models at the grid points. Potential problems are therefore expected to arise between grid points. Following the standard approach for gridded LPV systems, piece-wise linear interpolation is used to recover models between grid points. Figure 5.5 shows the frequency response from left outboard flap to right wing aft acceleration measurement, parameterized over airspeed, for the full-order and the reduced-order model. Both plots agree qualitatively and are indeed continuous in their dependence on airspeed. The two prominent peaks correspond to the stable-unstable bifurcations occurring at 60 and 62 knots airspeed at a frequency of approximately $65 \text{ rad/s}$. The first stable-unstable bifurcation at 42 knots and a frequency of $25 \text{ rad/s}$ is also visible as a spike.

(a) Full-order model (148 state variables). (b) Modal interpolation (15 state variables).

Figure 5.5: Parameterized frequency response from left outboard flap to right wing aft acceleration.

To further verify that the stable-unstable transitions are accurately captured, the VFG plot of Figure 5.2 is reproduced for the reduced-order model. The results are shown in Figure 5.6. The short period frequency can be seen to differ marginally from about 54 knots onwards. The frequency of the first anti-symmetric wing bending mode is captured very well, but the damping starts to differ from about 70 knots onwards. All other modes are very precisely captured in the reduced-order model. Most importantly, the stable-unstable bifurcations occur at the exact same frequency and airspeed.

Another way of representing this result is shown in Figure 5.7, where the pole migration within the frequency range of interest is compared for the full-order model and the reduced-order model. In this plot, the smoothness of the interpolation is confirmed. Further, the evolution of the modes that become unstable can be seen to be captured very accurately.
5.1 Model Order Reduction for Aeroservoelastic Aircraft

Figure 5.6: Velocity/frequency/damping (VFG) plot of reduced-order model. Short period (•), 1st symmetric wing bending (•), 1st anti-symmetric wing bending (□), 1st symmetric wing torsion (△), 1st anti-symmetric wing torsion (▲), 2nd anti-symmetric wing bending (♦) with values for full-order model shown in gray.

(a) Full-order model (148 state variables). (b) Modal interpolation (15 state variables).

Figure 5.7: Pole migration across the flight envelope.

Comparison to LMI-based Model Reduction

To compare the proposed reduction method with the state-of-the-art technique, a second reduced-order LPV model of the Body Freedom Flutter Vehicle is generated following the procedure proposed by Moreno et al. [2014]. That is, an LMI-based coprime factor balanced truncation as described in Section 3.1.2 is performed after a heuristic pre-processing which is necessary to arrive at a tractable problem size. In a first step, all state variables that result in a \( \nu \)-gap error of less than 0.001 when residualized in the original state space coordinates are residualized. Second, the state variables that result in a \( \nu \)-gap error of less than 0.2 when truncated are truncated. Next, an approximate modal transformation is found by nonlinear optimization to minimize the \( \nu \)-gap metric introduced by truncating all modes below 10 rad/s and residualizing all modes above 200 rad/s. This pre-processing results in a model with 42 states, which is then coprime factorized and balanced. The
solution of the required LMIs to calculate this transformation takes about 3 hours on a
desktop PC. Using again a $\nu$-gap of less than 0.2 below $125\,\text{rad/s}$ and 0.3 at $200\,\text{rad/s}$ as a
criterion for the remaining number of state variables eventually leads to a reduced-order
LPV model with 26 state variables.

Figure 5.8 shows the $\nu$-gap of the LMI-based reduced-order model and the reduced-order
model from modal interpolation with respect to the full-order model evaluated over a
dense grid of airs speeds with steps of 0.1 knots. The overall accuracy of both reduced-order
models is very good within the frequency range of interest. The $\nu$-gap values are well below
0.2 at frequencies up to $160\,\text{rad/s}$ and equal 0.3 at $200\,\text{rad/s}$ for both cases. The LMI-based
model reduction leads to a smaller error in the low-frequency regime at high airspeed, but
to a larger error at low airspeed when compared to the interpolated model. Within the
critical frequency range of about $40–80\,\text{rad/s}$, where the unstable modes are located, the
interpolated model is more accurate than the LPV model. The $\nu$-gap varies smoothly across
the airspeed dimension for both reduced-order models. With the LMI-based approach
explicitly accounting for the parameter dependence, this is to be expected. The proposed
interpolation method, however, yields similarly smooth results between grid points.

![Graphs showing the $\nu$-gap variation](image)

(a) LMI-based model reduction (26 state
variables). (b) Modal interpolation (15 state variables).

**Figure 5.8:** Variation of the $\nu$-gap between the reduced-order models and the full-order model
(148 state variables).

Figure 5.9 depicts several Bode plots to provide a direct comparison between the two
reduced-order models and the full-order model. The frequency responses are obtained for
airspeeds not included in the original grid (43 knots and 77 knots). Hence, they provide a
meaningful test for the quality of the interpolation. Both reduced-order models capture
the input-output behavior of the full-order model very well in the frequency range of
interest for all considered input-output maps. The proposed reduced-order model with 15
state variables can therefore be concluded to achieve the same level of accuracy as the
26th-order model obtained by the state-of-the-art LMI-based reduction method.
5.1 Model Order Reduction for Aeroservoelastic Aircraft

Figure 5.9: Frequency response of full-order model (148 state variables) at 43 knots (—) and 77 knots (—) compared to the proposed reduced-order model with 15 state variables (—/—) and LMI-based reduced-order model with 26 state variables (—/—/—).
5 Application Examples

5.1.2 X56A MUTT

The X56A MUTT is another unmanned test platform developed by the U.S. Air Force Research Laboratory and Lockheed Martin Aeronautics Company for the investigation of aeroservoelastic effects and demonstration of active aeroservoelastic control [Beranek et al. 2010]. It is currently used by NASA for research and flight testing [Ryan et al. 2014]. Figure 5.10 shows the aircraft.

![X56A MUTT aircraft](https://commons.wikimedia.org/wiki/File:Lockheed_Martin_X-56A_first_landing.jpg)

The aircraft is a flying wing with a wing span of 8.4 m and built in a modular fashion such that different sets of wings can be attached. In the considered configuration, the aircraft has ten control surfaces, two at the trailing edge of the body and eight at the trailing edge of the wings. Its empty weight is 86 kg with about 25 kg maximum fuel weight. Similar to the Body Freedom Flutter Vehicle, the X56A MUTT features six accelerometers, two located at each wing tip and two in the center body, fore and aft, respectively. The aircraft is also equipped with a Pitot tube to measure airspeed and with an IMU to measure rates of rotation and attitude.

A high-fidelity model of the longitudinal dynamics was developed by Schulze et al. [2016]. It combines rigid-body flight dynamics from first principle modeling, structural dynamics from finite element modeling and unsteady aerodynamics from computational fluid dynamics modeling. The rigid-body states are described in the moving body frame and are represented by angle of attack $\alpha$ and pitch rate $q$. The flexible modal displacements are represented in terms of assumed mode shapes and generalized coordinates $\eta$. Unsteady aerodynamic states are represented by state variables $w$ and are related to the rigid and flexible degrees of freedom of the system. Specifically, every degree of freedom is coupled to a third-order system that describes the unsteady aerodynamic forces caused by, and acting on, modal displacement. There are 8 structural modes (16 states), the 2 rigid-body states and 30 aerodynamic states, which totals to 48 state variables.

Aiming at a model suitable for longitudinal attitude control and flutter suppression, only a subset of the actual available sensors and actuators is considered for the present model order reduction. As inputs, symmetric deflection of the two outboard wing flap pairs ($\delta_1$ and $\delta_2$), highlighted in Figure 5.11, are used. The outputs are a pitch rate measurement $q_{\text{meas}}$ and an acceleration signal obtained at the center body ($a_{z,\text{center}}$), as well as a wing

1NASA public domain image [https://commons.wikimedia.org/wiki/File:Lockheed_Martin_X-56A_first_landing.jpg]
tip acceleration signal \((a_{z,\text{wing}})\) that averages the measurements from the four sensors shown in Figure 5.11.

The dynamics of the aircraft depend parametrically on the airspeed \(V_{\infty}\) and hence the state space model is of the form

\[
\dot{x} = A(V_{\infty}) \, x + B(V_{\infty}) \, \delta \\
y = C(V_{\infty}) \, x + D(V_{\infty}) \, \delta,
\]

with state vector \(x = [w^T \mid \alpha \, q \mid \dot{\alpha}^T \mid \eta^T]^T\), output vector \(y = [q_{\text{meas}} \, a_{z,\text{center}} \, a_{z,\text{wing}}]^T\), and input vector \(\delta = [\delta_1 \, \delta_2]^T\). A grid representation with 12 uniformly spaced points is used to cover the domain \(V_{\infty} \in [30.6 \, 68] \, m/s\). The aircraft is naturally stable in this domain but the damping ratio of the lowest-frequency aeroelastic mode decreases dramatically with higher airspeeds. Hence, the dynamics change rapidly.

The available bandwidth of the control system again provides an upper frequency limit on the fidelity requirement and fifth-order Butterworth filters with a cut-off frequency of 100 \(\text{rad/s}\) are selected. The augmented Lyapunov equations (3.29) are solved using the Matlab routine \texttt{lyapchol}\footnote{The Matlab routine \texttt{lyapchol} at each grid point and a reduced-order model is calculated using Algorithm 2 (p. 50). The calculation takes only seconds and hence the order of the reduced model can be determined by trial and error. A 12th-order model yields satisfactory results.}

Figure 5.12 shows the step response of both the full-order and the reduced-order model along a time-varying parameter trajectory. The trajectory covers the complete parameter range from 31 to 68 \(m/s\) airspeed. The rate of variation in the simulation is further selected to be unreasonably high with a maximum of 40 \(m/s^2\). The reduced-order model nevertheless approximates the response extremely well in all outputs. This confirms that there are absolutely no rate-dependent errors introduced by the oblique projection method. The most prominent difference in the outputs are the high-frequency transients occurring immediately after the step input is applied. These high-frequency dynamics are not included in the reduced-order model as a consequence of the frequency-weighted approximation.

Figure 5.13 shows the pole migration of both models over the parameter space with piece-wise linear interpolation between grid points. The plot confirms that the reduced-order
model obtained by parameter-varying projection indeed retains continuous dependence on the parameter. It further shows that the loci of the lightly damped modes in the reduced-order model almost exactly coincide with those of the original full-order model.

Figures 5.14 and 5.15 finally show frequency response plots of the full-order and the reduced-order model evaluated for frozen parameter values. The reduced-order model agrees very well with the full-order model up the specified frequency of $100 \text{ rad/s}$. 

Figure 5.12: LPV simulation of a step response with varying parameter: full-order model (--- 48 state variables) and reduced-order model (— 12 state variables).
5.1 Model Order Reduction for Aeroservoelastic Aircraft

Figure 5.13: Pole migration across flight envelope.

(a) Full-order model (48 state variables).
(b) Reduced-order model (12 state variables).

Figure 5.14: Parameterized frequency response from outboard flaps $\delta_1$ to wing tip acceleration.

(a) Full-order model (48 state variables).
(b) Modal interpolation (12 state variables).
5 Application Examples

Figure 5.15: Frequency response of full-order model (48 state variables) at 33 m/s (---) and 65 m/s (---) airspeed compared to the proposed reduced-order model with 12 state variables (---/---).
5.2 Model Order Reduction for Far Wakes of a Wind Turbine

As a large-scale example for LPV model order reduction, an unsteady aerodynamics problem known as the actuator disk model [Sørensen & Myken 1992, Sørensen & Kock 1995, Amnoni et al. 2016] is considered. It can be used to accurately model the far wake of a wind turbine by solving the two-dimensional Navier-Stokes equations for incompressible flows. Such models are relevant to study the aerodynamic interactions in a wind farm and to develop control strategies to maximize the overall power output of several turbines that are located close to each other.

The streamwise ($x$) and spanwise ($y$) velocity components are denoted $u$ and $v$ and their dynamics are governed by the partial differential equations

\[
\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -\frac{1}{\rho} \frac{\partial P}{\partial x} + \nu \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + f \tag{5.2a}
\]

\[
\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} = -\frac{1}{\rho} \frac{\partial P}{\partial y} + \nu \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) \tag{5.2b}
\]

where $\nu$ is the kinematic viscosity and $P$ is the pressure distribution. The forcing term $f$ depends linearly on the thrust coefficient $C_T$ of the turbine. This coefficient can be changed on a wind turbine via blade pitch or a change of the tip speed ratio and is thus a control input.

The particular configuration studied in this section consists of two wind turbines, each with rotor diameter $d$, that are located $5d$ apart from each other in a two-dimensional stream of air. A prescribed inflow and a convective outflow condition are used, leading to the boundary conditions

\[
\begin{align*}
    u|_{x=0} &= U_\infty, & \frac{\partial u}{\partial x}|_{x=20d} + U_\infty \frac{\partial u}{\partial x}|_{x=20d} &= 0, \\
    v|_{x=0} &= 0, & \frac{\partial v}{\partial x}|_{x=20d} + U_\infty \frac{\partial v}{\partial x}|_{x=20d} &= 0. 
\end{align*} \tag{5.2d}
\]

The upstream turbine runs with a constant thrust coefficient, while the downstream turbine’s thrust coefficient is considered as a control input. The output is a measurement of the spanwise velocity component $v$, located $5d$ downstream from the second turbine, indicating far wakes. The partial differential equations are solved following standard computational fluid dynamics methods with a central difference scheme for spatial discretization. The grid is defined by 201 points in the streamwise direction and 51 points in the spanwise direction. The discretization yields an ordinary differential equation system with 20502 state variables that depends parametrically on the freestream velocity $U_\infty$, or in nondimensionalized form on the Reynolds number $Re$.

Solving Lyapunov equations for a linearization of this large-scale system is intractable and hence empirical Gramians are constructed from simulation data as described in
Section 3.1.3. The matrices of sampled impulse responses are calculated for constant parameter values on the grid \( Re = \{10, 20, 30, 40, 50\} \). A forward Euler scheme and time steps of \( t_s = 0.01 \text{s} \) are used for time propagation with a time horizon of \( t_N = 50 \text{s} \). The resulting trajectories are sampled every 0.5 s and consequently the factorizations of the empirical Gramian at each grid point are of size 20502-by-100. These factors are then used in Algorithm 2 instead of the Cholesky factors to construct a parameter-varying oblique projection.

Simulation results for a sequence of step inputs and two different parameter trajectories are shown in Figure 5.16a and Figure 5.16b. The responses of the reduced-order model are in excellent agreement with the full-order model and the speed of parameter variation has no impact on the quality of the approximation. This once more confirms that there is no neglected rate-dependence in the reduction using the proposed parameter-varying oblique projection.

![Graphs showing simulation results](image)

**Figure 5.16:** Nonlinear simulation of full-order model (--- 20502 state variables) and reduced-order model (--- 6 state variables).

While the quality of the reduction should be strictly judged by how well the reduced-order system captures the considered input-output behavior, it remains insightful to also take a look at the approximated state vector \( \hat{\mathbf{x}} = \mathbf{Vz} \) which can be calculated from the reduced-order model state \( \mathbf{z} \) and the basis \( \mathbf{V} \). For the considered problem, the original state vector has a clear physical interpretation, namely velocities in \( x \) and \( y \) direction at each of the 10251 nodes in the domain. Figures 5.17a and 5.17b depict the baseflow that corresponds to a frozen-in-time snapshot taken from the simulation shown in Figure 5.16b with a constant thrust coefficient \( C_T = 30 \) for both turbines. As expected, the reduced-order model is not able to completely resolve the full state accurately. Still,
characteristic features of the stream are preserved up to the measurement point. The states of the reduced-order model can thus still be related to physically meaningful quantities. Velocities further downstream have little importance for the considered output and are hence less accurately resembled by the reduced-order model.

Figure 5.17: Flow at $t = 40s$ corresponding to the simulation shown in Figure 5.16b with locations of turbines (|) and the measurement point (⊕).
5.3 Active Flutter Suppression

This section describes the process of designing a controller for active flutter suppression on the University of Minnesota’s miniMUTT aircraft. The miniMUTT is a small, remote-piloted, flexible aircraft that resembles Lockheed Martin’s Body Freedom Flutter Vehicle (Section 5.1.1) and NASA’s X56A MUTT aircraft (Section 5.1.2). For the miniMUTT aircraft, collocated acceleration feedback tends to destabilize the short period dynamics. This is attributed to two facts. First, the collocated control surfaces (outboard flaps) have a much higher pitch effectiveness on a flying-wing compared to ailerons on a conventional aircraft. Second, the frequencies of the short period dynamics and the aeroelastic modes are very close to each other, making a frequency separation difficult to achieve. Therefore, the mixed sensitivity loopshaping approach of Section 4.2.3 is used.

A mathematical model of the miniMUTT aircraft is first described, with an emphasis on accurately capturing phase loss due to parasitic dynamics. A robust LTI controller that increases structural damping and suppresses flutter is then designed using the systematic procedure which was developed in Section 4.2.3. In the absence of a high-fidelity nonlinear model, the controller is finally analyzed with respect to a variety of linear stability margins.

5.3.1 Model of the miniMUTT Aircraft

The miniMUTT, built at the University of Minnesota, Minneapolis, is an unmanned flying wing aircraft with a wing span of 3 m and a total mass of about 6.7 kg. The design closely resembles Lockheed Martin’s Body Freedom Flutter Vehicle but has the modular wing concept of NASA’s X56 MUTT aircraft. To the present day, four miniMUTT aircraft were build. The first build (Fenrir) is shown in Figure 5.18 and is now out of service after a successful proof-of-concept flight test campaign conducted in 2014 and 2015. The second build (Skoll) and third build (Hati) were destroyed during system identification flights in August 2015 and April 2016. The fourth build (Geri) is currently in use by the University of Minnesota as part of NASA’s Performance Adaptive Aeroelastic Wing research program.

![Figure 5.18: miniMUTT unmanned aircraft.](image-url)
5.3 Active Flutter Suppression

The miniMUTT is designed such that it exhibits strong coupling of rigid-body dynamics and structural dynamics at low airspeed. Flutter occurs at an airspeed of approximately 30 m/s. Without active flutter suppression, the inevitable result is catastrophic structural failure as shown in the picture sequence in Figure 5.19.

![Figure 5.19: Open-loop flutter and catastrophic failure during a flight test slightly above 30 m/s indicated airspeed at the University of Minnesota on August 25th, 2015.](image)

When flutter was observed at 30 m/s airspeed in the flight tests, the aircraft was already running on full throttle. An envelope expansion beyond the flutter speed thus is also limited by the propulsion system. As the flutter dissipated a certain amount of energy, it appears possible to fly at 33 m/s once the oscillations are controlled. Still, the operating range appears to be sufficiently narrow to justify an LTI control design with the main objective to stabilize flight at 33 m/s. The controller must also provide enough safety margin to maintain stability at higher velocities that might occur due to head wind gusts and unintended dive maneuvers.

**Airframe Model**

The airframe model is described in detail by Schmidt et al. [2016] and Pfifer & Danowsky [2016]. The model is based on a mean-axes description and considers only longitudinal dynamics for straight and level flight under small elastic deformations. It contains four state variables associated with rigid-body dynamics, namely the forward velocity $u$, angle of attack $\alpha$, pitch angle $\theta$, and pitch rate $q$. Additionally, the first three symmetric free vibration modes are included in the model. They are described by their generalized displacements $\{\eta_i\}_{i=1}^3$ and velocities $\{\dot{\eta}_i\}_{i=1}^3$ with reference to the mode shapes depicted in Figure 5.20.

![Figure 5.20: Mode shapes of the first three symmetric structural modes of the miniMUTT aircraft [Schmidt et al. 2016].](image)
A state space representation with state vector $x = [u \quad \alpha \quad q \quad \eta_1 \quad \eta_2 \quad \eta_3 \quad \dot{\eta}_3]^T$ which parametrically depends on the airspeed $V_{\infty}$ is

$$
\dot{x} = \begin{bmatrix}
X_u & X_\alpha & X_q & X_{\eta_1} & X_{\eta_2} & X_{\eta_3} & X_{\dot{\eta}_3} \\
Z_u & Z_\alpha & Z_q & Z_{\eta_1} & Z_{\eta_2} & Z_{\eta_3} & Z_{\dot{\eta}_3} \\
V_{\infty} & V_{\infty} & -\frac{g}{V_{\infty}^2} & 0 & 0 & 0 & 0 \\
\frac{V_{\infty}}{\alpha} & \frac{V_{\infty}}{q} & \frac{V_{\infty}}{\eta_1} & \frac{V_{\infty}}{\eta_2} & \frac{V_{\infty}}{\eta_3} & \frac{V_{\infty}}{\dot{\eta}_3} & \frac{V_{\infty}}{\dot{\eta}_3} \\
M_u & M_\alpha & M_q & M_{\eta_1} & M_{\eta_2} & M_{\eta_3} & M_{\dot{\eta}_3} \\
\frac{M_u}{\alpha} & \frac{M_\alpha}{q} & \frac{M_q}{\eta_1} & \frac{M_{\eta_1}}{\eta_2} & \frac{M_{\eta_2}}{\eta_3} & \frac{M_{\eta_3}}{\dot{\eta}_3} & \frac{M_{\dot{\eta}_3}}{\dot{\eta}_3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix} x
$$

The plant input $\delta = [\delta_1 \quad \delta_2 \quad \delta_3 \quad \delta_4]^T$ consists of symmetric deflections of the outboard ($\delta_1$), midboard ($\delta_2$), inboard ($\delta_3$), and body ($\delta_4$) flaps. The entries $X_i$, $Z_i$, $M_i$ and $\Xi_{k,i}$ for $i \in \{u, \alpha, q, \eta_1, \eta_2, \eta_3, \dot{\eta}_3\}$ and $k = 1, 2, 3$ are dimensional aerodynamic derivatives. The entries $\omega_k$ and $\zeta_k$ are the eigenfrequencies and damping ratios of the $k$th structural mode and $g$ denotes the gravitational acceleration. The values of the aerodynamic coefficients were initially computed by Schmidt et al. [2016] using a VLM. Flight data, obtained in system identification flights, were used to update the coefficients $Z_i$, $M_i$, $\Xi_{i,k}$ for $i \in \{\alpha, q, \eta_1, \eta_2, \eta_3\}$ by Pfifer & Danowsky [2016]. These coefficients are associated with the short period dynamics and the first structural mode.

In order to simplify the synthesis model, the state variables $u$ and $\theta$ are removed by truncation. The resulting model thus only consists of eight state variables, $\alpha, q, \eta_1, \eta_2, \eta_3$, and can be interpreted as an aeroelastic short period approximation. Two dominant dynamic modes are apparent and shown in Figure 5.21 for an airspeed $V_{\infty} = 33\text{ m/s}$. The first mode is well damped, has a frequency of around $30\text{ rad/s}$ and consists mostly of $\hat{\eta}_1, q, \hat{\eta}_2$, and $\alpha$ contributions. In particular, the relation of $\alpha$ lagging $q$ by about $90^\circ$ is reminiscent of a classical short period mode, such that this mode will also be referred to as “short period” mode. The second mode is highly oscillatory with

\footnote{The “magnitude” of the individual contributions depends on how the involved state variables are normalized, such that the phasor diagrams should be interpreted qualitatively.}
a frequency of $33 \text{rad/s}$ and is unstable. This mode consists almost entirely of the structural deformation velocity $\dot{\eta}_1$ and is hence termed “aeroelastic” mode in the following. These names are merely used for convenience and it must be emphasized that both the “short period” and “aeroelastic” mode contain contributions from all state variables, which shows that there is no clear separation between rigid-body and structural dynamics.

The midboard and inboard flaps remain reserved for pitch and roll control by the pilot. For the initial flutter suppression control design, only the outboard flaps are used, i.e., $u = \delta_1$. Keeping the flutter suppression control loop completely separate from pilot inputs reduces the risk of saturating the control surfaces and facilitates a simple control design. As measurements, the pitch rate and the vertical acceleration at both the center of gravity and at the wing tips are used, i.e., $y = [\dot{q}_{\text{meas}}, a_{z,\text{center}}, a_{z,\text{wing}}]^T$. The corresponding output equation $y = Cx + D\delta$ is obtained using the mode shapes of the structural modes as detailed by Schmidt et al. [2016]. A schematic showing the aircraft with the sensor and actuator positions is depicted in Figure 5.22.

### Time Delay and Phase Loss Modeling

For regular flight control systems, the sampling rate is much higher than the closed-loop bandwidth and the induced phase loss from sensors and actuators is often negligible. On the contrary, active suppression of the flutter instability at high frequency requires a very high closed-loop bandwidth. Actuator and sensor dynamics are not negligible in this frequency regime. Time delay, introduced by digitalization effects and computation, also has a big impact on the control loop. The goal of this subsection is to describe and model all known parasitic dynamics. The miniMUTT aircraft is designed as a low-cost research platform, leading to a relatively simple systems architecture. Figure 5.23 shows all components in the feedback loop and how they are grouped into three models $P_{\text{sens}}$.

---

4Later, the body flaps could be added as suggested for the Body Freedom Flutter Vehicle in Section 5.1.1.
5 Application Examples

$P_{\text{delay}}$, and $P_{\text{act}}$. Including these dynamics in the synthesis model allows the controller to compensate for known phase loss and hence to improve performance and robustness.

The pitch rate measurement on the miniMUTT is obtained by an IMU that includes a 50 Hz low-pass filter. The accelerometer signals are filtered by an analog first-order low-pass with a bandwidth of 35 Hz. These components are modeled by two first-order transfer functions $P_{\text{acc}}(s) = \frac{1}{s/(2\pi35) + 1}$ and $P_{\text{imu}}(s) = \frac{1}{s/(2\pi50) + 1}$. The signals provided by the sensors are processed by the miniMUTT’s flight computer that executes the control algorithm within a 6.6 ms frame. The controller output is passed on to a microcontroller that runs asynchronous with a 3.3 ms frame rate to generate a pulse width modulation (PWM) signal. This PWM signal is the input to a servo controller that runs, also asynchronous,
with a 3.3 ms frame rate. This results in a maximum of 13.2 ms total computational delay. Further, the physical inertia of the actuators introduces additional low-pass characteristics, described by a second-order model \( P_{\text{act}}(s) \).

To keep the controller order low, actuator dynamics, sensor dynamics, and delay are combined into a low-order equivalent model. Obtaining this model requires a shift of the sensor dynamics from the plant output to the input, which is only possible if all sensors are modeled identically. The slower dynamics of the accelerometers are therefore also assumed for the faster IMU and both are uniformly modeled as \( P_{\text{sens}}(s) = P_{\text{accel}}(s) \). Further, all computational frames are added up and a factor of 1.5 is included in order to anticipate the zero-order hold delay. To further account for actuator and sensor delays, a total delay of 25 ms is assumed and modeled as \( P_{\text{delay}}(s) = e^{-0.025 s} \). A second-order model is calculated from balancing and residualization of \( P_{\text{act}}(s) P_{\text{delay}}(s) P_{\text{sens}}(s) \), where a fifth-order Padé approximation is used for the time delay. The resulting model is shown in Figure 5.24a and captures the phase loss accurately up to about 100 rad/s. Figure 5.24b further illustrates the phase loss contributions of the known parasitic dynamics in the critical frequency range in detail. The largest contribution comes from the time delay, followed by the actuator and sensors. The resulting simplified loop is depicted in Figure 5.25.

![Graphs showing phase loss contributions](image)

**Figure 5.24:** Phase loss due to known parasitic dynamics: pure time delay (---), plus actuator dynamics (--), plus sensor dynamics (---), second-order approximation for synthesis model (---).

### 5.3.2 Control Law Design

The model \( P(s) \) combines the simplified eighth-order airframe model which represents the miniMUTT at 33 m/s airspeed and the second-order equivalent model for actuator
dynamics, delay, and sensor dynamics. It hence has ten state variables. The measurable output \( y \) used for feedback consists of pitch rate \( q_{\text{meas}} \) (in \( \text{rad/s} \)), vertical center acceleration \( a_{z,\text{center}} \), and vertical wing tip acceleration \( a_{z,\text{wing}} \) (both in \( \text{m/s}^2 \)). The plant input is the symmetric deflection of the outboard flaps \( \delta_1 \) (in rad).

The high-level objective of the flutter suppression controller is to stabilize and attenuate the aeroelastic mode without impairing handling of the aircraft by the pilot. As was shown in Figure 5.21c, the aeroelastic mode almost exactly coincides with the structural modal velocity \( \dot{\eta}_1 \). Since the model (5.3) explicitly contains \( \dot{\eta}_1 \) as a state variable, it is decided to use \( \dot{\eta}_1 \) instead of actual modal state space coordinates in the definition of an additional performance output \( y_p \) as introduced in Section 4.2.3. The resulting transfer function \( P_p(s) \) has a band-pass characteristic and a sharp peak at the flutter frequency. It is combined with the plant model to form the partitioned synthesis model \( \begin{bmatrix} P \, P_p \end{bmatrix} \).

The controller further needs to provide robustness against a wide class of possible...
uncertainties in the model. The proposed generalized plant interconnection which translates these goals into the objective of minimizing the closed-loop H_{\infty}-norm is depicted in Figure 5.27. It represents the input-output map
\[
\begin{bmatrix}
  z_1 \\
  z_2 \\
  z_3
\end{bmatrix} =
\begin{bmatrix}
  W_y D_y^{-1} & W_u D_u^{-1} & W_p
\end{bmatrix}
\begin{bmatrix}
  T_o & S_o P & D_y \\
  C S_o & T_i - T_i & D_u \\
  P P S_i C & P P S_i & D_u
\end{bmatrix}
\begin{bmatrix}
  w_1 \\
  w_2
\end{bmatrix}.
\tag{5.4}
\]

This generalized plant is slightly different from the one introduced in Section 4.2.3. The sensitivity function S_o in the upper left block is replaced by the complimentary sensitivity T_o, i.e., y is penalized, not e. This modification is only possible when no tracking or steady-state disturbance rejection requirements are included. That is, the weight W_y that penalizes both S_o P_y and T_o has to be static in order to avoid ill-posed specifications such as a simultaneous reduction of both sensitivity and complementary sensitivity. Further, disturbances are modeled at the plant input with the implicit assumption D_d = D_u to ensure that both the input and output complementary sensitivities are included in a sensible way in the optimization problem. In particular, this ensures symmetrically scaled sensitivity functions \( T_o = D_y^{-1} T_o D_y \) and \( T_i = D_u^{-1} T_i D_u \). This setup turned out to result in more robust controllers than using the generalized plant of Section 4.2.3. This result is believed to be attributable to the complementary sensitivity representing multiplicative uncertainty (cf. Section 4.4), which appears to be a better uncertainty model for the present system than the inverse multiplicative uncertainty that corresponds to S. It should also be noted that common performance and robustness criteria include peak values for T of less than 2 dB and peak values for S of less than 6 dB [Skogestad & Postlethwaite 2005, Sec. 2.4.3] and that further “for unstable plants, \( \|T\| \) is usually larger than \( \|S\| \)” [ibid., p. 36]. These statements also hint at including T in the problem formulation if robustness is the primary objective.

Figure 5.27: Generalized plant interconnection for the flutter suppression control design.
essentially set through the weight $W_p$. Larger values result in higher damping augmentation as the weight encourages the controller to reduce the sharp peak in the frequency response. The weight $W_u$ is used to limit control action such that interaction with rigid-body dynamics in the low-frequency regime and with unmodeled high-frequency dynamics is avoided. Selecting $W_u$ as a band-stop filter, as shown in Figure 5.27b, results in band-pass behavior for both the input complementary sensitivity and the control sensitivity. Thus, both for low and high frequencies $T_i \approx 0$ which implies $S_i \approx I$ and consequently $S_i C = C S_o \approx C$. The band-stop weight $W_u$ thus directly shapes $C$ both at low and high frequencies and results in wash-out and roll-off characteristics.

**Design and Tuning**

The weights for the mixed sensitivity formulation (5.4) are selected as follows. As the only objective on $T_o$ and $S_o P_y$ is to flatten sharp peaks, $W_y = I$ is a simple yet sensible choice. The weight for the control effort is selected as the series interconnection of a filter with (approximate) integral behavior up to $20 \text{rad/s}$ and a filter with (approximate) differentiating behavior beyond $40 \text{rad/s}$. The resulting band-stop filter $W_u = \frac{50 s^2 + 3732 s + 69280000}{s^4 + 4000 s + 0.6928}$, shown in Figure 5.27b, thus restricts activity of the flutter suppression controller to the frequency region around the aeroelastic mode. To increase robustness margins of the closed-loop, $W_u(s)$ can be increased in the frequency region where the margin is attained (either below or above the region of interest) to decrease the controller gain at that frequency.

The scalings are used as the main tuning knobs as described in Section 4.2.3. They provide an intuitive way of including qualitative knowledge about the system. The input scaling is selected as $D_u = \pi/180$ and the output scaling is selected as $D_y = \text{diag}(1, 9.81, 9.81)$. These values essentially normalize the individual transfer functions to gains around 0dB to be compatible with the dimensionless sensitivity functions $S$ and $T$. The choice can also be given the interpretation of 1g acceleration per $\text{rad/s}$ pitch rate and 1° flap deflection as relative values that agree quite well with simulations of the open-loop system. The desired increase in damping, and hence the flutter margin, is set by the weight $W_p$. The input-output map $P_p D_u$ already has its peak value slightly above 0dB as a consequence of the input scaling. Therefore, a choice of $W_p = 1$ is also sensible.

A controller is obtained using Matlab’s hinfsyn routine with a 10% suboptimal synthesis, resulting in a performance index $\gamma = 3.0$. The resulting controller $C(s)$ is shown in Figure 5.28. The desired band-pass behavior is apparent. The peak gain for both center acceleration and wing tip acceleration signals is attained at the same frequency around $40 \text{rad/s}$, but their phase differs considerably. The wing tip acceleration lags the center acceleration by up to 45°. This shows that the proposed controller would be impossible to obtain by a simple combination of the acceleration signals in a single loop. The controller has twelve state variables. Its fastest pole is at 126 $\text{rad/s}$ and thus well within the permissible region for digital implementation on the flight computer.\(^4\)

Figure 5.29a shows the open-loop and closed-loop transfer function $P_p S_i$ used to specify damping augmentation. The disturbance sensitivity $PS_i$ which relates inputs to the

\(^4\)It should be noted that without a suboptimal synthesis, the fastest controller pole is at 3500 $\text{rad/s}$.\(^3\)
5.3 Active Flutter Suppression

![Bode plot of the flutter suppression controller.](image)

Figure 5.28: Bode plot of the flutter suppression controller.

measurable outputs is shown in Figure 5.29b. The sensitivity is in both cases lowered at the frequency of the aeroelastic mode, but as a consequence increased at neighboring frequencies. One important aspect of the control design is to confine this sensitivity degradation to a specific frequency region. Figure 5.29b shows that this is indeed achieved and that neither the low-frequency phugoid nor the high-frequency elastic modes are affected by the flutter suppression controller.

![Sensitivity of first structural mode (PpSs).](image)  ![Load disturbance sensitivity (PSi).](image)

Figure 5.29: Open-loop (•••) and closed-loop (—) transfer functions.

In order to assess the interaction with pilot commands, a comparison of open-loop and closed-loop step responses to midboard flap deflection is shown in Figure 5.30 for two different airspeeds. These flaps are used by the pilot to control the longitudinal motion of the aircraft. The pilot essentially closes a pitch angle feedback loop, since his main
visual indicator for control is the vehicle’s attitude. Maintaining a pitch response as close as possible to the open-loop aircraft is thus considered desirable. At the design airspeed of 33 m/s where the aircraft is naturally unstable, the highly oscillatory and divergent pitch rate response is effectively damped out and stabilized by the controller as shown in Figure 5.30a. This is achieved without affecting the initial transients up to about 0.15 s. The pitch response at the lower airspeed of 24 m/s is barely affected at all by the presence of the flutter suppression controller. The aircraft’s immediate response to pilot inputs is thus identical with and without flutter suppression, both for low and high airspeeds. The flutter suppression controller introduces no additional delay or phase lag, that could impair handling. The effectiveness of the controller is further visible in the acceleration responses in Figure 5.30b.

![Pitch response](image1.png) ![Acceleration response](image2.png)

**Figure 5.30:** Open-loop responses at 24 m/s (---) and 33 m/s (•••), and closed-loop responses at 24 m/s (- - -) and 33 m/s (-----) to step input at midboard flaps.

The effect of the flutter suppression controller on the pole locations is shown in Figure 5.31. The open-loop model exhibits flutter at an airspeed above 30 m/s, indicated by the poles of the aeroelastic mode crossing into the right half plane at about 33 rad/s in Figure 5.31a. Figure 5.31b shows that the locus of the aeroelastic mode is altered by the controller to stay within the left half plane with a drastic improvement in damping. The controller poles appear as transmission zeros of the sensitivity function \( PS_i \) and are also shown. The short period frequency is lowered, with a marginal decrease in damping for an airspeed below 30 m/s and a slight increase in damping beyond that speed. The remaining loci correspond to closed-loop poles resulting from the controller poles. Extrapolation of the model to a higher airspeed, i.e., evaluating Equation (5.3) for different values of \( V_\infty \), further shows that flutter now occurs at 43 m/s airspeed. This corresponds to an envelope
5.3 Active Flutter Suppression

(a) Open-loop pole migration.

(b) Closed-loop pole migration.

Figure 5.31: Effect of the flutter suppression controller on pole locations. Controller poles (○) appear as closed-loop zeros.

expansion of 13 m/s (43%) and is deemed a more than sufficient safety margin for the desired flight point at 33 m/s.

5.3.3 Control Law Verification

Given the catastrophic consequences of flutter, it is paramount that the controller is highly robust. Without a high-fidelity nonlinear model for evaluation and with limited possibilities for testing outside of the critical flight regime, a thorough linear analysis is required. The robustness tests described in this subsection aim at maximizing the likelihood of a successful flight with the developed controller. Disk margins, both one-loop-at-a-time and for multivariable loops, are considered. Further, structured singular values are used to evaluate robustness with respect to parametric uncertainties in the aircraft model. Specifically, uncertainty in the structural model, the aerodynamics model, and the actuator model is considered. All robustness calculations are performed on a model that includes sensor dynamics, actuator dynamics, delay, the first three structural modes and complete rigid-body dynamics. The analysis results of this section are thus to be understood “on top” of all known parasitic dynamics.

First, a disk margin analysis is performed to capture generic model uncertainty. The resulting robustness margins are depicted in Figure 5.32 as a function of airspeed. All margins uniformly increase with lower airspeed to a similar extent. This indicates a smooth variation without any particular robustness bottlenecks. The worst input disk margin is around 8 dB (45°) and the worst single-loop output disk margin is slightly below 11 dB (60°). The multi-output disk margin, corresponding to simultaneous perturbation of all three outputs, is also calculated and remains above 6 dB (37°). For independent perturbations of all outputs and the input, the multi-input-multi-output disk margin
remains above 3.5 dB (23°). The lowest delay margins are also attained at 33 m/s airspeed. They are 39 ms at the output and 22 ms at the input. As a delay of 19.8 ms would correspond to one full dropped frame from every single computational unit and the induced zero-order hold delay, the achieved margin appears to be sufficient.

Figure 5.32: Minimum robustness margins as a function of airspeed: single-loop disk margin at input (↑) and output (↓), multi-output disk margin (〈), and multi-input-multi-output disk margin (→).

Next, the analysis is narrowed to specific sources of uncertainty within the model structure. The models for both structural dynamics and aerodynamics are best described as uncertain with respect to real parameters. Structured singular value analysis provides an efficient way to calculate stability margins for such structured uncertainties. Three different sets of uncertainties are considered. Structural model uncertainty in the following refers to a real parametric uncertainty in the eigenfrequency $\omega_1$ of the first structural mode. The parameter $\omega_1$ in Equation (5.3) is hence replaced by $(1 + \Delta)\omega_1$, where $\Delta \in \mathbb{R}$ with nominal value zero spans the range of possible variation, e.g., $|\Delta| < 0.1$ for 10% uncertainty. Likewise, aerodynamic uncertainty refers to real perturbations in the aerodynamic coefficients for pitch moment ($M_\alpha$, $M_q$), lift ($Z_\alpha$, $Z_q$), and influence of the first structural mode ($\Xi_{1\alpha}$, $\Xi_{1q}$, $\Xi_{1\delta}$) in Equation (5.3). All real parametric uncertainties are “complexified” with a 5% dynamic uncertainty to regularize the resulting computational problem [Packard & Pandey 1993, Balas et al. 2014]. Actuator uncertainty refers to a norm-bounded complex multiplicative uncertainty in the actuator model, i.e., $P_{\text{act}}$ is replaced by $(1 + \Delta)P_{\text{act}}$, where $\Delta \in \mathcal{H}_\infty$ is a norm-bound, stable LTI dynamic uncertainty with nominal value zero that represents the range of variation, e.g., $\|\Delta\| < 0.1$ for 10% uncertainty.

Figure 5.33a shows the stability boundaries for parameter variations along with a robust performance analysis. Instability occurs first for uncertainty in the structural mode frequency. This frequency was obtained from ground vibration tests for the present
5.3 Active Flutter Suppression

model by Gupta & Seiler [2016] and is expected to be known very accurately, up to at least 2%. Thus, the stability margin of over 12% is more than sufficient. The highest uncertainty is expected in the aerodynamics model. The analysis shows that the controller is highly robust with respect to this uncertainty, tolerating up to 50% perturbations. The permissible actuator uncertainty for robust stability is even higher and is slightly below 70%.

(a) Robust performance analysis for three sets of structured uncertainties.

(b) Robust performance analysis with a large uncertainty set.

Figure 5.33: Robust performance analysis for parametric uncertainties in the structural model (- - -), aerodynamic model (-----), actuator model (---), and a combination of these (●).

The robust performance index is calculated as the ratio \( \| P_p S_i \| / (\max_\sigma \sigma_{\max}(P_p)) \) of the worst-case \( H_\infty \)-norm of the uncertain closed-loop system and the largest gain of the nominal open-loop. It thus measures the amount of damping augmentation that is provided by the flutter suppression controller with values less than one indicating additional damping. The performance degradation for all three cases is qualitatively similar and can be characterized as graceful. Small variations result in small performance degradation which only starts to increase significantly close to the stability boundary. For individual uncertainties below 7% in the structural model, 28% in the aerodynamics model, and 50% in the actuator model, the ratio of closed-loop and open-loop gain is less than one. In these cases, the controller provides additional damping to the aeroelastic mode and hence achieves robust performance.

A fourth analysis is shown in Figure 5.33b for an uncertainty set that combines all aforementioned uncertainties. Even in this case, performance degradation is smooth and graceful. As could be expected, the stability margin is considerably lower than for the individual uncertainties. Stability is certified up to simultaneous 2.5% structural mode uncertainty, 25% aerodynamic uncertainty and 25% actuator uncertainty. Robust performance is achieved up to simultaneous 1.5% structural mode uncertainty, 15% aerodynamic uncertainty and 15% actuator uncertainty. These values are plausible for the aircraft under consideration and the results are therefore very encouraging.
5 Application Examples

5.4 Integrated Aeroservoelastic Control

This section focuses on integrating aeroelastic control objectives and rigid-body control. The proposed design combines stability augmentation, structural mode control and maneuver demand in a single controller. The problem has several similarities with the flutter suppression control problem discussed in Section 5.3, but merging structural control objectives with primary flight control objectives provides an additional challenge. Further, the operating range for the considered model is large such that an LPV controller is designed to automatically adjust for variations in the model due to different operating conditions.

As a representative example, a freely available nonlinear model resembling the Rockwell B-1 Lancer is used [Schmidt 2013]. The B-1 aircraft, shown in Figure 5.34, is a four-engine strategic bomber developed and manufactured by Rockwell International (now Boeing). The aircraft is 44.5 m long, has a wing span of 42 m, an empty weight of 87.1 t and a maximum takeoff weight of 216.4 t. For supersonic flight, the wings are swept back, but the considered model is only concerned with the subsonic extended-wing configuration. The B-1 aircraft had serious ride quality issues due to adverse aeroelastic effects causing structural vibrations at the cockpit. This necessitated the development of a structural mode control system (SMCS) with additional control vanes being placed at the nose of the aircraft as effectors [Wykes et al. 1977].

![Figure 5.34: B-1 aircraft.](https://en.wikipedia.org/wiki/File:B1s.jpg)

---

### 5.4 Integrated Aeroservoelastic Control

Such a dedicated SMCS can be very effective to reduce aeroelastic vibrations, but also prevents the conventional control surfaces from engaging in aeroelastic control and the surplus effectors from aiding in rigid-body control. It therefore does not use the full potential of multiple control surfaces. The present section seeks to overcome these limitations and to combine stability augmentation, maneuver demand and aeroelastic control into a fully integrated multivariable LPV controller.

#### 5.4.1 Model of the B-1 Aircraft

A freely available nonlinear Matlab/Simulink simulation resembling the B-1 aircraft was recently developed by Schmidt [2013] from data available in the open literature. It has sparked renewed interest in the B-1 as a benchmark example of an aeroelastic aircraft [e.g. Hjartarson et al. 2014]. The simulation model includes nonlinear rigid-body dynamics, nonlinear quasi-steady aerodynamic data, actuator and servo models, as well as the three lowest-frequency symmetric and two lowest-frequency antisymmetric elastic deformation modes of the airframe. Being modeled in a mean-axes formulation, its structure resembles that of the mini MUTT model (5.3) described in Section 5.3.1. The state variables associated with the rigid-body dynamics are the velocities in body-frame \((x, y, z)\)-direction \((u, v, w)\), pitch, roll, and yaw rates \((p, q, r)\) with respect to the center of gravity and \((\Theta, \Phi, \Psi)\) denoting pitch, bank, and yaw attitude. The generalized displacements of the elastic modes are denoted \((\eta_1, \eta_2, \eta_3)\) for the symmetric and \((\eta_4, \eta_5)\) for the antisymmetric deformation. Figure 5.35 depicts the geometric displacement shapes of the individual modes. The displacements \(\eta_1, \eta_3\) and \(\eta_5\), shown in Figure 5.35a, mainly cause fuselage bending which contributes to vibrations at the cockpit. The displacements \(\eta_2\) and \(\eta_4\), shown in Figure 5.35b, correspond primarily to wing bending modes.

![Figure 5.35: Structural modes in the B-1 model](Image)

(a) Mode shapes of structural modes \(\eta_1\) (vertical displacement of fuselage and wing), \(\eta_3\) (vertical displacement of fuselage and horizontal tail), and \(\eta_5\) (lateral displacement of fuselage and vertical tail). (b) Mode shapes of structural modes \(\eta_2\) (vertical displacement of fuselage and wing) and \(\eta_4\) (lateral displacement of fuselage and vertical tail and vertical displacement of right wing and right horizontal tail).
Available measurements include the airspeed $V_\infty$, the three rotational rates ($p, q, r$), and accelerations $a$ in both $y$ and $z$ direction at the center of gravity and cockpit. The available effectors are the symmetric ($\delta_H$) and anti-symmetric ($\delta_{\text{H1}}$) all-movable horizontal tail, wing upper-surface spoilers ($\delta_{\text{sp}}$), as well as upper ($\delta_{\text{RL}}$) and lower ($\delta_{\text{RL}}$) split rudder.

The existing control system in the simulation consists of SAS, active up to about $5 \text{ rad/s}$, and dedicated SMCS, only active in a frequency range of about $5-30 \text{ rad/s}$. The SAS is formed of pitch rate ($q$) damper, roll rate ($p$) damper, washed-out yaw rate ($r$) damper and lateral acceleration ($a_{y,\text{cg}}$) feedback for turn coordination. The SMCS employs the principle of collocated accelerometers and effectors and generates forces opposing the current deformation velocity to actively reduce vibrations at the cockpit, indicated by the accelerations $a_{y,\text{cockpit}}$ and $a_{z,\text{cockpit}}$ [Wykes et al. 1977]. It uses the symmetrically ($\delta_{\text{cv,sym}}$) and anti-symmetrically ($\delta_{\text{cv,anti}}$) deflectable control vanes at the cockpit as effectors.

Figure 5.36 shows the root-loci for the existing SMCS which increases the damping of the $\eta_1$ and $\eta_5$ structural modes. These modes contribute heavily to vibrations at the cockpit and are hence well controllable via collocated feedback. The other structural modes, however, are only marginally affected by the control loops and the corresponding closed-loop poles remain weakly damped.

![Figure 5.36: Root-loci for existing SMCS with SAS in the loop evaluated at Mach 0.6 and 5000 ft altitude; open-loop poles (•), zeros (○), and closed-loop poles (×).](image-url)
5.4.2 Control Law Design

For control design, the nonlinear model is linearized at operating points on the rectangular grid $\mathcal{M} \times h = \{0.6, 0.7, 0.8\} \times \{5000\, \text{ft}, 10000\, \text{ft}, 15000\, \text{ft}, 20000\, \text{ft}\}$ in cruise condition, where $\mathcal{M}$ denotes the Mach number and $h$ the altitude. This leads to a gridded LPV model with scheduling parameter vector $\rho := [\mathcal{M} \ h]$. Rate bounds are selected as $|\dot{\mathcal{M}}| < 0.1 \, \text{s}^{-1}$ and $|\dot{h}| < 1000 \, \text{ft} / \text{s}$ based on a recent study by Hjartarson et al. [2014] concerned with the sensitivity of the existing control system to these dynamic variations. Both rate limits are conservative, as can be verified in simulation. The cruise condition used for linearization decouples longitudinal and lateral-directional dynamics, such that two distinct models are obtained. The complete longitudinal model contains 14 state variables representing pitch angle $\Theta$, altitude $h$, pitch rate $q$, forward and vertical velocities $u$ and $w$, as well as the $\eta_1$, $\eta_2$, and $\eta_3$ elastic modes and the actuator and servo dynamics. The complete lateral-directional model with 16 state variables contains bank angle $\Phi$, roll rate $p$, yaw rate $r$, lateral velocity $v$, the $\eta_4$ and $\eta_5$ elastic modes, as well as actuator and servo dynamics.

The following control objectives are considered: The control system should provide the pilot with the possibility to control the aircraft’s attitude via pitch rate and roll rate command inputs. Good handling qualities are to be achieved by providing additional damping to the dutch roll mode and short period mode. Further, ride quality should be enhanced by reducing any vibrations at the cockpit through damping augmentation of the fuselage bending modes $\eta_1$, $\eta_3$, and $\eta_5$. Finally, structural loads should be reduced, which are primarily associated with the wing bending modes $\eta_2$ and $\eta_4$.

**Longitudinal Control**

The longitudinal control design is concerned with the short period and structural dynamics such that pitch angle, altitude, and forward velocity are truncated from the full model to obtain the synthesis model. This effectively removes the phugoid mode. As measured outputs, the pitch rate $q$ and the vertical accelerations $a_{z,\text{cockpit}}$ and $a_{z,\text{cg}}$ are used. The acceleration measurements are combined to form a differential signal $\Delta a_z = a_{z,\text{cockpit}} - a_{z,\text{cg}}$ that measures the deviation between the acceleration at the cockpit and at the center of gravity. The control inputs are symmetric horizontal stabilizer deflection $\delta_H$ and symmetric control vane deflection $\delta_{\text{cv,sym}}$. The servo dynamics are fast enough to be considered irrelevant for the synthesis model, but the actuator models are retained to account for phase-loss. As additional performance outputs, the three elastic deformation velocities $\dot{\eta}_1$, $\dot{\eta}_2$, and $\dot{\eta}_3$ are included. The longitudinal plant model for synthesis hence is

$$
\begin{align*}
y \{ q, \Delta a_z \} & = \begin{bmatrix} P \ P_p \ P_p \ \ P_p \end{bmatrix}_P \rho \begin{bmatrix} \delta_H \ \delta_{\text{cv,sym}} \end{bmatrix} u. \\
y_p \{ \eta_1, \eta_2, \eta_3 \} & = \begin{bmatrix} \dot{\eta}_1 \ \dot{\eta}_2 \ \dot{\eta}_3 \end{bmatrix}
\end{align*}
$$

(5.5)
A two-degrees-of-freedom controller is used to acknowledge the tracking task, i.e.,

\[ u = K_\rho \begin{bmatrix} q_{\text{ref}} \\ -y \end{bmatrix} = \begin{bmatrix} C_{\text{FF}} & C_{FB} \end{bmatrix}_\rho \begin{bmatrix} q_{\text{ref}} \\ -y \end{bmatrix}, \tag{5.6} \]

where \( y \) denotes the measurable output used for feedback control and \( q_{\text{ref}} \) denotes a reference pitch rate command. A load disturbance model with input disturbances at all control surfaces is assumed, i.e., \( P_d = P \). This model is a simple representation for both gusts acting on the aircraft and actuator model uncertainties.

The generalized closed-loop interconnection used for the design of the controllers is

\[
\begin{bmatrix}
    z_1 \\
    z_2 \\
    z_3
\end{bmatrix} =
\begin{bmatrix}
    W_e D_e^{-1} & 0 & 0 \\
    0 & W_u D_u^{-1} & 0 \\
    0 & 0 & W_p
\end{bmatrix}
\begin{bmatrix}
    S_o & S P & S_r \\
    S_i C_{\text{FB}} & T_i & S_i C_{\text{FF}} \\
    P_p S_i C_{\text{FB}} & P_p S_i & P_p S_i C_{\text{FF}}'
\end{bmatrix}
\begin{bmatrix}
    D_e & D_d & D_r \\
    D_e' & w_1 \\
    w_2
\end{bmatrix}, \tag{5.7}
\]

where \( S_r \) denotes the error function to a reference command and \( D_r \) the scaling of the tracked output, i.e.

\[
S_r = \begin{bmatrix}
    1 \\
    0
\end{bmatrix} - S_o P C_{\text{FF}} \quad \text{and} \quad D_r = \begin{bmatrix}
    1 & 0
\end{bmatrix} D_e \begin{bmatrix}
    1
\end{bmatrix}.
\]

The design task is to select weights \( W_e \) and scalings \( D_u \) as discussed in detail in Section 4.2.3 such that demands on tracking, disturbance rejection, damping augmentation, robustness, and control usage are expressed. LTI weights are used in order to achieve uniform dynamic properties across the flight envelope despite the varying plant dynamics.

The present design starts with setting \( W_e \) to enforce sensitivity reduction up to \( 3 \text{rad/s} \) for pitch rate tracking and an upper bound of 2 on the sensitivity related to \( \Delta a_z \). The filter \( W_u \) could be chosen with differentiating behavior beyond the physical actuator capacities to enforce a roll-off, but it turns out that robustness is not restrictive for the present design. Thus, a simple constant of 1 is selected for the input \( \delta H \). For the input \( \delta_{cv,\text{sym}} \), a weight with integral behavior up to \( 1 \text{rad/s} \) and gain 1 beyond that frequency is used such that the control vanes are not deflected in steady state. Enforcing a zero deflection in steady state reduces the risk of saturation and frees up actuator capacity for use in the lateral control problem where antisymmetric deflection of the control vanes is required.

After the frequency-dependent weights are selected on control objective related reasoning, tuning is performed by altering the scalings. Physical insight provides very good initial guesses for these design parameters. The weight \( D_u \) is understood here as an “allowable maximum error” where a reference change (and hence instantaneous error) of \( 3/\text{s} \) pitch rate is assumed. Relative to this error, the acceleration scaling is selected as \( 32.17 \text{ft/s}^2 \) (1 g). Similarly, the weight \( D_d \) is used to express “allowable inputs” relative to the specified errors. Deflections of \( 10^\circ \) for the horizontal stabilizer and \( 20^\circ \) for the control vanes are selected. Next, the scaling \( D_d \) is increased until robustness margins at the plant input are satisfactory, leading to values of \( 1^\circ \) and \( 2^\circ \). With respect to the scaling \( D_u \), this choice can be given the interpretation of 10% uncertainty at both inputs, although the use as a tuning knob should be emphasized. The static weight \( W_p \) penalizes the peaks that correspond to the elastic modes and is finally selected such that these peaks are “exposed”
in the optimization problem, i.e., such that they are slightly above unit magnitude. Doing so leads to values of 0.01 for $\eta_1$ and $\eta_2$, and 1 for $\eta_3$.

The weights that are used in the present design are summarized in Table 5.1. Static weights satisfy $W = M$ and (approximately) integral weights are parameterized to satisfy $W(0) = 10^3$, $|W(j\omega_b)| = 1$, and $W(\infty) = 0.5$, i.e., $W(s) = 0.5 + 8.6603 \cdot 10^{-1}/s$.

<table>
<thead>
<tr>
<th>Weight on</th>
<th>$D_c$</th>
<th>$D_u$</th>
<th>$D_d$</th>
<th>$\omega_b$</th>
<th>$M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_e$</td>
<td>$q$</td>
<td>$3/180\pi$</td>
<td>—</td>
<td>—</td>
<td>3</td>
</tr>
<tr>
<td>$\Delta a_z$</td>
<td></td>
<td>32.17</td>
<td>—</td>
<td>—</td>
<td>0.5</td>
</tr>
<tr>
<td>$W_\alpha$</td>
<td>$\delta_H$</td>
<td>10</td>
<td>1</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>$\delta_{cv,sym}$</td>
<td>20</td>
<td>2</td>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>$W_p$</td>
<td>$\dot{\eta}_1$</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>$\dot{\eta}_2$</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>$\dot{\eta}_3$</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>1</td>
</tr>
</tbody>
</table>

### Lateral-Directional Control Problem Formulation

The purpose of the proposed lateral-directional controller is to provide the pilot with means to control the aircraft’s attitude through roll rate commands. The unstable spiral mode is removed from the full model by truncation after applying a common coordinate transformation, calculated from the modal form at the first grid point. Further, all the servo dynamics are neglected. Thus, the synthesis model is stable and has 10 states. The measured feedback signals are roll rate $p$, yaw rate $r$, and the lateral accelerations $a_y,\text{cockpit}$ and $a_y,\text{cg}$. Again, a differential acceleration signal $\Delta a_y = a_y,\text{cockpit} - a_y,\text{cg}$ is formed. Hence, the plant output is $y =$ \begin{bmatrix} p \\ r \\ \Delta a_y \\ \dot{\eta}_4 \\ \dot{\eta}_5 \end{bmatrix} = \begin{bmatrix} P \\ \dot{P} \end{bmatrix} \begin{bmatrix} \delta_{DH} \\ \delta_{RL} \\ \delta_{RU} \\ \delta_{sp} \\ \delta_{cv,anti} \end{bmatrix}$, where $u =$ \begin{bmatrix} \delta_{DH} \\ \delta_{RL} \\ \delta_{RU} \\ \delta_{sp} \\ \delta_{cv,anti} \end{bmatrix}$.

Actuator models are included and the model is further augmented with additional performance outputs to represent the structural modal velocities $\dot{\eta}_4$ and $\dot{\eta}_5$. Hence, the lateral-directional plant model for synthesis is

$$y = \begin{bmatrix} p \\ r \\ \Delta a_y \\ \dot{\eta}_4 \\ \dot{\eta}_5 \end{bmatrix}, \quad \dot{y}_p = \begin{bmatrix} P \\ \dot{P} \end{bmatrix} \begin{bmatrix} \delta_{DH} \\ \delta_{RL} \\ \delta_{RU} \\ \delta_{sp} \\ \delta_{cv,anti} \end{bmatrix} u.$$ (5.8)

Again, a two-degrees-of-freedom controller as in Equation (5.6) is used, only that this time a reference roll rate command $p_{ref}$ is provided. A load disturbance model with input disturbances at all control surfaces is assumed, i.e., $P_d = P$. The generalized closed-loop interconnection used for the design of the controller is the same as for the longitudinal
5 Application Examples

control problem given in Equation (5.7), only that
\[
S_r = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - S_o P C_{FF} \quad \text{and} \quad D_r^e = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} D_e \begin{bmatrix} 1 \\ 0 \end{bmatrix}.
\]

Following the same procedure as in the longitudinal control problem, \(W_u\) is set to enforce sensitivity reduction up to \(2 \text{ rad/s}\) for roll rate tracking and a constant upper bound of 2 on the sensitivity of \(r\) and \(\Delta a_y\). The weight \(W_u\) is selected to contain constants of 1 for the inputs \(\delta_{RU}, \delta_{RL}\), and \(\delta_{sp}\), as well as integral weights with a bandwidth of \(1 \text{ rad/s}\) for \(\delta_{DH}\) and \(\delta_{cv,anti}\). These integral weights penalize the horizontal tail and control vanes in steady state as both effectors are also used for longitudinal control. Introducing these weights can thus be understood as a “soft” control allocation that prioritizes the horizontal stabilizer and the control vanes for structural control, while attitude control is primarily achieved using the spoilers and rudder.

Tuning is again performed by altering the scalings. The weight \(D_e\) of “allowable maximum errors” is selected such that a reference change of \(10\%\) roll rate is assumed. Relative to this error, the remaining scalings are selected as \(1\%\) yaw rate and \(32.17 \text{ ft/s}^2 (1 \text{ g})\) acceleration. The scaling \(D_u\) is used to express “allowable inputs” of \(10\%\) deflection for horizontal stabilizer and control vanes, \(3^\circ\) deflection for upper and lower rudder, and \(30^\circ\) deflection for the spoilers. Next, the scaling \(D_d\) is increased until robustness margins are satisfactory, leading to values between \(1^\circ\) and \(2^\circ\). Again, the interpretation of actuator uncertainty might be convenient, resulting in \(33\%\) uncertainty for the rudders, \(20\%\) uncertainty for stabilizer and control vanes and \(7\%\) uncertainty for the spoilers. The static weight \(W_p\) is finally selected such that the peaks related to the structural modes are slightly above unity gain and hence contribute to the optimization criterion at that frequency.

The weights are summarized in Table 5.2 using the same parameterization as before, i.e., static weights satisfy \(W = M\) and integral weights satisfy \(W(0) = 10^5, |W(j \omega_b)| = 1\), and \(W(\infty) = 0.5\).

<table>
<thead>
<tr>
<th>Weight on</th>
<th>(D_e)</th>
<th>(D_u)</th>
<th>(D_d)</th>
<th>(\omega_b)</th>
<th>(M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(W_e)</td>
<td>(p)</td>
<td>(10/180 \pi)</td>
<td>—</td>
<td>—</td>
<td>2</td>
</tr>
<tr>
<td>(W_u)</td>
<td>(r)</td>
<td>(1/180 \pi)</td>
<td>—</td>
<td>—</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>(\Delta a_y)</td>
<td>32.17</td>
<td>—</td>
<td>—</td>
<td>0.5</td>
</tr>
<tr>
<td>(W_a)</td>
<td>(\delta_{DH})</td>
<td>—</td>
<td>10</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(\delta_{RL})</td>
<td>—</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(\delta_{RU})</td>
<td>—</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(\delta_{sp})</td>
<td>—</td>
<td>30</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(\delta_{cv,anti})</td>
<td>—</td>
<td>10</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>(W_p)</td>
<td>(\eta_4)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(\eta_5)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 5.2: Weights for Lateral-Directional Design.
The function `lpvsyn` of the LPVTools toolbox is used for a 15% suboptimal controller synthesis. Achievable performance and computational effort of the LPV synthesis depend on the selection of basis functions for the matrix functions $X$ and $Y$ in Theorem 4.2. To quantify the achievable performance, a lower bound is first obtained using the $\mathcal{H}_\infty$ controller synthesis routine `hinfsyn` at each grid point. This lower bound is $\gamma_{LTI} = 1.46$ for the lateral-directional controller and $\gamma_{LTI} = 2.58$ for the longitudinal controller. An LPV synthesis with parameter-independent matrices $X$ and $Y$ results in a relative performance index $\gamma/\gamma_{LTI} = 1.48$ for the lateral directional and $\gamma/\gamma_{LTI} = 1.07$ for the longitudinal case. The latter is already inside the 15% backoff factor of the suboptimal synthesis routine. Hence, no additional improvement can be expected from using parameter-dependent basis functions for the longitudinal control problem. Using affine parameter-dependent matrices improves the value for the lateral-directional controller to 1.13, which also falls into the accepted suboptimality range. It should also be noted that tighter rate bounds (e.g., by a factor of 10) have no noticeable effect on the achievable performance.

The LPV controller is implemented as a piece-wise linearly interpolated state space model with its matrices stored in lookup tables. The lateral-directional controller depends on both parameters and rates as a consequence of the affine basis functions. Not only the synthesis, but also simulations nevertheless indicated that the rate dependence is negligible for all considered scenarios. Thus, the rate dependence is discarded, i.e., the controller is implemented only with data corresponding to zero rates, reducing the required storage capacity by a factor of four.
5 Application Examples

5.4.3 Control Law Verification

For comparison, both a pitch rate and a roll rate tracker are designed using classical loopshaping guidelines. These SISO controllers are concatenated on the existing SAS and SMCS of the B-1 model. Since these gain-scheduled inner loops already take into account variations due to different operating conditions, LTI controllers are deemed sufficient for the purpose.

Linear RobustnessMargins

Before any simulations are performed, linear robustness margins are verified. Disk margins, corresponding to simultaneous gain and phase variations, are calculated both at the physical inputs and outputs of the plant, i.e., with actuator and servo models in the loop and considering the acceleration measurements $a_{\text{cockpit}}$ and $a_{\text{cg}}$ that constitute the feedback signals $\Delta a_*$ as independent. The margins are calculated one-loop-at-a-time using Matlab’s `loopmargin` function for parameters frozen at each grid point and are summarized in Table 5.3. The margins of the integrated designs are similar to those of the reference designs with slightly lower margins for $\delta_{DH}$, $\delta_{sp}$, $\delta_{cv,sym}$, and $a_{y,\text{cockpit}}$, but slightly improved margins for $\delta_{RL}$, $a_{y,\text{cg}}$, $a_{z,\text{cg}}$, and $q$. Overall, all designs can be considered to be fairly robust in this metric with all input and output margins above 6.7 dB (40°).

<table>
<thead>
<tr>
<th>Input</th>
<th>Conventional Design</th>
<th>Integrated Design</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gain Margin</td>
<td>Phase Margin</td>
</tr>
<tr>
<td>$\delta_H$</td>
<td>6.7 dB</td>
<td>40°</td>
</tr>
<tr>
<td>$\delta_{DH}$</td>
<td>9.8 dB</td>
<td>54°</td>
</tr>
<tr>
<td>$\delta_{RL}$</td>
<td>11.2 dB</td>
<td>59°</td>
</tr>
<tr>
<td>$\delta_{RV}$</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>$\delta_p$</td>
<td>14.7 dB</td>
<td>69°</td>
</tr>
<tr>
<td>$\delta_{cv,sym}$</td>
<td>15.2 dB</td>
<td>70°</td>
</tr>
<tr>
<td>$\delta_{cv,anti}$</td>
<td>18.8 dB</td>
<td>76°</td>
</tr>
<tr>
<td>Output</td>
<td>Gain Margin</td>
<td>Phase Margin</td>
</tr>
<tr>
<td>$p$</td>
<td>7.9 dB</td>
<td>46°</td>
</tr>
<tr>
<td>$q$</td>
<td>6.7 dB</td>
<td>40°</td>
</tr>
<tr>
<td>$r$</td>
<td>19.9 dB</td>
<td>78°</td>
</tr>
<tr>
<td>$a_{y,\text{cg}}$</td>
<td>11.5 dB</td>
<td>60°</td>
</tr>
<tr>
<td>$a_{z,\text{cg}}$</td>
<td>11.4 dB</td>
<td>60°</td>
</tr>
<tr>
<td>$a_{y,\text{cockpit}}$</td>
<td>18.9 dB</td>
<td>77°</td>
</tr>
<tr>
<td>$a_{z,\text{cockpit}}$</td>
<td>15.8 dB</td>
<td>71°</td>
</tr>
</tbody>
</table>
Pitch Doublet Simulation

The controllers are first evaluated in nonlinear simulation for a 3°/s pitch doublet maneuver, starting from two different trim conditions: a) Mach 0.6 at 5000 ft and b) Mach 0.8 at 20000 ft altitude. In the first case, full throttle is engaged immediately after the simulation starts. In the second case, throttle is kept at the trim value. These inputs are selected so that parameter variation is caused during the simulation. Figure 5.37 shows the resulting altitude and airspeed variation during the simulation trials.

Figure 5.37: Parameter variation during nonlinear simulation of pitch doublet with conventional controller (---) and integrated controller (----). Grid points are marked .

The goals are to track the pitch rate command and to avoid structural vibrations which cause adverse acceleration at the cockpit. The responses of the aircraft are depicted in Figure 5.38. Both controllers achieve a similar pitch rate and attitude response at both operating points, with the integrated controller providing slightly better short period damping and hence a faster settling time than the classical controller. The adverse acceleration at the cockpit is also significantly reduced, without altering the general acceleration profile which determines the pilot’s perception.

The structural modal velocities caused by the pitch doublet maneuver are depicted in Figure 5.39. The modal velocities are shown in terms of the normalized displacement mode shapes which where introduced in Figure 5.35. For a better comparison, the response with disengaged SMCS is also shown, demonstrating that the conventional SMCS already achieves a large amount of damping augmentation. Still, the integrated controller is able to significantly reduce the deformation velocities and achieve a much faster decay of the oscillations. Hence, the number of load cycles is reduced, which is particularly apparent in the mode represented by the state variable \( \bar{\eta}_3 \). The conventional collocated controller does not at all affect this mode, as was expected from the root locus plot of Figure 5.36. In contrast, the integrated controller is able to improve all structural dynamics.

Figure 5.39 further shows that this improvement is achieved without using additional control effort. In particular, both controllers use the horizontal tail in a very similar way. On the contrary, the control vane deflection is completely different with the integrated
controller, even though similar magnitude is used. In fact, the control signals appear to be almost 180 degrees out of phase, which is a surprising result. This behavior is believed to be attributable to the fact that the conventional controller only acts in response to already existing vibrations at the cockpit, i.e., the pitch rate tracker which controls the horizontal tail and the SMCS which controls the vanes at the cockpit work against each other. On the contrary, the integrated controller uses both available effectors jointly in a coordinated way and anticipates the vibrations that are caused by a pitch demand. It is hence able to avoid structural excitation better.

**Figure 5.38:** Response to pitch doublet command (...) in nonlinear simulation with conventional controller (---) and integrated controller (—).
5.4 Integrated Aeroservoelastic Control

(a) Starting at 5000 ft altitude and Mach 0.6. (b) Starting at 20000 ft altitude and Mach 0.8.

Figure 5.39: Structural modal velocities and control surface deflection during nonlinear simulation of pitch doublet with conventional controller (SMCS engaged /disengaged ) and integrated controller ( ).
Roll Doublet Simulation

The controllers are next evaluated in nonlinear simulation for a bank-to-turn maneuver, initiated by applying a $10^\circ/\text{s}$ roll doublet with a length of 5 seconds. Again, the two different trim conditions: a) Mach 0.6 at 5000 ft and b) Mach 0.8 at 20000 ft altitude are used as starting points. Figure 5.40 shows the resulting altitude and airspeed trajectories.

![Graph showing altitude and airspeed trajectories](image)

Figure 5.40: Parameter variation during nonlinear simulation of roll doublet with conventional controller (—) and integrated controller (—). Grid points are marked ●.

The goals are to track the roll rate command, to provide turn coordination in terms of sideslip regulation, and to reduce structural vibrations. The responses of the aircraft, depicted in Figure 5.41, confirm that both controllers achieve a similar roll rate response at both operating points. The integrated controller reaches the demand value slightly faster than the classical controller and exhibits slightly less roll-to-yaw coupling (visible in the heading angle response). Turn coordination in terms of sideslip regulation is also slightly improved with the integrated controller. Most prominently, the adverse acceleration at the cockpit is drastically reduced.

In Figure 5.42, the structural deformation velocities caused by the roll doublet maneuver are depicted with reference to the normalized mode shapes of Figure 5.35 as generalized coordinates. Again, the conventional controller provides a clear improvement when compared to the simulation trials without active SMCS for the fuselage bending mode $\eta_5$. The wing bending mode $\eta_4$, on the other hand, is not affected at all by the collocated controller, exactly as predicted by the root-locus plot of Figure 5.36. The integrated controller addresses both structural modes simultaneously and provides extra damping augmentation when compared to the conventional controller: The deformation velocities are reduced in magnitude, decay much faster, and the number of load is greatly reduced. These results parallel the observations in the longitudinal control problem.

Figure 5.43 shows the corresponding control surface deflections. Evidently, the integrated controller overall requires less control action during the maneuver, which shows the benefits of jointly using all available effectors. In particular, the deflections of the control vanes and the rudder are significantly reduced in comparison with the conventional
Figure 5.41: Response to roll doublet command (⋯⋯) in nonlinear simulation with conventional controller (——) and integrated controller (—).
controller. The integrated controller also distributes the control effort differently compared to the conventional design. The control vanes are not only deflected less, but also almost completely out of phase with the conventional control signal. This behavior is similar to the longitudinal case. The purpose of the control vanes in the integrated control system appears to be to counterbalance the effect of spoiler deflection. The spoilers are clearly the main effectors for the integrated controller and their deflection exhibits an additional reversal compared to the deflections issued by the classical controller.

Overall, the lateral-directional simulation trials confirm the results from the longitudinal simulation trials. The maneuver characteristics are slightly improved while less control effort is required and all structural modes are effectively damped such that the adverse acceleration at the cockpit is reduced.

Figure 5.42: Structural modal velocities during nonlinear simulation of roll doublet with conventional controller (SMCS engaged / disengaged ) and integrated controller ( ).

(a) Starting at 5000 ft altitude and Mach 0.6.  (b) Starting at 20000 ft altitude and Mach 0.8.
5.4 Integrated Aeroservoelastic Control

(a) Starting at 5000 ft altitude and Mach 0.6.  (b) Starting at 20000 ft altitude and Mach 0.8.

Figure 5.43: Control surface deflections during nonlinear simulation of roll doublet with conventional controller (---) and integrated controller (---).
Simulation with Turbulence

Having examined lateral-directional and longitudinal control at the extreme points of the flight envelope, a final combined simulation trial is performed. The simulation is initialized at 5000 ft altitude and an airspeed of Mach 0.8. At a run time of \( t = 10 \) s a pitch-up maneuver is performed, followed by a roll doublet at \( t = 35 \) s and finally a pitch down maneuver at \( t = 80 \) s. Throttle is adjusted during the simulation to remain inside the flight envelope that is represented by the simulation. The resulting parameter trajectories are depicted in Figure 5.44. Additionally, turbulence is added to evaluate the damping augmentation provided by the controller in response to external disturbances. A standard Dryden gust spectrum is used for the simulation of vertical and lateral turbulence as described by Schmidt [2013].

\[ \begin{align*}
\text{Airspeed (Mach)} & \quad 0.6 \quad 0.7 \quad 0.8 \\
\text{Altitude (ft)} & \quad 5000 \quad 10000 \quad 15000 \quad 20000 \\
\text{Time (s)} & \quad 40 \quad 80 \\
\text{Ma} & \quad 0.6 \quad 0.7 \quad 0.8
\end{align*} \]

Figure 5.44: Parameter variation during nonlinear simulation with turbulence with conventional controller (---) and integrated controller (---). Grid points are marked ●.

Figure 5.45 shows the pitch rate and roll rate response during the simulation. As expected from the previous simulation trials, the conventional and the integrated controller yield very similar results. For both controllers, an undesirable roll-to-pitch coupling is present, owing to the separate design of the longitudinal and lateral control systems.

The structural modal velocities of all five structural modes are depicted in Figure 5.46. The \( \eta_1 \) and \( \eta_5 \) modes, that are well damped via the conventional collocated controller, are similarly apparent with both control systems. The integrated controller can be seen to result in lower deformation velocities during the maneuvers (\( \dot{\eta}_1 \) during pitch, \( \dot{\eta}_5 \) during roll) but in return appears to create a stronger coupling between longitudinal and lateral vibrations as the roll doublet also excites the \( \eta_1 \) mode in this case. The difference for the \( \eta_2 \) mode is marginal, while for the \( \eta_3 \) mode, a large amount of additional damping augmentation is provided by the integrated controller throughout the simulation. The results for the \( \eta_4 \) mode are twofold. During the roll maneuver, vibrations are very effectively suppressed by the integrated controller, but during the rest of the simulation, a slight increase in vibrations can be observed.

The control effort is depicted in Figure 5.47, where the symmetric and antisymmetric
5.4 Integrated Aeroservoelastic Control

deflection signals are combined to show the actual deflection of the individual control surfaces. Apart from the significantly reduced rudder deflection issued by the integrated controller, the control effort of both controllers is similar in magnitude. The integrated controller shows some additional spoiler activity, consistent with the observations from the lateral-directional simulation trial. The horizontal tail deflection is almost identical with both controllers, while the control vane deflection is distinctively different. The largest (symmetric) deflections occur during the roll doublet, which appears to be a consequence of the increased coupling of longitudinal and lateral structural dynamics.

Conclusions

The proposed integrated controller achieves similar robustness margins and tracking performance with less control effort compared to the conventional design. It greatly reduces structural vibrations through a combination of excitation avoidance and active damping augmentation. This was shown in all three considered simulation trials and for different regions of the flight envelope. This additional reduction is much greater when maneuver loads are the cause of structural excitation in comparison to when the excitation is due to external turbulence. The main benefit of an integrated design thus appears to be that it can anticipate structural vibrations due to maneuver demand and consequently avoid them. In addition, the proposed controller is able to affect the structural modes that are not controllable through a colocated design.
Figure 5.46: Structural modal velocities during nonlinear simulation with turbulence with conventional controller (---) and integrated controller (---).
Figure 5.47: Control surface deflections during nonlinear simulation with turbulence with conventional controller (---) and integrated controller (—). The upper rudder deflection is small and not shown.
6 Conclusions and Future Research Opportunities

The present dissertation details the use of robust and LPV control techniques for applications in the field of aeroservoelastic systems. Such systems are characterized by a strong interaction of rigid-body dynamics, structural dynamics, aerodynamics, and feedback control systems. Such interaction is often caused by increased structural flexibility. This flexibility is desirable from a design perspective, e.g., as in high aspect-ratio large-span aircraft for improved fuel efficiency, but its adverse effects on the system dynamics must be compensated by means of feedback control. Aeroservoelastic systems are often modeled with a parametric dependence on the surrounding fluid, e.g., in the form of airspeed. The framework of LPV systems is hence a natural approach to aeroservoelastic control.

Section 2.1 shows how LPV models are obtained by means of a parameter-dependent linearization. Some relevant fundamental properties of LPV models, such as stability and the induced $\mathcal{L}_2$-norm, are compiled in Section 2.2.

LPV models of aeroservoelastic systems are often of high order which prohibits their immediate use for LPV controller synthesis whose computational complexity scales badly with the model's number of state variables. Section 3.1 therefore reviews available model order reduction techniques for LPV systems and concludes that these techniques suffer from the same computational limitations as controller synthesis algorithms. To address this problem, an interpolation method based on local modal decomposition is proposed in Section 3.2. It results in a parameter-varying state space basis for the reduced-order model. Section 3.3 introduces a second novel method: the parameter-varying oblique projection. This subspace method can use local measures of controllability and observability, but nevertheless constructs a consistent, parameter-independent state space basis. The novelty of the method and a major theoretical contribution is the use of a parameter-varying basis for the test space of the approximation, which provides additional freedom in constructing reduced-order models. The projection resembles state-of-the-art balancing and truncation, but is computationally much more efficient. With these model order reduction techniques, it becomes possible to reduce the number of state variables in a model to the point where existing controller synthesis methods are applicable.

Section 4.1 reviews the fundamental architecture of two-degrees-of-freedom control systems and the implications for the sensitivity of the resulting control loops. LPV control builds on norm-optimal robust control theory which formulates objectives in terms of sensitivity redistribution. This loopshaping paradigm is reviewed in-depth in Section 4.2 and a generic mixed sensitivity loopshaping formulation with a manageable number of comprehensible design parameters is proposed. Further, the transition to concrete design guidelines is achieved by providing links between several common design methodologies.
In particular, clarification of the role of scalings and weighting filters in the design process is provided and modal sensitivities are introduced to formulate damping requirements in terms of reducing sensitivity peaks. Section 4.3 collects the relevant synthesis machinery and Section 4.4 concludes with a discussion of the inherent robustness properties of mixed sensitivity controllers.

The last chapter of the dissertation provides several application examples for the techniques which where developed in the previous chapters. Model order reduction for two different industry-grade high-fidelity models of unmanned aircraft is detailed in Section 5.1. Section 5.2 further demonstrates the versatility of the oblique projection method on the large-scale model of far wakes behind a wind turbine using empirical Gramian estimates. The problem of active flutter suppression is investigated in Section 5.3 on the model of the miniMUTT, a small research drone. Envelope expansion into the naturally unstable flight regime is achieved with a significant increase in damping of the aeroelastic oscillations. The fundamental difficulty of controlling high-frequency instabilities such as flutter are highlighted and a particular focus is put on parasitic dynamics of individual components such as actuators and computation devices which aggravate the design. Further, the variety of possible uncertainties in the model is thoroughly discussed and their impact on achievable performance is evaluated. The integration of aeroservoelastic control objectives into a more general flight control system is finally investigated in Section 5.4 on the model of a large flexible aircraft resembling the B-1. Both a complete longitudinal and lateral-directional control system are designed and compared in detail with a conventional cascaded control system which employs dedicated collocated feedback loops for structural mode attenuation. Several nonlinear simulations are performed and the proposed multivariable controller is shown to greatly reduce all structural vibrations, including modes that are not addressable via the conventional controller. These design studies not only substantiate the previously formulated design guidelines but also provide insight into the challenges and limitations that arise in aeroservoelastic control through in-depth evaluations and robustness analyses. They further show that loopshaping is ideally suited for aeroservoelastic control where damping augmentation under large model uncertainty with frequency-limited control activity is the main objective.

Remarks on and Future Research Opportunities in LPV Model Order Reduction

The modal interpolation method is shown to approximate the full-order system locally very well and to indeed result in a smoothly interpolatable state space representation. Nevertheless, the parameter-varying basis of this representation can cause problems as becomes apparent from the unsatisfactory global error bound in the numerical example of Section 3.2.4. This issue provides a major motivation and justification for the development of the parameter-varying oblique projection method with its parameter-independent state space basis and highlights the importance of a suitable representation of the reduced-order model. Another issue with the modal interpolation method is its restriction to a common number of modes in the local models. In direct comparison, the oblique projection methods therefore appears to be superior and, in general, preferable.

The possibility of a partially parameter-varying oblique projection for model order
reduction is believed to be a very general result whose implications might go far beyond
the specific context of Gramian-based definition of subspaces which is in the focus of
Section 3.3. One obvious alternative for a basis space is provided by the Krylov-SVD
method of Gugercin [2005, 2008]. Whether this different basis space would provide any
benefits over the (approximate) balanced coordinates which are proposed in this thesis is
up to debate and remains to be investigated. A possible way to make unstable models
amenable to the proposed method is to use local coprime factorization Gramians instead
of local Gramians for the definition of test space and basis space. Doing so simply requires
to replace the solutions of the Lyapunov equations (3.25) by the solutions of the Riccati
equations (3.27). It remains to investigate whether it is possible to incorporate (at least in
parts) the possibility to perform modal-like decompositions by means of a suitable choice
of basis and test spaces.

In many applications, residualization is preferred over truncation to preserve the steady
state gain of the original model. It is therefore a natural question whether the parameter-
varying oblique projection can be restated in terms of a parameter-varying residualization.
This question is surprisingly complex to answer. Residualization would require a transfor-
mation, which inherently has no nullspace. As the parameter-varying oblique projection
uses a varying nullspace, it appears to be impossible to construct such a transformation
while maintaining the additional freedom of the proposed method. In fact, it is readily
verified that the projection (3.44) simply becomes the identity map $I_{n_y}$, when the basis
space has dimension $n_y$, i.e., when the basis $V$ has full row rank as would have to be
the case for any nonsingular transformation. However, when taking the view point of an
oblique projection as a transformation followed by truncation, the proposed parameter-
varying oblique projection can be given the interpretation of a transformation which is
restricted precisely in such a way that parameter dependence only occurs in the null
space of the subsequent truncation. It therefore appears to be in principal possible to
construct such a transformation and then apply residualization instead of truncation
on the state variables associated with the parameter-varying part of the basis. Another
trail of thoughts that leads to this particular problem is related to the fact that, while
it is impossible to express residualization as a projection, there exist projections that
also preserve the steady-state gain of the original system, e.g., by projecting the system
onto $\text{span}([B \ A^{-1}B \cdots A^{-n_y+1}B])$ as is done in moment matching at zero frequency. The
extension of the proposed oblique projection method to obtain reduced-order models with
preserved steady state gain therefore certainly deserves more attention and it remains to
investigate whether useful connections can be established.

Finally, it should be noted that this dissertation is concerned with removing weakly
controllable and observable state variables from the model, which excludes the case of
strictly uncontrollable or unobservable state variables. Again, it is surprising how severe the
implications of this restriction are, as the existence of not only the generalized Gramians
but also of the projection (3.44) depends on observability and controllability of the full-
order system. The problem of obtaining such minimal realizations for general gridded
LPV systems appears to have not been addressed in the literature so that practitioners
are required to rely on physical understanding of the model to remove unnecessary state
variables before any model order reduction methods are applied.
Remarks on and Future Research Opportunities in Aeroservoelastic Control

The proposed parameterization of the mixed sensitivity design problem has not only very appealing properties in terms of intuitive and comprehensible design parameters that can be selected by the designer. It might also prove very useful for automated tuning approaches that employ nonlinear optimization to select the “weights” as is frequently done for industry-grade applications. The parameterization of such problems plays an important role, as the optimization algorithm usually tries out a large number of combinations of weights through pattern search or genetic algorithms. In a standard setup, many of these trial cases can become ill-posed problems. With the proposed parameterization where only the scalings are considered to be actual tuning knobs, it should be possible to significantly reduce the computational load and hence time required to tune controllers. Comparative design studies should be performed to evaluate potential benefits and drawbacks of the proposed parameterization.

In Section 4.2, modal sensitivities are introduced to express damping requirements. Their definition requires a modal transformation and hence the model is restricted to be LTI. For LPV models, a remedy is to apply the “modal matching” algorithm, introduced in Section 3.2 for the purpose of model order reduction, in order to construct an approximate modal transformation. Since the algorithm also makes use of mode-wise canonical forms, an output matrix which defines the modal sensitivities can be obtained exactly as in Equation (4.17). It is essential to note that if this is done, the augmented system is still in the exact same state space coordinates as the original model. That is, no actual state transformation is applied and therefore the state consistency of an LPV model is not at stake. Alternatively, surrogate outputs can be defined as was done in both control design studies in Sections 5.3 and 5.4. In these models, the structural modal velocities appear explicitly which makes them obvious choices. The essence of the modal sensitivities is their band-limited frequency response with a characteristic peak at the frequency where lightly-damped oscillations occur.

Modal sensitivities can effectively be used to provide damping to modes that are not adequately addressable through collocated feedback loops as is shown on the example of the B-1 model in Section 5.4. It is of interest to recall that the proposed controller has, in summary, only access to the same measurements and effectors as the conventional control system. It hence internally estimates “modal states” in order to improve damping. This observation provides an interesting link to control techniques that seek to explicitly provide estimates of modal states and then use them in a state feedback fashion to augment damping [e.g., Danowsky et al. 2013] and this connection deserves to further be explored.

Finally, it remains to report that on June 2\textsuperscript{nd} 2017, a controller based on the design of Section 5.3 was successfully flight tested on the miniMUTT for an airspeed approximately 5 m/s below the estimated flutter speed [Kotikalpudi, Danowsky, Schmidt, Theis & Seiler 2018]. It was shown to improve the damping ratio of the aeroelastic mode by a factor of 1.5–2 [ibid]. This first flight test provides evidence that the methods which where developed in this thesis are useful for actual applications.
Bibliography


[Balas & Seiler 2014] Balas, Gary; Seiler, Peter: Multivariable Robust Control. Unpublished Lecture Notes, University of Minnesota, Minneapolis, 2014


Bibliography


Bibliography


[Theis et al. 2015c] Theis, Julian; Takarics, Béla; Pfifer, Harald; Balas, Gary; Werner, Herbert: Modal Matching for LPV Model Reduction of Aeroservoelastic Vehicles. AIAA Atmospheric Flight Mechanics Conference, 2015, Paper No. 2015-1686


Bibliography


List of Publications

Lead-authored


List of Publications

Co-authored

[10] **Kotikalpudi, Aditya** ; **Danowsky, Brian** ; **Schmidt, David** ; **Theis, Julian** ; **Seiler, Peter**: Flutter Suppression Control Design for a Small, Flexible Flying-Wing Aircraft. *AIAA Aviation Forum*, 2018 (to appear)


Curriculum Vitae

Julian Theis  born 29.07.1986 in Hamburg

**Employment History**

<table>
<thead>
<tr>
<th>Since</th>
<th>Position</th>
<th>Institute/Department</th>
</tr>
</thead>
<tbody>
<tr>
<td>08/2017</td>
<td>Post-Doctoral Research Associate</td>
<td>Institute of Aircraft Systems Engineering, Hamburg U. of T.</td>
</tr>
<tr>
<td>06/2016–05/2017</td>
<td>Research Assistant</td>
<td>Institute of Control Systems, Hamburg U. of T.</td>
</tr>
<tr>
<td>06/2015–05/2016</td>
<td>Research Fellow</td>
<td>Aerospace Engineering &amp; Mechanics Dept., Minneapolis</td>
</tr>
<tr>
<td>05/2013–04/2015</td>
<td>Research Assistant</td>
<td>Institute of Control Systems, Hamburg U. of T.</td>
</tr>
</tbody>
</table>

**Education**

<table>
<thead>
<tr>
<th>Date</th>
<th>Studies</th>
<th>Institution/Dept.</th>
</tr>
</thead>
<tbody>
<tr>
<td>05/2013–07/2017</td>
<td>Doctoral Studies in Control Engineering supervised by Prof. Herbert Werner</td>
<td>Hamburg U. of T.</td>
</tr>
<tr>
<td>06/2006</td>
<td>Certificate of General Qualification for University Entrance</td>
<td>Herzog-Ernst-Gymnasium, Uelzen</td>
</tr>
</tbody>
</table>

**Research Stays**

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Institute/Dept.</th>
</tr>
</thead>
<tbody>
<tr>
<td>06/2015–05/2016</td>
<td>University of Minnesota, Minneapolis</td>
<td>Aerospace Engineering &amp; Mechanics Dept. with Prof. Peter Seiler</td>
</tr>
<tr>
<td>04/2014–05/2014</td>
<td>University of Minnesota, Minneapolis</td>
<td>Aerospace Engineering &amp; Mechanics Dept. with Prof. Gary Balas</td>
</tr>
<tr>
<td>10/2012–04/2013</td>
<td>German Aerospace Center (DLR), Oberpfaffenhofen</td>
<td>Institute of System Dynamics and Control</td>
</tr>
<tr>
<td>08/2011–03/2012</td>
<td>University of California, Berkeley</td>
<td>Mechanical Engineering Dept. with Prof. Andrew Packard</td>
</tr>
</tbody>
</table>