Frequency-limited model approximation of large-scale dynamical models

P. Vuillemin

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Institut Supérieur de l'Aéronautique et de l'Espace (ISAE)*

Présentée et soutenue le *24/11/2014* **par :**

Pierre Vuillemin

Approximation de modèles dynamiques de grande dimension sur intervalles de fréquences limités

---

JURY

M. Olivi	Chargée de Recherche	INRIA	Rapporteur
S. Gugercin	Professeur	Virginia Tech	Rapporteur
D. Henrion	Directeur de Recherche	CNRS	Examinateur
M. Zasadzinski	Professeur	Université de Lorraine	Examinateur
D. Alazard	Professeur	Université de Toulouse	Directeur de thèse
C. Poussot-Vassal	Ingénieur de Recherche	Onera	Co-directeur de thèse

---

École doctorale et spécialité :
EDSYS : Automatique 4200046

Unité de Recherche :
Onera – The French Aerospace Lab, F-31055 Toulouse, France

Directeurs de Thèse :
Daniel Alazard et *Charles Poussot-Vassal*

Rapporteurs :
Martine Olivi et *Serkan Gugercin*

Frequency-limited model approximation of large-scale dynamical models

Pierre Vuillemin

À mes parents

Acknowledgement

Firstly, I would like to express my warm thanks to my thesis advisors Charles Poussot-Vassal and Daniel Alazard for giving me the opportunity to achieve this PhD. thesis. I am also thankful for their guidance, constructive criticism and advice which have enabled me to successfully complete this project.

I would also like to thank Martine Olivi and Serkan Gugercin for accepting to review this manuscript. Their constructive comments have given me the key elements to improve this work and various leads to extend it. Likewise, I would like to thank Didier Henrion and Michel Zasadzinski for accepting to be part of my jury and for their interesting feedback.

I am also grateful to all the people at Onera, and especially within the Flight Dynamics and Systems Control department, who have helped me during my thesis either for technical matters or administrative tasks.

Special thanks to my fellow Phd. students of the department who have helped make these three years a very enjoyable and agreeable period through the numerous coffee breaks and other activities.

I could not end without expressing my gratitude to Marion for her foolproof support during these three years which has enabled me to be fully focused on my work.

Résumé

Les systèmes physiques sont représentés par des modèles mathématiques qui peuvent être utilisés pour simuler, analyser ou contrôler ces systèmes. Selon la complexité du système qu'il est censé représenter, un modèle peut être plus ou moins complexe. Une complexité trop grande peut s'avérer problématique en pratique du fait des limitations de puissance de calcul et de mémoire des ordinateurs. L'une des façons de contourner ce problème consiste à utiliser *l'approximation de modèles* qui vise à remplacer le modèle complexe par un modèle *simplifié* dont le comportement est toujours représentatif de celui du système physique.

Dans le cas des modèles dynamiques Linéaires et Invariants dans le Temps (LTI), la complexité se traduit par une dimension importante du vecteur d'état et on parle alors de *modèles de grande dimension*. L'approximation de modèle, encore appelée *réduction de modèle* dans ce cas, a pour but de trouver un modèle dont le vecteur d'état est plus petit que celui du modèle de grande dimension tel que les comportements entrée-sortie des deux modèles soient proches selon une certaine norme. La norme $\mathcal{H}_2$ a été largement considérée dans la littérature pour mesurer la qualité d'un modèle réduit. Cependant, la bande passante limitée des capteurs et des actionneurs ainsi que le fait qu'un modèle est généralement représentatif d'un système physique dans une certaine bande fréquentielle seulement, laissent penser qu'un modèle réduit dont le comportement est fidèle au modèle de grande dimension dans un intervalle de fréquences donné, peut être plus pertinent. C'est pourquoi, dans cette étude, *la norme $\mathcal{H}_2$ limitée en fréquence*, ou norme $\mathcal{H}_{2,\Omega}$, qui est simplement la restriction de la norme $\mathcal{H}_2$ sur un intervalle de fréquences Ω, a été considérée. En particulier, le problème qui vise à trouver un modèle réduit minimisant la norme $\mathcal{H}_{2,\Omega}$ de l'erreur d'approximation avec le modèle de grande dimension a été traité.

Deux approches ont été proposées dans cette optique. La première est une approche empirique basée sur la modification d'une méthode sous-optimale pour l'approximation $\mathcal{H}_2$. En pratique, ses performances s'avèrent intéressantes et rivalisent avec certaines méthodes connues pour l'approximation de modèles sur intervalles de fréquences limités.

La seconde est une méthode d'optimisation basée sur la formulation pôles-résidus de la norme $\mathcal{H}_{2,\Omega}$. Cette formulation généralise naturellement celle existante pour la norme $\mathcal{H}_2$ et permet également d'exprimer deux bornes supérieures sur la norme $\mathcal{H}_\infty$ d'un modèle LTI, ce qui est particulièrement intéressant dans le cadre de la réduction de modèles. Les conditions d'optimalité du premier ordre pour le problème d'approximation optimale en norme $\mathcal{H}_{2,\Omega}$ ont été exprimées et utilisées pour créer un algorithme de descente visant à trouver un minimum local au problème d'approximation. Couplée aux bornes sur la norme $\mathcal{H}_\infty$ de l'erreur d'approximation, cette méthode est utilisée pour le contrôle de modèles de grande dimension.

D'un point de vue plus pratique, l'ensemble des méthodes proposées dans cette étude ont été appliquées, avec succès, dans un cadre industriel comme élément d'un processus global visant à contrôler un avion civil flexible.

Mots-clés : modèles linéaires invariants dans le temps, modèles de grande dimension, réduction de modèles, approximation de modèles, approximation de modèles sur intervalles de fréquences limités

Abstract

Physical systems are represented by mathematical models in order to be simulated, analysed or controlled. Depending on the complexity of the physical system it is meant to represent and the way it has been built, a model can have a varying complexity. This complexity can become an issue in practice due to the limited computational power and memory of computers. One way to alleviate this issue consists in using *model approximation* which is aimed at finding a *simpler* model that still represents faithfully the physical system.

In the case of Linear Time Invariant (LTI) dynamical models, complexity translates into a large dimension of the state vector and one talks about *large-scale models*. Model approximation is in this case also called *model reduction* and consists in finding a model with a smaller state vector such that the input-to-output behaviours of both models are close with respect to some measure. The $\mathcal{H}_2$-norm has been extensively used in the literature to evaluate the quality of a reduced-order model. Yet, due to the limited bandwidth of actuators, sensors and the fact that models are generally representative on a bounded frequency interval only, a reduced-order model that faithfully reproduces the behaviour of the large-scale one over a bounded frequency interval only, may be more relevant. That is why, in this thesis, *the frequency-limited* $\mathcal{H}_2$-norm, or $\mathcal{H}_{2,\Omega}$-norm, which is the restriction of the $\mathcal{H}_2$-norm over a frequency interval Ω, has been considered. In particular, the problem of finding a reduced-order model that minimises the $\mathcal{H}_{2,\Omega}$-norm of the approximation error with the large-scale model has been addressed here.

For that purpose, two approaches have been developed. The first one is an empirical approach based on the modification of a sub-optimal $\mathcal{H}_2$ model approximation method. Its performances are interesting in practice and compete with some well-known frequency-limited approximation methods.

The second one is an optimisation method relying on the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm. This formulation naturally extends the one existing for the $\mathcal{H}_2$-norm and can also be used to derive two upper bounds on the $\mathcal{H}_\infty$-norm of LTI dynamical models which is of particular interest in model reduction. The first-order optimality conditions of the optimal $\mathcal{H}_{2,\Omega}$ approximation problem are derived and used to built a complex-domain descent algorithm aimed at finding a local minimum of the problem. Together with the $\mathcal{H}_\infty$ bounds on the approximation error, this approach is used to perform control of large-scale models.

From a practical point of view, the methods proposed in this study have been successfully applied in an industrial context as a part of the global process aimed at controlling a flexible civilian aircraft.

Keywords : linear time invariant models, large-scale models, model reduction, model approximation, frequency-limited model approximation

Notations and acronyms

Mathematical Notations

j	the square root of -1
$\mathbf{Re}(z)$	the real part of the complex z
$\mathbf{Im}(z)$	the imaginary part of the complex z
M^T	the transpose of M
M^*	the conjugate of M
M^H	the conjugate transpose of M
$[M]_{i,k}$	the element of M located at the i-th row and k-th column
$\lambda_i(M)$	the i-the eigenvalue of M
$\mathbf{tr}\,(M)$	the trace of M
I_n	the identity matrix of size n
$\mathbf{e_i}$	the i-th column canonical vector, *i.e.* the vector with zero entries except for a 1 at the i-th position
$M \odot N$	the Hadamard product, or element-wise product, between M and N
$\mathbf{diag}(M)$	the column vector containing the diagonal of the square matrix M
$\mathbb{1}_n$	a column vector of ones of size n
$\mathbf{a}_{\omega,\lambda}$	denotes $\frac{2}{\pi}\mathbf{atan}(\frac{\omega}{\lambda})$
$\mathbf{vec}(M)$	denotes the vectorisation of the matrix M, *i.e.* the column vector formed by vertically concatenating the columns of M
$\mathcal{F}_l()$	represents the lower Linear Fractional Representation
$\mathcal{F}_u()$	represents the upper Linear Fractional Representation

Acronyms

BT	Balanced Truncation
FW-BT	Frequency-Weighted Balanced Truncation
FL-BT	Frequency-Limited Balanced Truncation
IRKA	Iterative Rational Krylov Algorithm, *also called* Iterative Tangential Interpolation Algorithm (**ITIA**)
ISRKA	Iterative SVD-Rational Krylov Algorithm, *also called* Iterative SVD-Tangential Interpolation Algorithm (**ISTIA**)
FL-ISTIA	Frequency-Limited SVD Tangential Interpolation Algorithm
DARPO	Descent Algorithm for Residues and Poles Optimisation
LPV	Linear Parameter Varying
LFR	Linear Fractional Representation
SISO	Single Input Single Output
SIMO	Single Input Multiple Output
MISO	Multiple Input Single Output
MIMO	Multiple Input Multiple Output

Contents

Acknowledgement i

Résumé iii

Abstract v

Notations and acronyms vii

I Introduction 1

1 Introduction to model approximation 3
- 1.1 Context and motivations . . . 3
- 1.2 Motivating examples . . . 4
  - 1.2.1 Simulation of a 3D cantilever Timoshenko beam . . . 4
  - 1.2.2 Control of an industrial aircraft . . . 6
  - 1.2.3 Standard benchmarks . . . 8
- 1.3 Problem formulation . . . 10
- 1.4 Overview of the contributions . . . 10
- 1.5 Manuscript overview . . . 11

II State of the art 15

2 Preliminary in LTI systems theory 17
- 2.1 Generalities . . . 17
  - 2.1.1 Representation of LTI dynamical models . . . 17
  - 2.1.2 Gramians and balanced realisation . . . 20
- 2.2 Norms of systems . . . 25
  - 2.2.1 $\mathcal{H}_2$-norm . . . 25
  - 2.2.2 Frequency-limited $\mathcal{H}_2$-norm . . . 26
  - 2.2.3 $\mathcal{H}_\infty$-norm . . . 29

3 Standard model approximation methods 31
- 3.1 Model approximation by truncation . . . 31
  - 3.1.1 Truncation and residualisation of state-space representation . . . 31
  - 3.1.2 Modal and balanced truncation . . . 32
- 3.2 Model approximation by moment matching . . . 38
  - 3.2.1 Moment matching problem . . . 39
  - 3.2.2 Implicit moment matching : the SISO case . . . 40
  - 3.2.3 Tangential interpolation . . . 43

4 Optimal $\mathcal{H}_2$ model approximation 47
- 4.1 First order optimality conditions . . . 48
  - 4.1.1 $\mathcal{H}_2$ approximation error . . . 48
  - 4.1.2 Formulation of the first-order optimality conditions . . . 49
- 4.2 Algorithms for optimal $\mathcal{H}_2$ approximation . . . 54
  - 4.2.1 Iterative Rational Krylov Algorithm (IRKA) . . . 54
  - 4.2.2 Iterative SVD Rational Krylov Algorithm (ISRKA) . . . 58
  - 4.2.3 Optimisation algorithm for optimal $\mathcal{H}_2$ model approximation . . . 59

5 Frequency weighted and frequency-limited model approximation 63
5.1 Frequency weighted model approximation . . . 63
5.1.1 Frequency weighted balanced truncation . . . 63
5.1.2 Frequency weighted $\mathcal{H}_2$ model approximation . . . 65
5.2 Frequency-limited model approximation . . . 67
5.2.1 Frequency-limited balanced truncation . . . 67
5.2.2 Gramian-based $\mathcal{H}_{2,\Omega}$ optimal model approximation . . . 69

III Frequency-limited approximation of linear dynamical models 73

6 Development of a first approach for frequency-limited model approximation 75
6.1 Modification of **ISRKA** . . . 75
6.1.1 Presentation of the method . . . 75
6.1.2 Properties . . . 77
6.1.3 Numerical improvement of the method . . . 78
6.2 Numerical illustrations . . . 79
6.2.1 Comparison with the standard version of **ISRKA** . . . 79
6.2.2 Impact of restart and error watching . . . 83
6.2.3 Comparison with other methods . . . 86

7 Formulation of the $\mathcal{H}_{2,\Omega}$-norm with the poles and residues of the transfer function 89
7.1 Models with semi-simple poles only . . . 89
7.1.1 Preliminary results on complex functions . . . 90
7.1.2 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm . . . 91
7.1.3 Numerical illustration of the formulation . . . 93
7.2 Models with higher order poles . . . 94
7.2.1 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm for models with high order poles . . . 95
7.2.2 Special case 1 : n eigenvalues of multiplicity 1 (semi-simple case) . . . 97
7.2.3 Special case 2 : 1 eigenvalue of multiplicity n . . . 97
7.3 Upper bounds of the $\mathcal{H}_\infty$-norm . . . 98
7.3.1 Formulation of the bounds . . . 98
7.3.2 Computation of the bounds . . . 101
7.3.3 Experimental study of the bounds quality . . . 102
7.3.4 Construction of a frequency template . . . 105

8 Formulation of the $\mathcal{H}_{2,\Omega}$ error and differentiation of its gradient 111
8.1 Expression of the approximation error with the poles and residues of the models 111
8.1.1 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$ approximation error . . . 111
8.1.2 On the computation of the approximation error . . . 117
8.2 Gradient of the error . . . 119
8.2.1 Reminder on Wirtinger Calculus . . . 119
8.2.2 Gradient of the $\mathcal{H}_{2,\Omega}$ approximation error . . . 122

9 Development of a descent algorithm for the optimal $\mathcal{H}_{2,\Omega}$ approximation problem 129
9.1 A descent algorithm in the complex domain . . . 129
9.1.1 Reminder of unconstrained optimisation . . . 130
9.1.2 Descent Algorithm for Residues and Poles Optimisation . . . 133
9.1.3 Initialisation of **DARPO** . . . 137
9.2 Numerical illustration . . . 139

9.2.1 Illustration of the behaviour of **DARPO** . . . 139
9.2.2 Improvement provided by **DARPO** . . . 146
9.2.3 Application for the control of a large-scale model . . . 151

10 Industrial aeronautical use case **159**
10.1 Vibration control for one business jet aircraft model . . . 160
10.1.1 General industrial framework . . . 160
10.1.2 Problem formulation and business jet aircraft model approximation . . . 161
10.1.3 Anti-vibration control design . . . 162
10.1.4 Numerical results . . . 166
10.2 Global stability validation of a parameter varying business jet aircraft . . . 167
10.2.1 Problem statement . . . 167
10.2.2 Model approximation & error bounding . . . 168
10.2.3 Low-order LFR model creation . . . 170
10.2.4 IQC-based global stability analysis . . . 173

IV Conclusion **177**

11 Discussion **179**
11.1 **FL-ISTIA** . . . 179
11.2 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm . . . 180
11.2.1 Numerical robustness . . . 180
11.2.2 Scalability . . . 182
11.3 Bounds on the $\mathcal{H}_\infty$-norm and its applications . . . 183
11.3.1 Improvement of the construction of frequency templates . . . 183
11.3.2 Other possible applications for the $\mathcal{H}_\infty$ bound Γ_Ω . . . 185
11.4 First-order optimality conditions for the $\mathcal{H}_{2,\Omega}$ approximation problem . . . 185
11.5 **DARPO** . . . 186
11.5.1 Issues related to the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm . . . 186
11.5.2 Optimisation scheme . . . 188
11.5.3 Approximation for a fixed error . . . 189
11.5.4 Model approximation for the control of large-scale models . . . 190

12 Outlook **193**

V Appendix **195**

A Proofs related to the poles-residues expression of the $\mathcal{H}_{2,\Omega}$-norm **197**
A.1 Poles-residues expression of the $\mathcal{H}_{2,\Omega}$-norm in the semi-simple case . . . 197
A.2 Poles-residues expression of the $\mathcal{H}_{2,\Omega}$-norm for models with higher order poles . . . 199
A.3 Poles-residues expression of the $\mathcal{H}_{2,\Omega}$ approximation error . . . 200

B Elements about the principal value of the complex arctangent **201**
B.1 Difference between the two definitions of the inverse tangent function . . . 201
B.2 Limit of **atan** $\left(\frac{\omega}{\lambda}\right)$ as ω tends towards infinity . . . 201

C Code samples **205**
C.1 $\mathcal{H}_{2,\Omega}$-norm computation . . . 205
C.2 Example 11 : non-linearity of the optimal $\mathcal{H}_2$ approximation problem . . . 206
C.3 Use of the **FL-ISTIA** . . . 206
C.4 Example 15 : computation of the bounds on the $\mathcal{H}_\infty$-norm . . . 207

C.5 Example 16 : construction of a frequency template 208
C.6 Example 17 : behaviour of the $\mathcal{H}_{2,\Omega}$ approximation error when approximating to a second-order model 208
C.7 Use of **DARPO** 210
C.8 Use of **DARPO** for the control of a large-scale model 210
C.9 Use of **DARPO** for the approximation at a fixed error 213

Part I

Introduction

Chapter 1

Introduction to model approximation

1.1 Context and motivations

Physical systems or phenomena are generally represented by mathematical models in order to be simulated, analysed or controlled. Depending on *(i)* the complexity of the physical system to be modelled, *(ii)* the means used to build the mathematical model and *(iii)* the desired accuracy of the model, this model can be more or less complex and representative.

A very accurate model is intuitively desirable, however this often yields an important complexity which might not be tractable in practice. Indeed, the limited computational power of computers and their limited storage capabilities might make some numerical tools not tractable or not in an acceptable time. In addition, the errors induced by the floating point arithmetic might significantly perturb some theoretical results. Hence, the complexity of the model has to be restrained.

It may seem adequate to restrain the complexity of a model from the very beginning, *i.e.* , during its construction, however it is not always possible nor suitable. Indeed, on the one hand numerical modelling tools (such as identification methods, finite elements methods, etc.) do not necessarily enable to restrain the complexity without loosing too much information. And on the other hand, models can be used for different purposes (simulation, control, analysis, etc.) which limitations with respect to the complexity vary. Hence, a standard process in industrial settings consists in creating one single high-fidelity model which is then transformed to match the application. Thus, an a posteriori method is generally preferred to diminish the complexity of the original, high-fidelity model. *Model approximation* serves this purpose. Indeed, the underlying idea behind model approximation consists in replacing a complex model by a simpler one which preserves its main characteristics and which is suitable for a specific application (*e.g.* simulation, analysis or control).

In this study, dynamical systems are considered. They can be represented by finite difference equations, ordinary differential equations, differential algebraic equations or partial differential equations which can be linear, non-linear, time-invariant or time-variant. Linear Time Invariant (LTI) models are widely used, both in the industry and in research. Indeed, for many physical systems, they are sufficiently representative around an equilibrium point and numerous tools exist in order to analyse and control them. For this kind of models, complexity results in a large state-space vector and one talks of *large-scale model*.

Approximation of large-scale LTI models has been extensively studied over the years and two main steps can be distinguished. Initially, some well-known methods such as the Balanced Truncation and the Hankel norm approximation have been developed [Moore, 1981; Glover, 1984]. Then, the extensive use of numerical modelling tools has led to modify the conception of *large-scale* models which can now have thousands or even millions of states. Standard model approximation methods (in their basic form) were not adapted anymore for very large case due to their inherent numerical complexity. Hence, original techniques that are numerically cheaper have been developed. They are mainly based on interpolation through Krylov subspaces [Grimme, 1997] and have more recently led to some interesting development concerning optimal $\mathcal{H}_2$ model approximation [Gugercin et al., 2008; Van Dooren et al., 2008b]. This problem has also been addressed in a different way using non-linear optimisation schemes [Marmorat et al.,

2002; Beattie and Gugercin, 2009]. Initially, these methods were developed to reproduce the behaviour of the large-scale model over the whole frequency range. However, in some cases, it may seem more relevant to match the behaviour of the large-scale model only on a bounded frequency range.

Indeed *(i)* the limited bandwidth of sensors used to identify models from measured data might lead to an inaccurate (and thus irrelevant) model at some frequencies, *(ii)* similarly, actuators cannot act on some dynamics which make them less important for control purpose and *(iii)* by avoiding to take into accounts some dynamics, the large-scale model may be reduced even more without loosing accuracy in the considered frequency band. That is why we have chosen in this thesis to address the problem of approximation over a bounded frequency range.

The most intuitive way to address such a problem consists in applying *frequency filters* to the large-scale model so that the reduced-order model is built in order to match the *filtered* large-scale model. This is called the *frequency-weighted approximation problem* and several methods have been developed to solve it [Enns, 1984; Lin and Chiu, 1990; Zhou, 1995; Leblond and Olivi, 1998; Wang et al., 1999; Breiten et al., 2014]. Note that the filtered large-scale model is augmented by the order of the filters. Hence, and even if it is often marginal compared to the order of the model, the complexity is increased for the approximation algorithms. More importantly, the design of adequate filters can be a tedious task. An approach that does not involve any explicit filtering, called the *frequency-limited balanced truncation*, has been proposed in [Gawronski and Juang, 1990]. It is based on frequency-limited gramians which act as frequency-weighted gramians [Enns, 1984] considered with perfect filters.

This thesis tries to bring together the methodology used in optimal $\mathcal{H}_2$ model approximation [Beattie and Gugercin, 2009] with the criterion derived from frequency-limited gramians [Gawronski and Juang, 1990] in order to perform *optimal frequency-limited $\mathcal{H}_2$ model approximation of large-scale LTI dynamical models*. Note that a similar study has been conducted in parallel of this work in [Petersson, 2013] with another formulation of the criterion.

From a practical point of view, the methods developed during this study are meant to be applied in an industrial aeronautic context which problematic are quite specific as presented below in the second motivating example.

1.2 Motivating examples

In Section 1.2.1, the benefits of model approximation for the simulation[1] of LTI models is highlighted with a 3D cantilever Timoshenko beam model. Then in Section 1.2.2, some issues encountered in the context of industrial aircraft control are introduced. These issues form the industrial application of this thesis and are addressed more in depth in Chapter 10. Finally, in Section 1.2.3, the main models used as academic benchmarks in this thesis are introduced.

1.2.1 Simulation of a 3D cantilever Timoshenko beam

In [Panzer et al., 2009], a simple finite elements method to model a 3D Cantilever Timoshenko beam (see Figure 1.1) loaded by some force $F(t)$ at its tip is described. The provided script enables to arbitrarily choose some parameters as the number of elements N used for discretization, the length of the beam L, etc. Hence it can conveniently be used to build state-space models of

[1] The time measurement presented throughout this study are mainly aimed at being compared to each other and should not be considered as absolute references but for sake of clarity, the technical details related to the numerical computations are presented here.

All the numerical computations have been performed with Matlab® R2013a on an Intel® Core™ i7-3610QM processor under a Linux environment and with 8Gb of RAM.

In Matlab®, the Just In Time (JIT) compiler has been disabled in order to have stable time measurements. The elapsed time is measured through the CPU time.

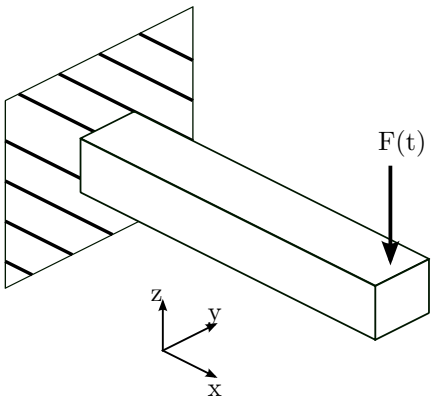


Figure 1.1: 3D Cantilever Timoshenko beam.

Model	Order	T_i (s)	T_f (s)
H_{large}	720	9.86	0.91
H_{small}	12	0.028	0.012
$H_{reduced}$	12	0.018	0.011

Table 1.1: Mean computation times of the impulse responses (T_i) and of the frequency responses (T_f) for the different models.

varying complexity. Note that by default, the output of the system is the vertical (along z axis) displacement at the tip of the beam.

In this example, three models are considered :

- $\mathbf{H}_{large}$: a high fidelity model built with $N = 60$ elements. It has $n_l = 720$ states.
- $\mathbf{H}_{small}$: a model built with $N = 1$ element. It has $n_s = 12$ states.
- $\mathbf{H}_{reduced}$: an approximation of $\mathbf{H}_{large}$[2]. It has also $r = 12$ states.

To evaluate the time taken to simulate each model, the impulse responses over 10s and the frequency responses over 100 points linearly spaced between 10^{-2} and 10^{2} rad/s are computed 50 times for each model. The mean times T_i and T_f for computing the impulse and frequency responses, respectively, are reported in Table 1.1.

As expected, both low-order models $\mathbf{H}_{small}$ and $\mathbf{H}_{reduced}$ have similar simulation times. However, the difference appears clearly with the large-scale model $\mathbf{H}_{large}$ which temporal and frequency simulations take respectively about 300 and 90 times more time than for the low-order models.

The low-order models $\mathbf{H}_{small}$ and $\mathbf{H}_{reduced}$ have the same order but they are not identical. Indeed, by considering the large-scale model $\mathbf{H}_{large}$ as a reference, the low-order model obtained by approximation is more accurate than the one obtained by limiting the number of elements in the finite elements method. This is illustrated in Figure 1.2, where the impulse responses of the different models as well as the impulse responses of the error models $\mathbf{H}_{large} - \mathbf{H}_{reduced}$ and $\mathbf{H}_{large} - \mathbf{H}_{small}$ are plotted on the left and right, respectively. The response of $\mathbf{H}_{large}$ is better matched by the response of $\mathbf{H}_{reduced}$ than the response of $\mathbf{H}_{small}$. This illustrate the fact that the complexity cannot necessarily be restrained directly when building the model without loosing some accuracy.

[2]The reduced-order model is obtained with **DARPO**, the main method developed during this thesis which is detailled in Chapter 9.

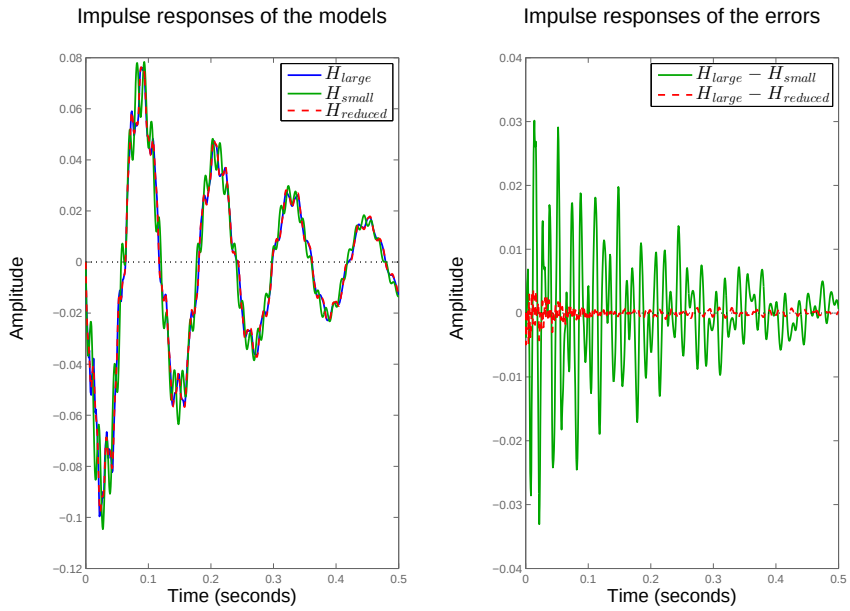


Figure 1.2: Impulse responses of the models (left) and of the errors (right).

Note that in that case the simulation times of the large model $\mathbf{H}_{large}$ is not unacceptable. However, for very large scale models, such as models obtained by discretisation of partial differential equations, it might be the case [Li and Kamon, 2005; Lienemann et al., 2006; Dergham et al., 2011]. Besides, some tasks are far more demanding in terms of computational power and memory requirement than simulation, for instance, even with a model like $\mathbf{H}_{large}$, modern robust, optimal or predictive control tools [Zhou et al., 1995; Scherer et al., 1997; Bemporad and Morari, 1999; Apkarian and Noll, 2006; Burke et al., 2006] might fail to deliver a result depending on the specifications of the controller.

1.2.2 Control of an industrial aircraft

Modelling of a flexible aircraft

The modelling process of an aircraft can roughly be divided into three steps as illustrated in Figure 1.3. Firstly, the model of the structure is built by taking into account the materials, the geometry, etc. Then, in a second step, the aerodynamics parameters are determined in wind tunnel for *some flight operating conditions* (such as the mass, the Mach, etc.) while the loads and gust propagation are modelled by simulation. At this point, the model of the aircraft consists actually in a set of models, each one being valid at one single flight operating condition. Finally, these models are adjusted with identification methods during flight tests. Since these tests are expensive, the number of flight operating conditions for which a model is available is quite low.

Simply speaking, the model of an aircraft represents both its *rigid* and *flexible* dynamics. The former translate the movement of the aircraft while the latter concern the deformation of the structure. Depending on the type of aircraft, the flexible dynamics are more or less important to describe accurately the behaviour of the aircraft. For instance, due to their large wingspan and their low weight, current electrical aircraft are extremely flexible [Gadient et al., 2012; Wang et al., 2014; Hesse and Palacios, 2014].

While the flexibility of industrial transport aircraft is not as pronounced, it cannot be ne-

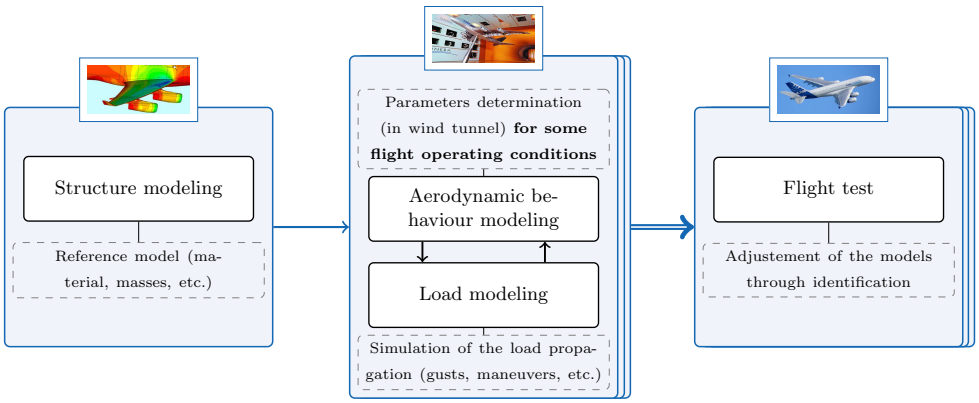


Figure 1.3: Overview of the modelling process in aeronautic.

glected since it can disturb the stability of the aircraft. This is especially true with modern aircraft which use more composite materials to decrease their weight and are therefore more flexible. Modelling accurately these flexible dynamics generally leads to a large dimension (see Remark 1 for more information on what *large* means in this context).

Finally, the global model of an aircraft is composed by a set of large-scale and possibly non-linear models representing the behaviour of the aircraft around different flight operating conditions. Designing control-laws from this global model is a challenging task and the underlying idea of a commonly used approach to achieve such an objective is introduced thereafter.

Remark 1. *The aircraft models used in the aeronautic industry have in general between* 500 *and* 5000 *states. This range of dimensions is relatively small in comparison to the very large-scale models that can be obtained by discretising partial differential equations [Li and Kamon, 2005; Lienemann et al., 2006; Dergham et al., 2011].*

However, for control purpose and considering the demanding specifications for control-laws in the aeronautical context, such a dimension is prohibitive. In particular, modern control or analysis tools might not be applicable directly on these models.

Commonly used approaches for the control of a flexible aircraft

Historically, control in aeronautic implicitly generally refers to *robust control*. Indeed, *(i)* the concept of robustness makes sense when considering the large domain (altitude, speed, etc.) in which an aircraft can evolve, *(ii)* due to the constraints imposed by certification, strong guarantees must be ensured by control-laws and *(iii)* the generic design framework associated with robust control is convenient for engineers (generalised plant, [Zhou et al., 1995]).

Based on the robust control framework, one can distinguish two main approaches to design a controller from a set of linear models,

- the first one consists in designing a controller for each model and to interpolate this set of controllers,
- the second one consists in first interpolating the models into an uncertain, parameter dependent model represented for instance as a Linear Fractional Representation (LFR) (see [Magni, 2006; Poussot-Vassal and Roos, 2012; Vizer et al., 2013]) on which one controller is designed.

While the first approach is still widely used in the industry, one believes that the second one is more promising due to the fact that it directly builds the LFR model of the aircraft which is mandatory to perform global stability or performance analysis. That is why it has been privileged here.

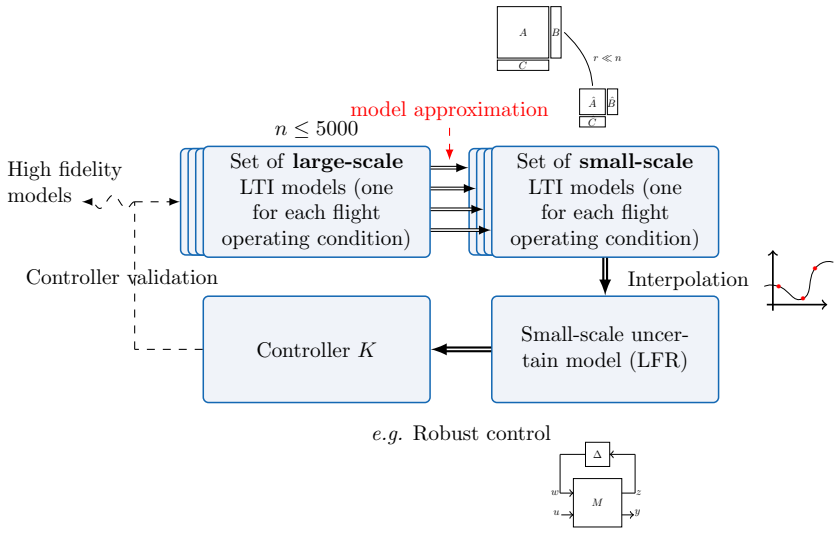


Figure 1.4: Global process for the control of industrial aircraft

In any case, for both approaches, the dimension of the initial models can become an issue. Indeed, on the one hand, a controller cannot efficiently be designed on a too large model and on the other hand the requirements concerning the interpolation of the set of models representing the aircraft cannot be achieved in a satisfactory way on large models. More specifically, the LFR must be stable over the whole parametric domain (since civilian aircraft are structurally stable), the parametric dependence must be of low complexity (otherwise control/analysis tools cannot be applied anymore) and the variation of the model between two parametric points must be as smooth as possible. Hence, for both approaches to be applicable on large-scale models, a preliminary approximation step is mandatory.

This approximation step too must satisfy some constraints :

- civilian aircraft are structurally stable, hence the reduced-order models must be stable too,
- as much as possible, the poles of the reduced-order model must not be too far from some poles of the large-scale model. Indeed, the poles of the original large-scale model have a physical meaning that is important to preserve for engineers.

The overall process for designing a controller from an initial set of large-scale linear models based on the interpolation of the models is summarised in Figure 1.4. Such a process has been applied on real case study detailed in Chapter 10.

1.2.3 Standard benchmarks

The numerical illustrations presented throughout this thesis are mainly based on three standard academic benchmarks available in the library $COMPL_eib$ [Leibfritz and Lipinski, 2003]. These benchmarks are not particularly high dimension, but each one has some interesting characteristics:

- the Los Angeles Hospital model (LAH) is a strictly proper and stable SISO model of order $n = 48$. Its frequency-response and its poles are plotted in Figure 1.5. This model has several significant dynamics which require a relatively high reduced order to be reproduced accurately.
- The Clamped Beam model (CBM) is a strictly proper and stable SISO model of order $n = 348$. Its frequency-response and its poles are plotted in Figure 1.6. This model has a

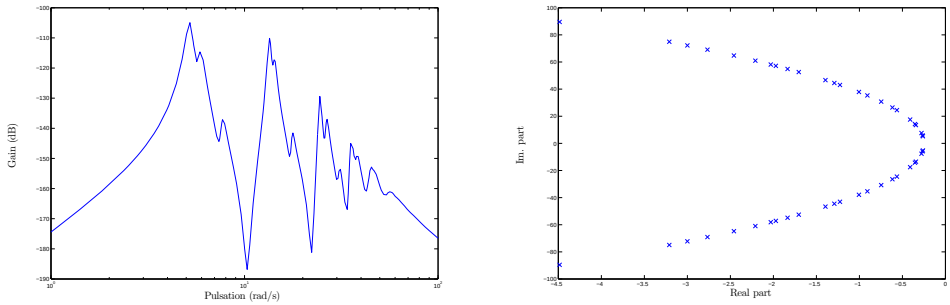


Figure 1.5: Frequency response (left) and eigenvalues (right) of the LAH model.

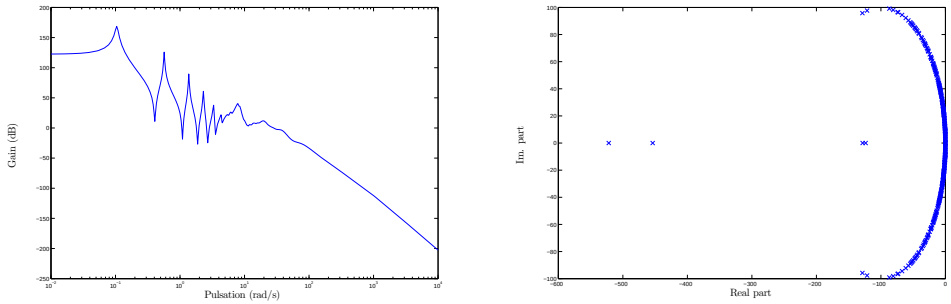


Figure 1.6: Frequency response (left) and eigenvalues (right) of the CBM model.

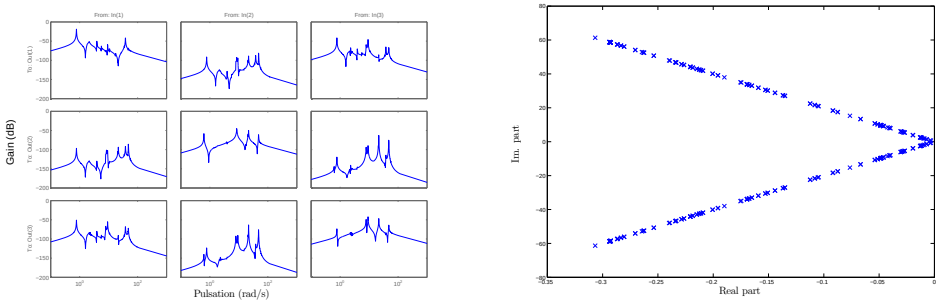


Figure 1.7: Frequency response (left) and eigenvalues (right) of the ISS model.

high gain dynamic at low frequency and several other small dynamics at higher frequencies. Catching accurately the low frequency dynamic directly yields a low approximation error.

- The International Space Station model (ISS) is a strictly proper and stable MIMO ($n_y = n_u = 3$) model of order $n = 270$. Its frequency-response and its poles are plotted in Figure 1.7. Aside from being MIMO, this model has high gain dynamics at several frequencies which makes it particularly well suited to illustrate the relevance of frequency-limited approximation techniques.

Additional benchmarks are described when used if they are publicly available.

1.3 Problem formulation

The approximation problem, as considered in this study, is formulated in a general way in Problem 1 thereafter.

Problem 1 (General LTI model approximation problem)**.** *Given a continuous, LTI, Multiple Inputs Multiple Outputs (MIMO), large-scale dynamical model* $\mathbf{H}$ *of order* n *described by its state-space representation,*

$$\mathbf{H} := \left\{ \begin{array}{rcl} \dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = & C\mathbf{x}(t) + D\mathbf{u}(t) \end{array} \right. ,$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n_u}$, $C \in \mathbb{R}^{n_y\times n}$ *and* $D \in \mathbb{R}^{n_y\times n_u}$. *The model approximation problem consists in finding a model* $\hat{\mathbf{H}}$ *of order* $r \ll n$ *given by the following realisation,*

$$\hat{\mathbf{H}} := \left\{ \begin{array}{rcl} \dot{\hat{\mathbf{x}}}(t) & = & \hat{A}\hat{\mathbf{x}}(t) + \hat{B}\mathbf{u}(t) \\ \hat{\mathbf{y}}(t) & = & \hat{C}\hat{\mathbf{x}}(t) + \hat{D}\mathbf{u}(t) \end{array} \right. ,$$

where $\hat{A} \in \mathbb{R}^{r\times r}$, $\hat{B} \in \mathbb{R}^{r\times n_u}$, $\hat{C} \in \mathbb{R}^{n_y\times r}$ *and* $\hat{D} \in \mathbb{R}^{n_y\times n_u}$, *which accurately reproduces the behaviour of the full-order model* $\mathbf{H}$ *in some sense.*

The evaluation of the quality of the reduced-order model as formulated in Problem 1 is purposely left vague at this stage and is detailed later using standard norms of LTI models. The important point implied by the formulation of Problem 1 lies in the fact that in this study, it is assumed that a realisation of the full-order model is available. This hypothesis therefore differentiates the methods presented in this study from the recently developed realisation-free (or data driven) methods such as those presented in [Antoulas et al., 2012; Ionita, 2013] which are on the frontier between model approximation and model identification [Unbehauen and Rao, 1990; Vayssettes et al., 2014].

1.4 Overview of the contributions

This study is focused on the approximation of continuous, LTI, stable, large-scale dynamical models over a bounded frequency range. In order to address this problem, several steps have been followed during this thesis. In a first time, an empirical method has been developed by modifying an existing sub-optimal $\mathcal{H}_2$ model approximation method to perform $\mathcal{H}_{2,\Omega}$ oriented model approximation [Vuillemin et al., 2013a,b]. The method has proven to be efficient in practice but the lack of theoretical guarantee and control over the approximation error has led us to consider another approach based on optimality considerations.

In particular, the approximation over a bounded frequency range has been formulated as an optimisation problem in terms of the frequency-limited $\mathcal{H}_2$-norm. In the $\mathcal{H}_2$ optimal approximation problem, the poles-residues formulation of the $\mathcal{H}_2$-norm had led to convenient optimality

conditions. Hence, we have firstly proposed a similar formulation for the frequency-limited $\mathcal{H}_2$-norm [Vuillemin et al., 2012b, 2014c]. Then, based on this formulation of the norm, the first-order optimality conditions for the optimal frequency-limited approximation problem have been derived [Vuillemin et al., 2014b]. However, unlike the $\mathcal{H}_2$ problem, these optimality conditions could not be expressed as convenient interpolation conditions. Hence, an unconstrained complex domain optimisation algorithm aimed at finding one optimal reduced-order model has been developed [Vuillemin et al., 2014b]. Finally, this method has been brought one step further towards some control objective by using the two upper bounds on the $\mathcal{H}_\infty$-norm that we have proposed [Vuillemin et al., 2014d] in order to build a frequency template of the approximation error that can be exploited in the robust control framework.

Besides, from a practical point of view, all the proposed methods and tools have been integrated in the MORE Toolbox[3] [Poussot-Vassal and Vuillemin, 2012] and have successfully been applied in several industrial applications. In a first time, the approximation methods developed in this study have been applied on industrial large-scale aircraft models to evaluate their efficiency on this type of models [Vuillemin et al., 2012a, 2013b]. Then, they have been used for the control of large-scale models to design *(i)* an anti-vibration control law for a business jet aircraft [Poussot-Vassal et al., 2013], *(ii)* a control law ensuring flight performance and load clearance in presence of input saturation on a model representing the longitudinal behaviour of a flexible civilian aircraft [Burlion et al., 2014]. Additionally, a process for the creation of a low-order uncertain, parameter varying model from a set of large-scale dynamical models has been developed based on model approximation [Poussot-Vassal and Vuillemin, 2013; Poussot-Vassal et al., 2014]. Based on a similar method, the global stability and performance of the set of controlled large-scale models representing a business jet aircraft at different flight operating conditions subject to actuator saturation have been proven by taking into account the error induced by the approximation and the interpolation of the initial large-scale models [Vuillemin et al., 2014a].

1.5 Manuscript overview

This manuscript is divided into three parts in addition of the introduction that gather chapters which concern the state of the art, the contributions of this thesis with regard to frequency-limited model approximation and the conclusions of this study.

Part II : State of the Art

Chapter 2 : Preliminary in LTI systems theory

This chapter is aimed at recalling some general elements about linear systems theory and at introducing the general notations used in this thesis. In particular, two elements that form the basis of this thesis are recalled : the partial fraction decomposition of a transfer function and the frequency-limited $\mathcal{H}_2$-norm of LTI models.

Chapter 3 : Standard model approximation methods

In this chapter, some well-known model approximation techniques based either on state-space truncation or moment matching are recalled. In particular, the balanced truncation is recalled since it is one of the most popular approximation method and that it serves as reference in the benchmarks. Implicit moment matching methods are presented as an introduction for the optimal $\mathcal{H}_2$ model approximation methods.

[3] The toolbox is available from `w3.onera.fr/more` and its use is illustrated by several code samples available in Appendix C.

Chapter 4 : Optimal $\mathcal{H}_2$ model approximation

The optimal $\mathcal{H}_2$ approximation problem is introduced together with some methods to address it. The methods that are presented are those based on the interpolation of the large-scale model through projection on some specific Krylov subspaces and one optimisation method relying on the partial fraction decomposition of the reduced-order transfer function.

Chapter 5 : Frequency-weighted and frequency-limited model approximation

In this chapter, some methods aimed at model approximation over a bounded frequency range are presented. Methods based on the use of filters are classified here as frequency-weighted model approximation methods while those that do not require any weight falls into the frequency-limited model approximation methods. Among these methods, the frequency-weighted and frequency-limited balanced truncation are presented.

Part III : Frequency-limited approximation of linear dynamical models

Chapter 6 : development of a first approach for frequency-limited model approximation

In this chapter, an empirical method built by modification of a sub-optimal $\mathcal{H}_2$ model approximation method to achieve model approximation over a bounded frequency-range is presented. The improvement of performances in terms of $\mathcal{H}_{2,\Omega}$-norm in comparison to the original method is illustrated through various examples. Its efficiency is also illustrated by comparison with the frequency-limited balanced truncation.

Chapter 7 : Formulation of the $\mathcal{H}_{2,\Omega}$-norm with the poles and residues of the transfer function

The only formulation available for the computation of the frequency-limited $\mathcal{H}_2$-norm was the gramian one. In this chapter, a formulation based on the poles and residues of the transfer function is developed thus generalising the poles-residues formulation of the $\mathcal{H}_2$-norm. This is done both for models with semi-simple poles only and for model with higher order poles. Besides, two upper bounds on the $\mathcal{H}_\infty$-norm of LTI dynamical models are derived from this formulation.

The results presented in this chapter constitute the basis of the major contributions of this thesis. Hence, this chapter plays a pivotal role in this manuscript.

Chapter 8 : Formulation of the $\mathcal{H}_{2,\Omega}$-norm of the error and differentiation of its gradient

The poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm is used to express the approximation error between the large-scale and reduced-order models thus leading to a formulation of the optimal $\mathcal{H}_{2,\Omega}$ model approximation problem in terms of the poles and residues of the reduced-order model. The gradient of the error is derived with respect to the reduced-order model parameters and the first-order optimality conditions are presented.

Chapter 9 : Development of DARPO, a descent algorithm for the optimal $\mathcal{H}_{2,\Omega}$ approximation problem

Since the first-order optimality conditions of the optimal $\mathcal{H}_{2,\Omega}$ problem cannot be expressed as convenient interpolation conditions, an optimisation scheme is developed to find a local minimum of the approximation problem instead. In this chapter, this optimisation method is described and compared to other frequency-limited approximation techniques on academic benchmarks. It

is also used to address a fictive problem of control of large-scale model by exploiting the bounds on the $\mathcal{H}_\infty$-norm of the approximation error.

Chapter 10 : Industrial aeronautical use case

In this chapter, the model approximation methods and tools developed in this thesis are used within a global process aimed at designing an anti-vibration control-law for an industrial business jet aircraft. More specifically, the preliminary study in which the control design problem is addressed on one single large-scale model representing the business jet aircraft at one flight operating condition is presented. Then, based on a controller designed with the extension of this approach, the global stability of the set of controlled large-scale models is assessed over the whole parametric domain.

Part IV : Conclusion

Chapter 11 : Discussion

This chapter is aimed at recalling the contributions of this thesis as well as their limitations and to present some improvement leads to alleviate them. In addition, some short-term extensions are presented.

Chapter 12 : Perspectives

In this chapter, some long-term outlook concerning the extension of the methods and tools developed in this thesis to other type of models or their application for control purpose are drawn.

Part II

State of the art

Chapter 2

Preliminary in LTI systems theory

In this chapter, some general elements about LTI systems theory are recalled and the associated notations introduced. The material is standard and can be found in many books such as [Zhou et al., 1995]. Additional references are mentioned for less standard material when required.

Section 2.1 is aimed at presenting some elements about model representation and gramians, and Section 2.2 concerns the norms of LTI models.

Contents

2.1 Generalities 17
2.1.1 Representation of LTI dynamical models 17
2.1.2 Gramians and balanced realisation 20
2.2 Norms of systems 25
2.2.1 $\mathcal{H}_2$-norm 25
2.2.2 Frequency-limited $\mathcal{H}_2$-norm 26
2.2.3 $\mathcal{H}_\infty$-norm 29

2.1 Generalities

In Section 2.1.1, some elements about the representation of LTI models are recalled. In particular, the poles-residues representation of LTI models is presented. Then in Section 2.1.2, the infinite gramians, the balanced realisation and the frequency-limited gramians are detailed.

2.1.1 Representation of LTI dynamical models

Time-domain representation

A n-th order, MIMO (n_y outputs and n_u inputs) LTI dynamical model $\mathbf{H}$ can be represented in the time-domain by a state-space realisation,

$$\mathbf{H} := \left\{ \begin{array}{rcl} \dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = & C\mathbf{x}(t) + D\mathbf{u}(t) \end{array} \right. ,$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n_u}$, $C \in \mathbb{R}^{n_y\times n}$ and $D \in \mathbb{R}^{n_y\times n_u}$. This realisation may also be written as $\mathbf{H} := (A, B, C, D)$ or as

$$\mathbf{H} := \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \in \mathbb{R}^{(n+n_y)\times(n+n_u)}.$$

Frequency-domain representation

The transfer function $H(s)$ associated with the model $\mathbf{H}$ is given by

$$H(s) = C\left(sI_n - A\right)^{-1} B + D \in \mathbb{C}^{n_y\times n_u},$$

and represents the model $\mathbf{H}$ in the frequency-domain. Depending on the multiplicity of the poles of the model, the transfer function can be written in different forms, in particular :

- if the model has semi-simple poles only, *i.e.* , *A is diagonalisable*, then the transfer function $H(s)$ can be decomposed as

$$H(s) = \sum_{i=1}^{n} \frac{\Phi_i}{s - \lambda_i} + D, \tag{2.1}$$

where $\Phi_i \in \mathbb{C}^{n_y \times n_u}$ with $rank(\Phi_i) = 1$ are the residues of H associated with the poles λ_i, $i = 1, \ldots, n$. For semi-simple poles, the residues Φ_i are defined as

$$\Phi_i = \lim_{s \to \lambda_i} (s - \lambda_i) H(s).$$

They can also be computed thanks to the state-space matrices as

$$\Phi_i = \mathbf{c_i}^T \mathbf{b_i} = CX\mathbf{e_i}\mathbf{e_i}^T X^{-1} B,$$

where $\mathbf{e_i} \in \mathbb{R}^n$ is the i-th canonical vector and $X = \begin{bmatrix} \mathbf{x_1} & \ldots & \mathbf{x_n} \end{bmatrix} \in \mathbb{C}^{n \times n}$ is the matrix containing the right eigenvectors $\mathbf{x_i} \in \mathbb{C}^n$ $(i = 1, \ldots, n)$ of A.

Note that by imposing a condition on the norms of either $\mathbf{c_i}$ or $\mathbf{b_i}$ $(i = 1, \ldots, n)$, then the representation (2.1) has the minimal number of parameters required to represent a LTI dynamical model of order n. This formulation of the model transfer function is sometimes called the diagonal canonical form [Van Dooren et al., 2008a] and may be referred here as the poles-residues formulation of the transfer function.

- If the matrix A *is defective*, then the decomposition is slightly more complicated. Indeed, let us consider a model $\mathbf{H}$ with n_b Jordan blocks $J_i \in \mathbb{C}^{n_i \times n_i}$ associated with $\lambda_i \in \mathbb{C}$. In that case, the partial fraction decomposition of $H(s)$, is given by

$$H(s) = \sum_{i=1}^{n_b} H_{\lambda_i}(s) + D = \sum_{i=1}^{n_b} \underbrace{\sum_{k=1}^{n_i} \frac{\Phi_i^{(k)}}{(s - \lambda_i)^k}}_{H_{\lambda_i}(s)} + D, \tag{2.2}$$

where the residues $\Phi_i^{(k)} \in \mathbb{C}^{n_y \times n_u}$ $k = 1, \ldots, n_i$, corresponding to the pole λ_i are given by

$$\Phi_i^{(k)} = \lim_{s \to \lambda_i} \frac{1}{(n_i - k)!} \frac{d^{n_i - k}}{ds^{n_i - k}} (s - \lambda_i)^{n_i} H(s). \tag{2.3}$$

Again, these residues can be expressed with the state-space matrices A, B and C. Indeed, let us consider an invertible matrix $T \in \mathbb{C}^{n \times n}$ that transforms A into its Jordan canonical form :

$$T^{-1} A T = J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_{n_b} \end{bmatrix},$$

where each Jordan block J_i $(i = 1, \ldots, n_b)$ is given by

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \in \mathbb{C}^{n_i \times n_i},$$

By denoting $C_i \in \mathbb{R}^{n_y \times n_i}$ (resp. $B_i \in \mathbb{R}^{n_i \times n_u}$) the columns (resp. lines) of CT (resp. $T^{-1}B$) associated with the block J_i [1], the transfer function $H_{\lambda_i}(s)$ $(i = 1, \ldots, n_b)$ can be

[1] In other words, $C_i = CTP_i^T$ and $B_i = P_i T^{-1} B$ with $P_i = \begin{bmatrix} 0_{n_i \times (n_1 + \ldots + n_{i-1})} & I_{n_i} & 0_{n_i \times (n_{i+1} + \ldots + n_b)} \end{bmatrix} \in \mathbb{R}^{n_i \times n}$.

expressed as

$$H_{\lambda_i}(s) = C_i(sI_{n_i} - J_i)^{-1}B_i. \tag{2.4}$$

Due to the structure of the matrix J_i in (2.4), the inverse $(sI_{n_i} - J_i)^{-1}$ can be written as a sum of rational functions of s,

$$(sI_{n_i} - J_i)^{-1} = (s-\lambda_i)^{-1}F_1 + (s-\lambda_i)^{-2}F_2 + \ldots + (s-\lambda_i)^{-n_i}F_{n_i},$$

where $F_k \in \mathbb{R}^{n_i \times n_i}$ is the matrix with 1 on the $(k-1)$-th upper diagonal and 0 elsewhere, *i.e.*

$$F_1 = I_{n_i},\ F_2 = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}, \ldots,\ F_{n_i} = \begin{bmatrix} 0 & 0 & & 1 \\ & 0 & \ddots & \\ & & \ddots & 0 \\ & & & 0 \end{bmatrix}.$$

Hence, for $i = 1, \ldots, n_b$ and $k = 1, \ldots, n_i$, the residues $\Phi_i^{(k)}$ are given by :

$$\Phi_i^{(k)} = C_i F_k B_i.$$

This decomposition, also called Jordan canonical form of $H(s)$ [Van Dooren et al., 2008a], is illustrated on a simple case in Example 1.

Example 1 (Simple model with one double pole)**.** *Let us consider the model* **H** *which realisation is the following :*

$$A = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix},\ B = C^T = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ D = 0.$$

The matrix A is already in its Jordan form and has a double eigenvalue at $\lambda = -1$. The transfer function $H(s)$ associated with the model **H** *is given by*

$$H(s) = \frac{2s+3}{(s+1)^2},$$

and can be decomposed as

$$H(s) = \frac{\Phi_\lambda^{(1)}}{s+1} + \frac{\Phi_\lambda^{(2)}}{(s+1)^2}.$$

The residues $\Phi_\lambda^{(1)}$ and $\Phi_\lambda^{(2)}$ associated with the pole λ can either be computed with the equation (2.3),

$$\Phi_\lambda^{(1)} = \lim_{s\to-1} \frac{1}{(2-1)!}\frac{d^{2-1}}{ds^{2-1}}(s+1)^2H(s) = \lim_{s\to-1}\frac{d}{ds}2s+3 = 2,$$

and,

$$\Phi_\lambda^{(2)} = \lim_{s\to-1} \frac{1}{(2-2)!}\frac{d^{2-2}}{ds^{2-2}}(s+1)^2H(s) = \lim_{s\to-1} 2s+3 = 1,$$

or with the state-space matrices. Indeed, here,

$$F_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \textit{and} \quad F_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

thus,

$$\Phi_1^{(\lambda)} = CF_1B = 2 \quad \textit{and} \quad \Phi_2^{(\lambda)} = CF_2B = 1.$$

2.1.2 Gramians and balanced realisation

Infinite gramians

The controllability and observability gramians $\mathcal{P}$ and $\mathcal{Q}$ (see Definition 1) are important matrices in system theory, they are related to the concepts of state reachability and observability (see for instance [Zhou et al., 1995] for further details). In particular, they give information about the input-output energy of the model :

- the minimal energy ϵ_r required to make the system move from the state 0 to the state $\mathbf{x}$ is given by
$$\epsilon_r = \mathbf{x}^T \mathcal{P}^{-1} \mathbf{x}.$$
If ϵ_r is large, then the state $\mathbf{x}$ is said to be hard to reach.
- The maximal observation energy ϵ_o obtained by releasing the system from an initial state $\mathbf{x}$ without any input excitation is given by
$$\epsilon_o = \mathbf{x}^T \mathcal{Q} \mathbf{x}.$$
It the energy ϵ_o is small, the state $\mathbf{x}$ is said to be hard to observe.

These concepts of energy transfer are the basis of some model approximation methods.

Definition 1 (Infinite gramians)**.** *Given an asymptotically stable LTI dynamical model* $\mathbf{H} := (A, B, C, D)$*, the infinite controllability and observability gramians associated with* $\mathbf{H}$*, denoted* $\mathcal{P}$ *and* $\mathcal{Q}$ *respectively, are defined as,*

$$\begin{aligned} \mathcal{P} &= \int_0^\infty e^{At} B B^T e^{A^T t} dt, \\ \mathcal{Q} &= \int_0^\infty e^{A^T t} C^T C e^{At} dt, \end{aligned}$$

in the time domain and as,

$$\begin{aligned} \mathcal{P} &= \frac{1}{2\pi} \int_{-\infty}^\infty T(\nu)^H B B^T T(\nu) d\nu, \\ \mathcal{Q} &= \frac{1}{2\pi} \int_{-\infty}^\infty T(\nu)^H C^T C T(\nu) d\nu, \end{aligned}$$

with $T(\nu) = (j\nu I_n - A)^{-1}$*, in the frequency domain.*

The controllability and observability gramians can be computed by solving two Lyapunov equations. Indeed the controllability gramian $\mathcal{P}$ is the solution of

$$A\mathcal{P} + \mathcal{P}A^T + BB^T = 0, \tag{2.9}$$

while the observability gramian $\mathcal{Q}$ is the solution of

$$A^T\mathcal{Q} + \mathcal{Q}A + C^T C = 0. \tag{2.10}$$

It is important to note that the gramian depends of the basis in which the model is expressed. Hence a state $\mathbf{x}$ which yields a small observation energy ϵ_o (*i.e.* it is *hard to observe*) might only requires a small energy ϵ_r to be reached (*i.e.* it is *easy to reach*) and conversely in another basis. However, the model can be put in a basis in which both notions coincide. This basis is called the *balanced realisation*.

Balanced realisation

The balanced realisation $(A_{bal}, B_{bal}, C_{bal}, D_{bal})$ of a LTI dynamical model $\mathbf{H} := (A, B, C, D)$ is obtained through a state transformation $\mathbf{x}(t) = T\mathbf{x_{bal}}(t)$, where $T \in \mathbb{R}^{n\times n}$ is a nonsingular transformation matrix, as,

$$A_{bal} = T^{-1}AT, \quad B_{bal} = T^{-1}B, \quad C_{bal} = CT, \quad D_{bal} = D.$$

In this basis, the gramians are given by

$$\mathcal{P}_{bal} = T^{-1}\mathcal{P}T^{-T} \quad \text{and} \quad \mathcal{Q}_{bal} = T^{T}\mathcal{Q}T.$$

If both $\mathcal{P}_{bal}$ and $\mathcal{Q}_{bal}$ are diagonal, the transformation T is called a *contragredient* transformation. If in addition, they are equals,

$$\mathcal{P}_{bal} = \mathcal{Q}_{bal} = \Sigma = \mathbf{diag}\left(\sigma_1, \ldots, \sigma_n\right),$$

the state coordinates are said to be internally balanced and T is called the *balanced transformation* [Moore, 1981; Laub et al., 1987]. The diagonal entries σ_i $(i = 1, \ldots, n)$ of Σ are called the *Hankel singular values* (see for instance [Antoulas, 2005, chap. 5] for further details on the Hankel operator). For reachable, observable and stable models, they can be computed as the square roots of the eigenvalues of the product of the two gramians, *i.e.* for $i = 1, \ldots, n$,

$$\sigma_i = \sqrt{\lambda_i\left(\mathcal{P}\mathcal{Q}\right)}.$$

Unlike the gramians, the Hankel singular values of the model are invariant by state transformation.

Remark 2. *Note that more generally, the concept of balancing consists in simultaneously diagonalising two positive definite matrices. Depending on which matrices are diagonalised, the type of balancing changes. In this section, only the* Lyapunov balancing *is presented, but there exist also the* stochastic balancing, *the* bounded real balancing *and the* positive real balancing *which are described more in depth in [Antoulas, 2005, chap.7].*

As mentioned before, when the model is balanced, the reachability and observability energy coincide. Indeed, consider the state $\mathbf{x_0} = \mathbf{e_i} = \begin{bmatrix} 0 & \ldots & 1 & \ldots & 0 \end{bmatrix}^T$, then

$$\epsilon_r = \mathbf{x_0}^T\mathcal{P}^{-1}\mathbf{x_0} = \mathbf{e_i}^T\Sigma^{-1}\mathbf{e_i} = {\sigma_i}^{-1},$$

and similarly,

$$\epsilon_o = \mathbf{x_0}^T\mathcal{Q}\mathbf{x}_0 = \sigma_i.$$

In the balanced realisation, $\epsilon_o = \epsilon_r^{-1}$, hence the states that are easy to observe (large ϵ_o) are also easily reachable (small ϵ_r).

Different approaches exist to find the balanced transformation, here two standard methods presented in [Laub et al., 1987] are recalled in Algorithm 1 and in Algorithm 2. For a more in depth study of the balanced transformations and their numerical differences, see [Antoulas, 2005, chap. 7] and references therein.

Algorithm 1 Balanced Transformation

1: Compute directly the lower Cholesky factorizations of the gramians with the method presented in [Hammarling, 1982] :

$$\mathcal{P} = L_c L_c^T \text{ and } \mathcal{Q} = L_o L_o^T.$$

2: Compute the singular value decomposition of $L_o^T L_c$,

$$L_o^T L_c = U \Sigma V^T,$$

where Σ is the diagonal matrix containing the Hankel singular values of $\mathbf{H}$

3: The balanced transformation is given by

$$T = L_c V \Sigma^{-\frac{1}{2}}, \text{ and } T^{-1} = \Sigma^{-\frac{1}{2}} U^T L_o^T.$$

One can verify that the transformation T constructed in Algorithm 1 is indeed a contragredient transformation by projecting the gramians,

$$\begin{aligned} T^{-1}\mathcal{P}T^{-T} &= \Sigma^{-\frac{1}{2}} U^T L_o^T \mathcal{P} L_o U \Sigma^{-\frac{1}{2}} \\ &= \Sigma^{-\frac{1}{2}} U^T \underbrace{L_o^T L_c}_{U\Sigma V^T} \underbrace{L_c^T L_o}_{V\Sigma U^T} U \Sigma^{-\frac{1}{2}} = \Sigma. \end{aligned}$$

Similarly,

$$\begin{aligned} T^T \mathcal{Q} T &= \Sigma^{-\frac{1}{2}} V^T L_c^T \mathcal{Q} L_c V \Sigma^{-\frac{1}{2}} \\ &= \Sigma^{-\frac{1}{2}} V^T \underbrace{L_c^T L_o}_{V\Sigma U^T} \underbrace{L_o^T L_c}_{U\Sigma V^T} V \Sigma^{-\frac{1}{2}} = \Sigma. \end{aligned}$$

Algorithm 2 Balanced Transformation

1: Compute the gramians $\mathcal{P}$ and $\mathcal{Q}$ with (2.9) and (2.10).

2: Compute the lower Cholesky factorization of $\mathcal{P}$,

$$\mathcal{P} = L_c L_c^T.$$

3: Solve the following symmetric eigenvalue problem :

$$L_c^T \mathcal{Q} L_c = V \Sigma^2 V^T,$$

where Σ contains the Hankel singular values of the system.

4: The balanced transformation is given by

$$T = L_c V \Sigma^{-\frac{1}{2}} \text{ and } T^{-1} = \Sigma^{\frac{1}{2}} V^T {L_c}^{-1}$$

Again, the transformation T given by the Algorithm 2 is a contragredient transformation,

$$T^{-1}\mathcal{P}T^{-T} = \Sigma^{\frac{1}{2}} V^T {L_c}^{-1} \underbrace{\mathcal{P}}_{L_c L_c^T} L_c^{-T} V \Sigma^{\frac{1}{2}} = \Sigma,$$

and

$$T^T \mathcal{Q} T = \Sigma^{-\frac{1}{2}} V^T \underbrace{L_c^T \mathcal{Q} L_c}_{V\Sigma^2 V^T} V \Sigma^{-\frac{1}{2}} = \Sigma.$$

In Algorithm 2, the gramians are explicitly computed whereas in Algorithm 1, only their Cholesky factorizations are required [Hammarling, 1982]. The latter method is numerically

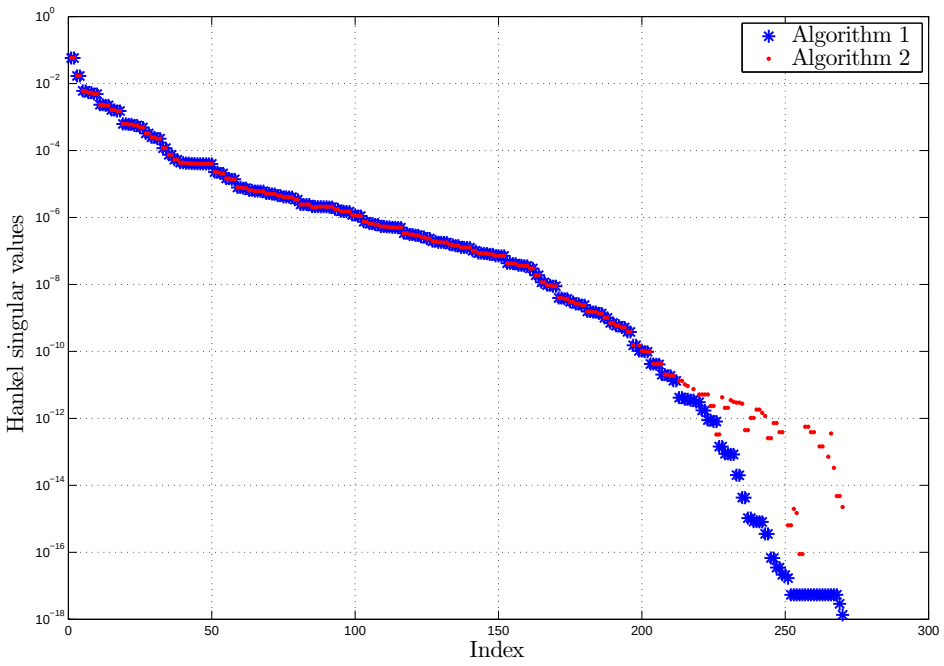


Figure 2.1: Hankel singular values of the ISS model obtained with Algorithm 1 and Algorithm 2.

more reliable because it does not involve the explicit computation of the products BB^T and C^TC arising in the Lyapunov equations (2.9) and (2.10) which can lead to discard some small singular values. An illustration of the difference between the two algorithms is shown in Example 2. The counterpart is that Algorithm 2 requires less operations than Algorithm 1 (see [Laub et al., 1987] for a detailed description of the number of operations required for both algorithms). For similar reasons, the product in step 2 of Algorithm 1 could also be avoided by using a dedicated method for computing the singular value decomposition of a product of matrices [Heath et al., 1986].

Example 2. *In this example, the Hankel singular values of the ISS model are computed with the algorithms 1 and 2 and plotted in Figure 2.1.*

The largest singular values are not affected by the algorithm, but one can see that differences appears for small singular values. Indeed for the singular values smaller than 10^{-10}*, the algorithm that explicitly computes the gramians* $\mathcal{P}$ *and* $\mathcal{Q}$ *gives erroneous results.*

Frequency-limited gramians

In [Gawronski and Juang, 1990], the authors have proposed to restrict the frequency-domain definition of the gramian in order to build frequency-limited gramians (see Definition 2).

Definition 2 (Frequency-limited gramians)**.** *Given a frequency interval* $\Omega = [\omega_1, \omega_2]$*,* $0 \leq \omega_1 < \omega_2$ *and a LTI model* $\mathbf{H} := (A, B, C)$*. The frequency-limited controllability and observability gramians of* $\mathbf{H}$ *over* Ω*, denoted respectively as* $\mathcal{P}_\Omega$ *and* $\mathcal{Q}_\Omega$*, are defined as*

$$\begin{aligned} \mathcal{Q}_\Omega &= \mathcal{Q}(\omega_2) - \mathcal{Q}(\omega_1), \\ \mathcal{P}_\Omega &= \mathcal{P}(\omega_2) - \mathcal{P}(\omega_2), \end{aligned}$$

where

$$\begin{aligned}\mathcal{P}(\omega) &= \frac{1}{2\pi}\int_{-\omega}^{\omega} T(\nu)BB^T T^*(\nu)d\nu,\\ \mathcal{Q}(\omega) &= \frac{1}{2\pi}\int_{-\omega}^{\omega} T^*(\nu)C^T CT(\nu)d\nu,\end{aligned}$$

with $T(\nu) = (j\nu I_n - A)^{-1}$.

$\mathcal{P}(\omega)$ and $\mathcal{Q}(\omega)$ are the solutions of the following Lyapunov equations

$$\begin{aligned}A\mathcal{P}(\omega) + \mathcal{P}(\omega)A^T + W_c(\omega) &= 0,\\ A^T\mathcal{Q}(\omega) + \mathcal{Q}(\omega)A + W_o(\omega) &= 0,\end{aligned}$$

where the last terms are given by

$$\begin{aligned}W_c(\omega) &= S(\omega)BB^T + BB^T S^H(\omega),\\ W_o(\omega) &= S^H(\omega)C^T C + C^T CS(\omega),\end{aligned} \tag{2.11}$$

and where, by denoting $\mathbf{logm}(M)$ the matrix logarithm of M,

$$\begin{aligned}S(\omega) &= \frac{1}{2\pi}\int_{-\omega}^{\omega} T(\nu)d\nu,\\ &= \frac{j}{2\pi}\mathbf{logm}\left((A + j\omega I_n)(A - j\omega I_n)^{-1}\right).\end{aligned}$$

By denoting $W_c(\Omega) = W_c(\omega_2) - W_c(\omega_1)$ and $W_o(\Omega) = W_o(\omega_2) - W_o(\omega_1)$, it turns out that similarly to the infinite gramians, the frequency-limited gramians $\mathcal{Q}_\Omega$ and $\mathcal{P}_\Omega$ can be computed by solving Lyapunov equations,

$$\begin{aligned}A\mathcal{P}_\Omega + \mathcal{P}_\Omega A^T + W_c(\Omega) &= 0,\\ A^T\mathcal{Q}_\Omega + \mathcal{Q}_\Omega A + W_o(\Omega) &= 0.\end{aligned} \tag{2.12}$$

Alternatively, they can be computed directly from the infinite gramians as,

$$\begin{aligned}\mathcal{Q}_\Omega &= W_o(\Omega)^H\mathcal{Q} + \mathcal{Q}W_o(\Omega),\\ \mathcal{P}_\Omega &= \mathcal{P}W_c(\Omega)^H + W_c(\Omega)\mathcal{P}.\end{aligned}$$

Remark 3 (Time limited gramians)**.** *By following the same idea, the authors in [Gawronski and Juang, 1990] have also built time-limited gramians.*

Remark 4 (Balancing of the frequency-limited gramians)**.** *Note that the last terms* $W_c(\Omega)$ *and* $W_o(\Omega)$ *of the Lyapunov equations* (2.12) *are not necessarily positive definite, hence the Lyapunov solver that directly gives the Cholesky factorisation of the gramians [Hammarling, 1982] is not applicable here.*

The frequency-limited gramians are also positive definite and can be balanced. Hence they can be used to compute quantities similar to the Hankel singular values as illustrated in Example 3.

Example 3 (Singular values obtained from frequency-limited gramians)**.** *The idea behind this example comes from [Gawronski, 2004, chap. 4]. Let us consider a* 6*-th order model which poles are* $\lambda_1 = -0.1 + 3j$, $\lambda_2 = -0.05 + 10j$, $\lambda_3 = -0.01 + 20j$ *and their complex conjugate. The transfer function associated with this model is given by*

$$H(s) = \frac{1}{(s^2 + 0.2s + 9.01)(s^2 + 0.1s + 100)(s^2 + 0.02s + 400)}.$$

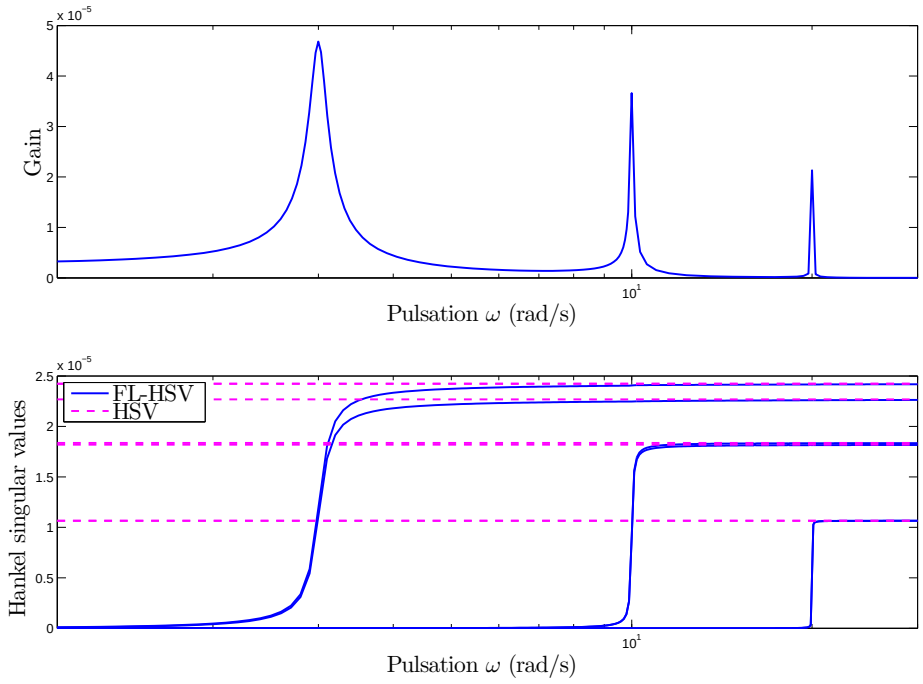


Figure 2.2: Frequency response of $H(s)$ (top) and Hankel singular values obtained from the frequency-limited gramians for a varying pulsation ω (bottom).

The frequency-limited gramians $\mathcal{Q}_\Omega$ and $\mathcal{P}_\Omega$ are computed for $\Omega = [0, \omega]$ where ω varies from 0 *to* 20 *and for each value of ω, the square roots of the eigenvalues of the product $\mathcal{Q}_\Omega\mathcal{P}_\Omega$ are computed,* i.e. *what may be called the* frequency-limited Hankel singular values *(FL-HSV) of the model.*

The frequency response of the model, its Hankel singular values (HSV) and its frequency-limited Hankel singular values (FL-HSV) are plotted in Figure 2.2.

One can observe that the frequency-limited singular values are null (or very small) before they cross the natural frequency of their corresponding mode, then, they are almost constant as ω increases and have the same value as the Hankel singular values.

Note that this correspondence between the poles and the singular values is obvious in this case because the modes are poorly damped but it would not be as obvious with highly damped modes.

2.2 Norms of systems

Various norms with different interpretation can be used for judging the quality of a reduced-order model. Here, the definitions of the $\mathcal{H}_2$-norm (Section 2.2.1), its frequency-limited counterpart (Section 2.2.2) and the $\mathcal{H}_\infty$-norm (Section 2.2.3) are quickly recalled.

2.2.1 $\mathcal{H}_2$-norm

Definition and properties

The $\mathcal{H}_2$-norm can be interpreted, in the time domain, as the energy of the model output signal when its input is a Dirac, and equivalently in the frequency domain, as the output energy of its

transfer function when the input is a white noise. Its frequency domain definition is recalled in Definition 3.

Due to its interesting physical interpretation, its properties and the relative simplicity to compute it, the $\mathcal{H}_2$-norm has been widely considered in model approximation [Meier and Luenberger, 1967; Wilson, 1974; Fulcheri and Olivi, 1998; Van Dooren et al., 2008b; Gugercin et al., 2008] (further information about optimal $\mathcal{H}_2$ model approximation are presented in Chapter 4).

Definition 3 ($\mathcal{H}_2$-norm)**.** *Given a LTI dynamical model* $\mathbf{H}$ *whose transfer function is* $H(s) \in \mathbb{C}^{n_y \times n_u}$*, the* $\mathcal{H}_2$*-norm of* $\mathbf{H}$*, denoted* $\|H\|_{\mathcal{H}_2}$ *is defined, in the frequency-domain, as*

$$\|H\|_{\mathcal{H}_2}^2 := \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{tr}\left(H(j\nu)H(-j\nu)^T\right) d\nu.$$

The $\mathcal{H}_2$-norm is infinite for models which have poles in the right half plane or a non-null direct feedthrough matrix D.

Computation

The $\mathcal{H}_2$-norm of a model can be expressed either with the infinite gramians of the model as presented in Theorem 1 or with its poles and residues, as shown in Theorem 2.

Theorem 1 (Gramian formulation of the $\mathcal{H}_2$-norm)**.** *Given an asymptotically stable LTI dynamical model* $\mathbf{H}$ *which infinite observability and controllability gramians are* $\mathcal{Q}$ *and* $\mathcal{P}$*, respectively. The* $\mathcal{H}_2$*-norm of* $\mathbf{H}$*, denoted* $\|H\|_{\mathcal{H}_2}$ *is given by*

$$\begin{aligned} \|H\|_{\mathcal{H}_2}^2 &= \mathbf{tr}\left(C\mathcal{P}C^T\right) \\ &= \mathbf{tr}\left(B^T\mathcal{Q}B\right). \end{aligned}$$

Theorem 2 (Poles-residues formulation of the $\mathcal{H}_2$-norm)**.** *Given a* n*-th order asymptotically stable and strictly proper LTI dynamical model* $\mathbf{H}$ *whose transfer function is* $H(s)$ *and which has only semi-simple eigenvalues. Then, by denoting* $\Phi_i \in \mathbb{C}^{n_y \times n_u}$*, the residues of the transfer function associated with the pole* $\lambda_i \in \mathbb{C}$ *(*$i = 1, \ldots, n$*), one obtains*

$$\|H\|_{\mathcal{H}_2}^2 = \sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i H(-\lambda_i)^T\right) = -\sum_{i=1}^{n}\sum_{k=1}^{n} \frac{\mathbf{tr}\left(\Phi_i\Phi_k^T\right)}{\lambda_i + \lambda_k}.$$

The hypothesis on the multiplicity of the poles in Theorem 2 can be alleviated as shown in [Antoulas, 2005, chap 5.].

2.2.2 Frequency-limited $\mathcal{H}_2$-norm

Definition and properties

The frequency-limited $\mathcal{H}_2$-norm, denoted $\mathcal{H}_{2,\Omega}$-norm, has been suggested in [Anderson et al., 1991] in order to estimate the $\mathcal{H}_2$-norm of nominally unstable models. Its definition is recalled below in Definition 4. It has been used recently in [Garulli et al., 2013] to perform robustness analysis on an aircraft model and in [Petersson, 2013] to perform optimal model approximation. Its behaviour is illustrated in Example 4.

Definition 4 ($\mathcal{H}_{2,\Omega}$-norm)**.** *Given a LTI dynamical model* $\mathbf{H}$ *whose transfer function is* $H(s) \in \mathbb{C}^{n_y \times n_u}$ *and a frequency interval* $\Omega = [0, \omega]$*, the frequency-limited* $\mathcal{H}_2$*-norm of* $\mathbf{H}$*, denoted* $\|H\|_{\mathcal{H}_{2,\Omega}}$ *is defined as the restriction of its* $\mathcal{H}_2$*-norm over* $[-\omega, \omega]$, i.e.

$$\begin{aligned} \|H\|_{\mathcal{H}_{2,\Omega}}^2 &:= \frac{1}{2\pi} \int_{-\omega}^{\omega} \mathbf{tr}\left(H(j\nu)H(-j\nu)^T\right) d\nu \\ &:= \frac{1}{\pi} \int_{\Omega} \mathbf{tr}\left(H(j\nu)H(-j\nu)^T\right) d\nu \end{aligned}$$

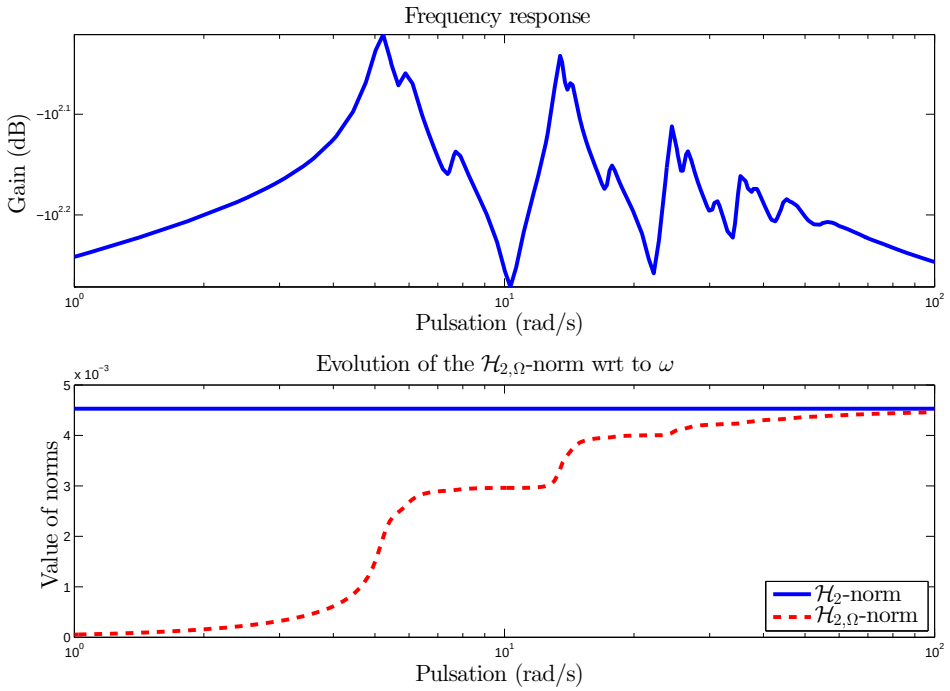


Figure 2.3: Frequency-response of the LAH model (top) and evolution of its $\mathcal{H}_{2,\Omega}$-norm with $\Omega = [0, \omega]$ with respect to ω (bottom).

Example 4 (Illustration of the behaviour of the $\mathcal{H}_{2,\Omega}$-norm (code available in Appendix C.1))**.** *To illustrate the evolution of the $\mathcal{H}_{2,\Omega}$-norm with respect to the frequency interval Ω, the $\mathcal{H}_{2,\Omega}$-norm of the LAH model is computed for the interval $\Omega = [0, \omega]$ for varying ω and plotted in Figure 2.3 together with its $\mathcal{H}_2$-norm and its frequency-response.*

One can observe that the frequency-limited $\mathcal{H}_2$-norm gives insight about the evolution of the transfer function since it increases more or less quickly depending on the gain of the transfer function. Also, in that case, the frequency-limited $\mathcal{H}_2$-norm tends towards the $\mathcal{H}_2$-norm as ω increases.

Note that using the frequency interval $\Omega = [0, \omega]$ does not yield any loss of generality. Indeed, as stated in Property 1, the $\mathcal{H}_{2,\Omega}$-norm of a model $\mathbf{H}$ over an union of multiple frequency intervals can be expressed as a sum of its $\mathcal{H}_{2,\Omega}$-norms over simpler intervals.

Property 1 (Multiple frequency intervals)**.** *Let us consider a model $\mathbf{H}$ and a frequency interval Ω defined as the union of $K \in \mathbb{N}^*$ separated frequency intervals $\Omega_i = \left[\omega_i^{(0)}, \omega_i^{(1)}\right] = \left[0, \omega_i^{(0)}\right] \bigcap \left[0, \omega_i^{(1)}\right] = \Omega_i^{(0)} \bigcap \Omega_i^{(1)}$ with $\omega_i^{(0)} < \omega_i^{(1)}$,* i.e. *:*

$$\Omega = \bigcup_{i=1}^{K} \Omega_i = \bigcup_{i=1}^{K} \left(\Omega_i^{(0)} \bigcap \Omega_i^{(1)} \right).$$

Then the $\mathcal{H}_{2,\Omega}$-norm of H can be rewritten with the frequency-limited $\mathcal{H}_2$-norms of $\mathbf{H}$ over $\Omega_i^{(0)}$ and $\Omega_i^{(1)}$ $(i = 1, \ldots, K)$ as

$$\|H\|_{\mathcal{H}_{2,\Omega}}^2 = \sum_{i=1}^{K} \|H\|_{\mathcal{H}_{2,\Omega_i^{(1)}}}^2 - \sum_{i=1}^{K} \|H\|_{\mathcal{H}_{2,\Omega_i^{(0)}}}^2.$$

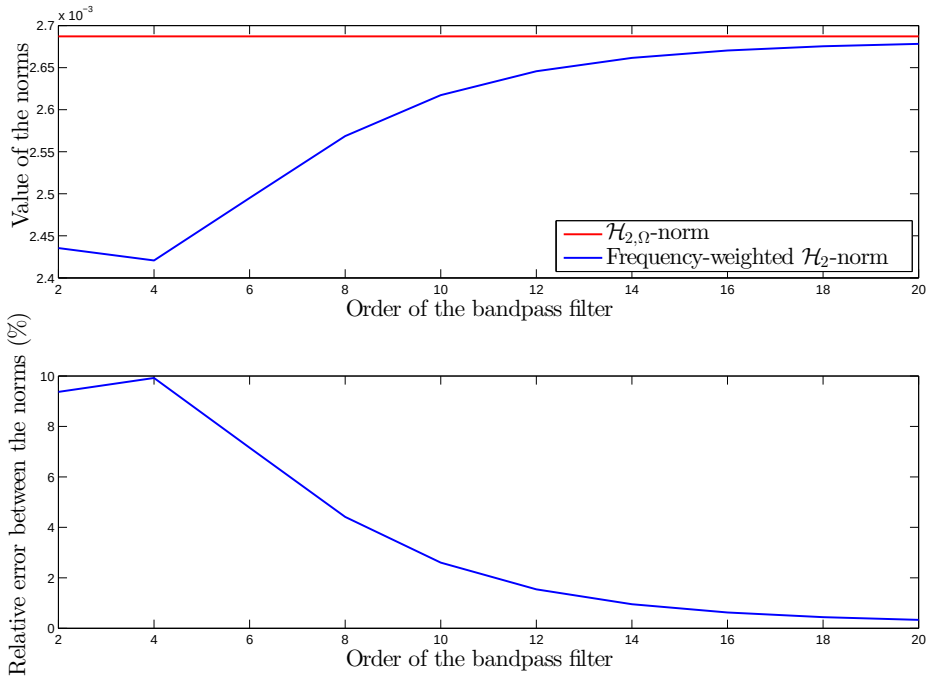


Figure 2.4: $\mathcal{H}_{2,\Omega}$-norm and frequency-weighted $\mathcal{H}_2$-norm of the LAH model computed over $\Omega = [10, 20]$ for varying order of the Butterworth filters.

A similar measurement to the frequency-limited $\mathcal{H}_2$-norm is the frequency-weighted $\mathcal{H}_2$-norm, denoted $\mathcal{H}_{2,W}$-norm (see for instance [Anić et al., 2013]). The latter is defined as the $\mathcal{H}_2$-norm of the model $\mathbf{H}$ weighted by a filter $W \in \mathcal{H}_\infty$, *i.e.*

$$\|H\|_{\mathcal{H}_{2,W}} = \|HW\|_{\mathcal{H}_2}.$$

The frequency-limited $\mathcal{H}_2$-norm is equivalent to the frequency-weighted $\mathcal{H}_2$-norm considered with perfect filters. This point is illustrated in Example 5 where both norms are compared.

Example 5 (Comparison of the $\mathcal{H}_{2,\Omega}$-norm and frequency-weighted $\mathcal{H}_2$-norm)**.** *In this example, the frequency-limited $\mathcal{H}_2$-norm of the LAH model is computed over the frequency interval $\Omega = [10, 20]$ and compared to the $\mathcal{H}_2$-norm computed on the weighted model obtained by applying an input bandpass filter to it. The filter is constructed with two Butterworth filters which orders are increased from 1 to 10. The top frame of Figure 2.4 shows the $\mathcal{H}_{2,\Omega}$-norm and frequency-weighted $\mathcal{H}_2$-norm for varying order of the bandpass filter and the bottom frame represents the relative error of the frequency-weighted $\mathcal{H}_2$-norm compared to the $\mathcal{H}_{2,\Omega}$-norm.*

The frequency-weighted $\mathcal{H}_2$-norm tends towards the $\mathcal{H}_{2,\Omega}$-norm as the order of the filter increases. With a 8-th order bandpass filter, the relative error falls below 5%. The frequency-weighted $\mathcal{H}_2$-norm is not necessarily inferior or superior to the $\mathcal{H}_{2,\Omega}$-norm, both cases can be observed, depending on the model. Note that the required order of the filter strongly depends on the considered model and the frequency interval Ω. Besides, multiple frequency intervals might be difficult to handle with filters whereas they are indifferently handled with the $\mathcal{H}_{2,\Omega}$-norm.

Computation

The frequency-limited $\mathcal{H}_2$-norm is closely related to the frequency-limited gramians introduced in [Gawronski and Juang, 1990] and recalled in Definition 2. Indeed, the frequency-limited $\mathcal{H}_2$-

norm can be computed with the frequency-limited gramians similarly to the $\mathcal{H}_2$-norm as stated in Theorem 3.

Theorem 3 (Gramian formulation of the $\mathcal{H}_{2,\Omega}$-norm)**.** *Given the frequency interval Ω and a strictly proper LTI dynamical model $\mathbf{H}$. Let $\mathcal{P}_\Omega$ and $\mathcal{Q}_\Omega$ be its frequency-limited controllability and observability gramians, then frequency-limited $\mathcal{H}_2$-norm of $\mathbf{H}$ is given as*

$$\begin{aligned} \|H\|^2_{\mathcal{H}_{2,\Omega}} &= \mathbf{tr}\left(C\mathcal{P}_\Omega C^T\right) \\ &= \mathbf{tr}\left(B^T\mathcal{Q}_\Omega B\right). \end{aligned}$$

In Theorem 3, the hypothesis is made that the model is strictly proper. This assumption can be alleviated and the formulation of the norm adapted consequently as shown in [Petersson, 2013].

2.2.3 $\mathcal{H}_\infty$-norm

The definition of the $\mathcal{H}_\infty$-norm of a LTI dynamical model is recalled below in Definition 5. For SISO models, it represents the maximum gain of the transfer function across all frequencies and for MIMO models, it is the maximum singular value of the transfer function across all frequencies.

Definition 5 ($\mathcal{H}_\infty$-norm)**.** *Given an asymptotically stable LTI dynamical model $\mathbf{H}$ whose transfer function is $H(s) \in \mathbb{C}^{n_y \times n_u}$, the $\mathcal{H}_\infty$-norm of $\mathbf{H}$, denoted $\|H\|_{\mathcal{H}_\infty}$ is defined as the maximum of the largest singular value over the imaginary axis,* i.e.

$$\|H\|_{\mathcal{H}_\infty} := \max_{\omega \in \mathbb{R}} \sigma_{max}\left(H(j\omega)\right).$$

Computing the $\mathcal{H}_\infty$-norm is a complex task that is usually achieved by an iterative bisection algorithm (see [Zhou et al., 1995, chap. 4]).

The problem that consists in finding a reduced-order model minimising the $\mathcal{H}_\infty$-norm of the error is complex and was still considered as unsolved in 2004 [Blondel and Megretski, 2004]. In fact, only few methods have been proposed to address this problem :

- the non-smooth $\mathcal{H}_\infty$ control synthesis methods presented in [Apkarian and Noll, 2006; Burke et al., 2006] can also be used to perform optimal $\mathcal{H}_\infty$ model approximation. Yet, they have not been designed for this purpose and might not be tractable in large-scale settings.
- Methods based on Linear Matrix Inequalities (LMIs) that solve a relaxation of the $\mathcal{H}_\infty$ model approximation problem have been proposed (see for instance [Grigoriadis, 1995] and reference therein). Generally, these approaches are intractable in large-scale settings and the published application examples are generally of low order ($n \leq 10$).
- An approach that adjusts the feedthrough term $\hat{D}$ of the reduced-order model in an $\mathcal{H}_\infty$-optimal way has been proposed in [Flagg et al., 2013]. This method relies on an efficient interpolation framework that is tractable even in large-scale cases but does not solve the actual $\mathcal{H}_\infty$ approximation problem.

Conclusion

In this chapter, some tools from LTI systems theory about the representation of models, their gramians and their norms, have been recalled. Concerning the representation of LTI dynamical models, the poles-residues one is convenient for addressing the optimal $\mathcal{H}_2$ approximation

problem described in Chapter 4 and is also the basis of several elements developed during this thesis.

The infinite gramians and their energetic interpretation are the foundations of the well known balanced truncation and their frequency-limited counterpart is of particular interest for model approximation over a bounded frequency interval.

The metrics used throughout this study, namely the $\mathcal{H}_2$, $\mathcal{H}_{2,\Omega}$ and $\mathcal{H}_\infty$ norms, have also been recalled. Each one of these norms has a different meaning and interest, but in the context of model approximation, the $\mathcal{H}_2$-norm has been, by far, the most considered one. Here, due to the objectives of this study, the contributions of the thesis are mainly focused on its frequency-limited counterpart.

The next chapter is dedicated at describing some well known model approximation methods which are important to understand more recent techniques.

Chapter 3

Standard model approximation methods

This chapter is aimed at recalling some well-known model approximation techniques. In particular, in Section 3.1, two methods based on the truncation of the state-space realisation, the modal truncation and the balanced truncation, are recalled and illustrated through some numerical examples. Then, in Section 3.2, model approximation methods based on implicit moments matching are presented. These methods are the basis of some of the $\mathcal{H}_2$ model approximation methods presented in the next chapter.

Note that there exist many other model approximation methods and this chapter is not meant to be an exhaustive overview of all these methods but rather a glimpse of the ones that have, in some way, influenced this study. Therefore, some techniques such as the Hankel norm approximation [Glover, 1984] are not presented in this chapter. For a much wider overview of all the approximation methods, refer to [Antoulas, 2005; Benner et al., 2005].

Contents

3.1 Model approximation by truncation 31
3.1.1 Truncation and residualisation of state-space representation 31
3.1.2 Modal and balanced truncation 32
3.2 Model approximation by moment matching 38
3.2.1 Moment matching problem 39
3.2.2 Implicit moment matching : the SISO case 40
3.2.3 Tangential interpolation 43

3.1 Model approximation by truncation

In this Section, the modal truncation and the balanced truncation are presented. They are both based on the truncation of the state-space representation which principle is recalled in Section 3.1.1. Then, the actual model approximation methods are described in Section 3.1.2 with some numerical illustrations.

3.1.1 Truncation and residualisation of state-space representation

Truncation

Given a n-th order linear time-invariant (LTI) dynamical model $\mathbf{H}$ described by the following state-space representation :

$$\mathbf{H} := \left\{ \begin{array}{rcl} \dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = & C\mathbf{x}(t) + D\mathbf{u}(t) \end{array} \right. , \tag{3.1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n_u}$, $C \in \mathbb{R}^{n_y\times n}$ and $D \in \mathbb{R}^{n_y\times n_u}$. The state vector $\mathbf{x}(t)$ can be divided between the states which must be retained $\mathbf{x_1}(t)$ and those which must be discarded $\mathbf{x_2}(t)$, *i.e.*

$$\mathbf{x}(t) = \left[\begin{array}{c} \mathbf{x_1}(t) \\ \mathbf{x_2}(t) \end{array} \right].$$

The state-space representation (3.1) can be structured accordingly to the partitioning of $\mathbf{x}(t)$,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

Then, the truncated state-space realisation $\hat{\mathbf{H}}$ of $\mathbf{H}$ is obtained by keeping only the sub-matrices associated with the states $\mathbf{x_1}(t)$, *i.e.* $\hat{\mathbf{H}} := (A_{11}, B_1, C_1, D)$. The order of $\hat{\mathbf{H}}$ is then given by the number of states which have been retained.

In general, the truncated model $\hat{\mathbf{H}}$ has no any particular property. In particular, any property such as stability, observability, etc. might be lost. Since the direct feedthrough is not affected by truncation, the only guaranty is that

$$H(\infty) = \hat{H}(\infty).$$

The state-space truncation is generally interesting when the states have some specific physical meaning or when it is performed on particular realisations such as the modal or balanced ones.

Residualisation

The state-space truncation procedure guarantees a perfect matching between $\mathbf{H}$ and $\hat{\mathbf{H}}$ at infinity. However, the steady-state error is not equal to zero, indeed,

$$H(0) - \hat{H}(0) = C_1 {A_{11}}^{-1} B_1 - C A^{-1} B.$$

If a perfect matching is required at low frequency, then state residualisation (or singular perturbation approximation) should be used instead of state truncation. It consists in saying that the dynamics of $\mathbf{x_2}(t)$ are fast compared to those of $\mathbf{x_1}(t)$, *i.e.* $\dot{\mathbf{x}}_\mathbf{2}(t) = 0$. This leads to the equation :

$$-A_{22}\mathbf{x_2}(t) = A_{21}\mathbf{x_1}(t) + B_2\mathbf{u}(t).$$

If A_{22} is nonsingular, then the reduced order model $\hat{\mathbf{H}}$ is then given by

$$\hat{\mathbf{H}} = \left(\begin{array}{c|c} A_{11} - A_{12}{A_{22}}^{-1}A_{21} & B_1 - A_{12}{A_{22}}^{-1}B_2 \\ \hline C_1 - C_2{A_{22}}^{-1}A_{21} & D - C_2{A_{22}}^{-1}B_2 \end{array}\right),$$

and with this method,

$$H(0) = \hat{H}(0).$$

This result can be explained by the fact that the state residualisation is related to the state truncation by the bilinear transformation $s \to s^{-1}$. Indeed a residualisation of $H(s)$ can be obtained by performing a state-truncation to $G(s) = H(s^{-1})$ and setting $\hat{H}(s) = \hat{G}(s^{-1})$. That is why, in the sequel, no distinction is made between truncation and residualisation. The reader should refer to [Kokotovic et al., 1986] or [Liu and Anderson, 1989] for further information on this topic.

3.1.2 Modal and balanced truncation

Modal truncation

Truncation of modal realisations is widely used by engineers since it is conceptually simple and enables to keep the meaningful modes of the initial system. For instance, high frequency modes are often discarded because they play a secondary role in the system main dynamics and they might also be out of the actuators' bandwidths.

Let consider the LTI dynamical model $\mathbf{H}$ described by the state space representation (3.1). For simplicity, let us assume that the matrix A is diagonalisable with semi-simple eigenvalues

denoted by $\lambda_i \in \mathbb{C}$, $i = 1, \ldots, n$. By projection onto the eigenspace of A, the realisation of $\mathbf{H}$ becomes

$$A_\Delta = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \quad B_\Delta = \begin{bmatrix} \mathbf{b_1} \\ \vdots \\ \mathbf{b_n} \end{bmatrix}, \quad C_\Delta = \begin{bmatrix} \mathbf{c_1}^T & \ldots & \mathbf{c_n}^T \end{bmatrix}, \tag{3.3}$$

The reduced order model $\hat{\mathbf{H}}$ is then obtained by retaining only some modes in (3.3). Let consider that the first r modes of $\mathbf{H}$ are retained, then, the error between the two models can be expressed as

$$H(s) - \hat{H}(s) = \sum_{i=r+1}^{n} \frac{\mathbf{c_i}^T \mathbf{b_i}}{s - \lambda_i},$$

where H and $\hat{H}$ are the transfer functions associated with $\mathbf{H}$ and $\hat{\mathbf{H}}$. In particular, the $\mathcal{H}_\infty$-norm of the error is given by,

$$\|H - \hat{H}\|_{\mathcal{H}_\infty} = \max_{\omega \in \mathbb{R}} \left\| \sum_{i=r+1}^{n} \frac{\mathbf{c_i}^T \mathbf{b_i}}{j\omega - \lambda_i} \right\|_2 \tag{3.4a}$$

$$\leq \max_{\omega \in \mathbb{R}} \sum_{i=r+1}^{n} \frac{\left\|\mathbf{c_i}^T \mathbf{b_i}\right\|_2}{|j\omega - \lambda_i|} \tag{3.4b}$$

$$\leq \sum_{i=r+1}^{n} \frac{\left\|\mathbf{c_i}^T \mathbf{b_i}\right\|_2}{|\mathbf{Re}\,(\lambda_i)|} \tag{3.4c}$$

Remark 5. *The bound of the approximation error (3.4c) can be very pessimistic (see Example 6). This comes from the use of the triangular inequality (from (3.4a) to (3.4b)) and the transition from (3.4b) to (3.4c) which consists in saying that the maximum of a sum of functions is smaller than the sum of each maximum.*

There are several ways to select the modes that should be retained. Engineers often select those which have a particular physical meaning. This approach implies a good knowledge of the physical system behind the equations. For more systematic selection methods, some criterion can be formulated :

- to discard modes which have a fast decay rate, the modes with the largest real part $\mathbf{Re}\,(\lambda_i)$ should be retained.
- to discard modes which have a natural high frequency, the modes with the smallest absolute value $|\lambda_i|$ should be retained.
- to minimize the $\mathcal{H}_\infty$ bound of the error (3.4c), the dominant modes, *i.e.* those with the highest ratio $\frac{\left\|\mathbf{c_i}^T \mathbf{b_i}\right\|_2}{|\mathbf{Re}(\lambda_i)|}$, should be retained.

In large-scale settings, the computation of the eigenvalues and corresponding eigenvectors can become a tedious task. Iterative algorithms that only compute some eigenvalues can be used to alleviate this issue :

- if the eigenvalues with the smallest or largest magnitude are required, then the Arnoldi or Lanczos procedures can directly be used (see Algorithm 3 in Section 3.2). Using some shift point in these algorithms enable to compute the poles in a neighbourhood of that point,
- dedicated algorithms have been developed to efficiently compute iteratively the dominant poles of a model [Rommes and Martins, 2006].

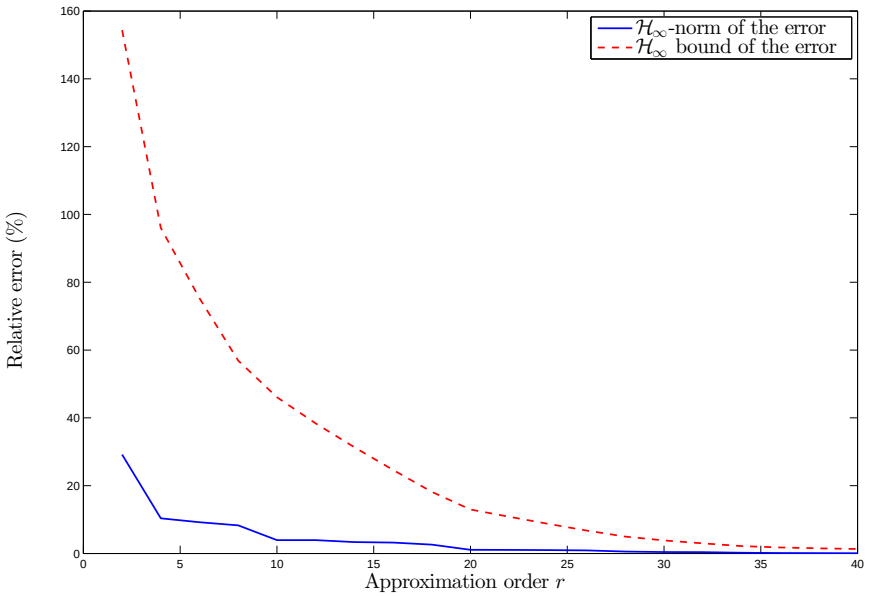


Figure 3.1: Relative $\mathcal{H}_\infty$ approximation error and its bound obtained after reduction of the ISS model with the modal truncation for r going from 2 to 40.

Example 6. *Let us consider the International Space Station (ISS) model. This model is reduced to several orders r going from 2 to 40 with the modal truncation by considering the last selection method presented above. The relative $\mathcal{H}_\infty$-norm of the approximation error and its bound,* i.e.

$$\frac{\|H - \hat{H}\|_{\mathcal{H}_\infty}}{\|H\|_{\mathcal{H}_\infty}} \quad \textit{and} \quad \frac{1}{\|H\|_{\mathcal{H}_\infty}} \sum_{i=r+1}^{n} \frac{\left\|\mathbf{c_i}^T \mathbf{b_i}\right\|_2}{|\mathbf{Re}\,(\lambda_i)|},$$

are computed for each value of r and plotted in Figure 3.1. As expected, the upper bound is very pessimistic and obviously, the smaller the reduced-order model is, the worse the bound is. In Figure 3.2, the magnitude of the frequency responses of the full-order model and reduced-order one (for $r = 10$) are plotted. Due to the modes selection criterion, the input-output transfers with the highest magnitude (like the 1^{st} input to 1^{st} output) are quite well matched while other transfers with smaller magnitude (such as the 2^{nd} input to 3^{rd} output) are not as well reproduced.

On this example, the approximation error is reasonably small which shows that modal approximation may be sufficient in some cases.

Modal approximation is a simple approximation technique that has the advantage of preserving some modes of the initial system. This is particularly interesting when some modes with a particular physical meaning must be retained. Yet, when the $\mathcal{H}_2$ or $\mathcal{H}_\infty$ norms of the approximation error are considered, modal truncation is often less efficient than other methods such as the balanced truncation described thereafter.

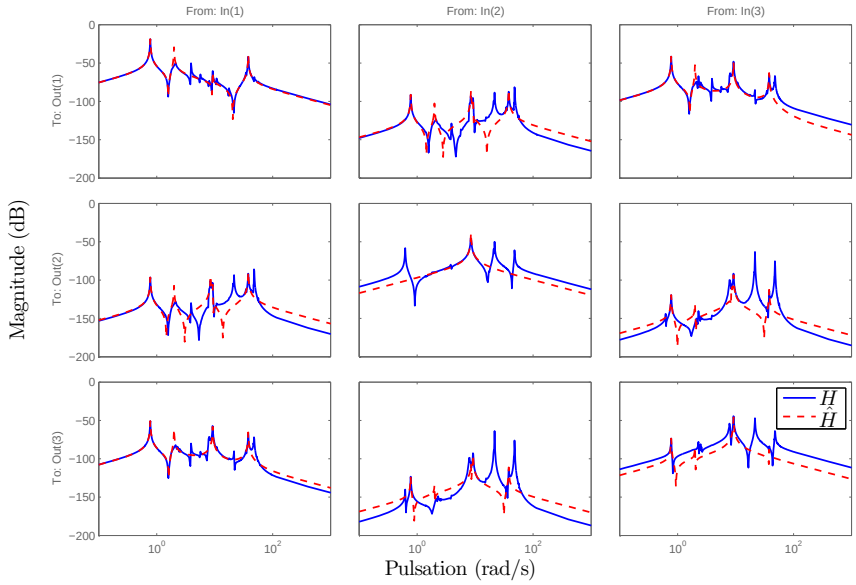


Figure 3.2: Magnitude of the frequency responses of the full-order and reduced-order models for $r = 10$.

Balanced truncation

The idea behind the balanced truncation[1] (abbreviated **BT** in the sequel) is to discard the states that require a lot of energy to be reached and/or the states that release only few energy because they play a minor role in the input to output energy transfer. As mentioned before in Section 2.1.2, the realisation in which states that are *hard to observe* are also *hard to reach* is the balanced realisation.

Once a model $\mathbf{H}$ is balanced, its associated gramians are diagonal and equals to the matrix Σ which diagonal entries are the Hankel singular values of the model (see Section 2.1.2). This matrix can be partitioned as,

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}, \tag{3.5}$$

and following this partitioning, the state-space representation of the model can be written accordingly as

$$A_{bal} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B_{bal} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C_{bal} = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \tag{3.6}$$

The reduced-order models obtained by balanced truncation are then given by,

$$\hat{\mathbf{H}}_i = \left(\begin{array}{c|c} A_{ii} & B_i \\ \hline C_i & D \end{array} \right). \tag{3.7}$$

The states that should be retained are those which correspond to the largest Hankel singular values because they have a larger impact on the input-output energy transfer. The balanced truncation has some interesting properties which are recalled in Theorem 4.

[1] Note that the Lyapunov balancing is considered here, the truncation procedure has different properties if other type of balancing are used. See [Antoulas, 2005, chap.7] or [Gugercin and Antoulas, 2004] for further details on this topic.

Theorem 4 (Balanced truncation for continuous-time models [Antoulas, 2005, chap.7]). *Given the stable, reachable and observable continuous system* $\mathbf{H}$. *The reduced-order models* $\hat{\mathbf{H}}_i$, $i = 1, 2$ (3.7) *obtained by balanced truncation have the following properties :*

1. $\hat{\mathbf{H}}_i$ *is balanced and has no pole in the open right half plane.*
2. *If the singular values contained in* Σ_1 *and* Σ_2 (3.5) *are different, then* $\hat{\mathbf{H}}_i$ *are reachable, observable and have no pole on the imaginary axis.*
3. *The* $\mathcal{H}_\infty$*-norm of the error between* $\mathbf{H}$ *and* $\hat{\mathbf{H}}_i$ *is bounded by twice the sum of the neglected Hankel singular values without their multiplicities [Pernebo and Silverman, 1982],* i.e. *if* $\hat{\mathbf{H}}_1$ *is of order* r,

$$\|H - \hat{H}_1\|_{\mathcal{H}_\infty} \leq 2 \sum_{i=r+1}^{n} \sigma_i$$

Remark 6 (About the stability of the reduced-order model). *The full proof of Theorem 4 is not detailed here and can be found in [Antoulas, 2005, chap.7], but understanding where the stability comes from helps to understand why it cannot be guaranteed with other methods presented in the sequel.*

Let us consider the reduced-order model $\hat{\mathbf{H}}_1 = (A_{11}, B_1, C_1)$ *from* (3.6). *By construction,* $\hat{\mathbf{H}}_1$ *is balanced and the Lyapunov equations*

$$A_{11}\Sigma_1 + \Sigma_1 A_{11}^T + B_1 B_1^T = 0 \quad \textit{and} \quad A_{11}^T\Sigma_1 + \Sigma_1 A_{11} + C_1^T C_1 = 0,$$

are satisfied. Then, since Σ_1 *is positive definite and* $B_1B_1^T$ *(or* $C_1^TC_1$*) positive semi-definite, inertia results (see [Ostrowski and Schneider, 1962] or [Antoulas, 2005, chap.6]) enable to conclude that* A_{11} *is stable. The asymptotic stability can then be proved by contradiction.*

Computing an exact solution to a Lyapunov equation is an ill-conditioned problem which might be difficult to achieve in large-scale settings. To overcome this issue,

- parallel computing capacities of modern computers can be exploited [Benner et al., 1999, 2000],
- low-rank solutions of the gramians can be computed [Penzl, 1999; Li and White, 2002; Stykel and Simoncini, 2012] thus leading to *approximate balanced truncation* [Sorensen and Antoulas, 2002; Badía et al., 2006; Gugercin and Li, 2005]. When approximate gramians are used, the stability of the reduced-order model is not guaranteed anymore.

Example 7. *In this example, the LAH model is approximated to several order* r *going from* 2 *to* 30 *with the modal approximation and the balanced truncation. For each order* r, *the* $\mathcal{H}_\infty$*-norm of the relative error is computed together with the corresponding upper bounds of each method and both are plotted in Figure 3.3. In Figure 3.4, the poles of the large-scale and reduced-order models are displayed for* $r = 20$.

On this example, the balanced truncation performs better than the modal truncation in terms of $\mathcal{H}_\infty$*-norm of the error. The bound given by the* ***BT*** *is also tighter than the one given by the modal truncation. In Figure 3.4, one can observe that the* ***BT*** *does not preserve the poles of the initial large-scale model unlike the modal truncation. However, the poles associated with important dynamics are generally preserved (such as those close to the imaginary axis in this case).*

Due to its simplicity, its efficiency and the fact that it offers interesting guarantees, the balanced truncation is one of the most popular model approximation methods. The method has been extended to other kind of models such as parameter-varying models [Wood et al., 1996], descriptor models [Mehrmann and Stykel, 2005], second-order linear models [Chahlaoui et al., 2006], etc. A similar concept can also be applied to non-linear models (see Remark 7).

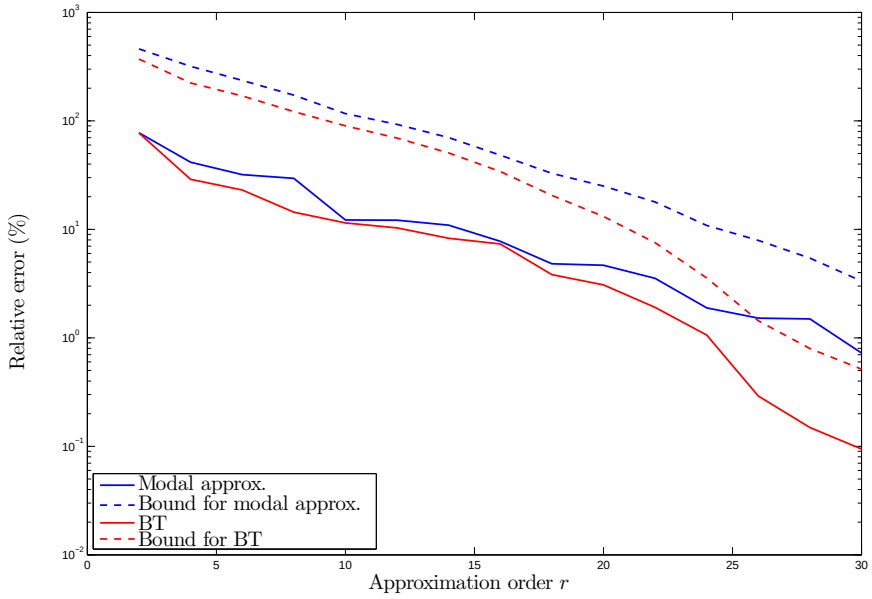


Figure 3.3: Relative $\mathcal{H}_\infty$ errors and bounds obtained with the balanced truncation and with the modal truncation on the LAH model for varying approximation order r.

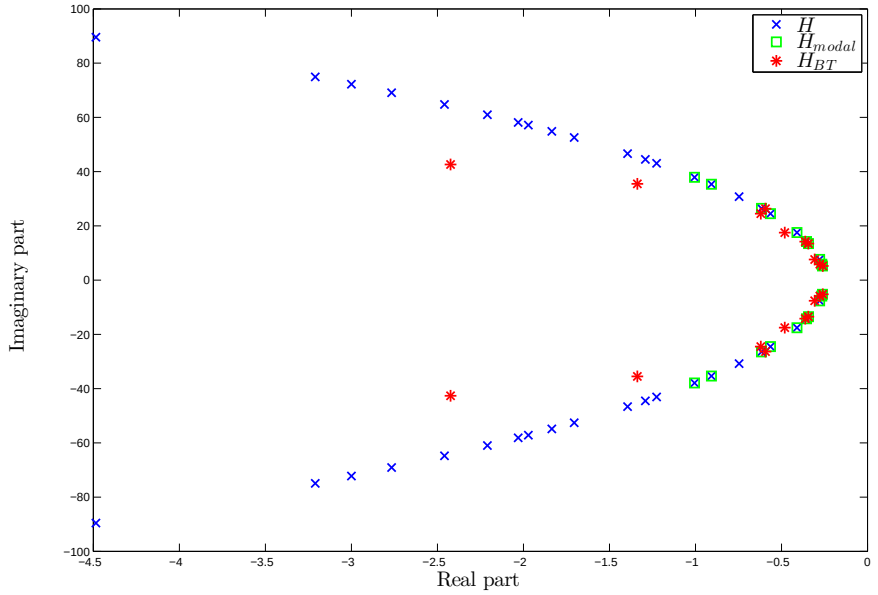


Figure 3.4: Poles of the large-scale and reduced-order models obtained by modal and balanced truncation for $r = 20$.

Remark 7 (Empirical gramians and proper orthogonal decomposition)**.** Empirical gramians *can be built from temporal simulations of a possibly non-linear model. Using these empirical gramians to project the initial model leads to the so-called* Proper Orthogonal Decomposition*. See [Antoulas, 2005, chap.9] and references therein for further information on this method. Note also that an open-source toolbox dedicated to model approximation based on empirical gramians is available in [Himpe, 2014].*

In the next Section, other projection methods are presented. The underlying idea is no longer to discard states that are barely involved in the input-output energy transfer but to match the moments of the initial transfer function at some specific interpolation points.

3.2 Model approximation by moment matching

The method described in this section greatly relies on the projection (see Remark 8 for further information on model approximation by projection) of the large-scale model over specific Krylov subspaces. These subspaces, denoted $\mathcal{K}_k$ and defined as,

$$\mathcal{K}_k\left(A, \mathbf{v}\right) = \mathbf{span}\left(\mathbf{v}, A\mathbf{v}, \ldots, A^{k-1}\mathbf{v}\right),$$

arises in many areas of research and are of particular interest for large-scale problems since they can be constructed iteratively. In particular, they are used

- to solve systems of linear equations $A\mathbf{x} = \mathbf{b}$ [Saad, 2003],
- to compute the eigenvalues of matrices [Ruhe, 1994; Sorensen, 1997],
- to compute approximate solutions of Lyapunov equations [Stykel and Simoncini, 2012],
- to reduce large-scale dynamical models [de Villemagne and Skelton, 1987; Grimme, 1997].

The last point is described in this Section. It consists in using the Krylov subspaces to perform an implicit matching of the full-order model moments at some points.

In Section 3.2.1, the moments of a transfer function are defined and the problem of moment-matching is introduced. The problem is then addressed for SISO models with a *rational interpolation* framework in Section 3.2.2. Even if this framework can theoretically be directly generalised to MIMO models, the associated algorithms become very complex. Instead, the so-called *tangential interpolation* framework is used to address the MIMO case in Section 3.2.3.

Remark 8 (Model approximation by projection)**.** *Let us consider a LTI dynamical model* $\mathbf{H} = (A, B, C, D)$*. In many cases, the trajectories* $\mathbf{x}(t)$ *of* $\mathbf{H}$ *are contained in a low dimensional subspace* $\mathcal{S}$*, that is to say* $\mathbf{x}(t)$ *can be expressed as,*

$$\mathbf{x}(t) = V_{\mathcal{S}}\hat{\mathbf{x}}(t),$$

where $V_{\mathcal{S}}$ *is a basis of* $\mathcal{S}$ *and* $\hat{\mathbf{x}}(t)$ *is the reduced state vector. The exact basis of* $\mathcal{S}$ *is in general not known, so an approximation* $V \in \mathbb{R}^{n\times r}$ *is used instead, thus,*

$$\mathbf{x}(t) \approx V\hat{\mathbf{x}}(t).$$

Then the ordinary differential equation representing $\mathbf{H}$ *becomes*

$$V\dot{\hat{\mathbf{x}}}(t) = AV\hat{\mathbf{x}}(t) + B\mathbf{u}(t) + \mathbf{res}(t), \tag{3.8}$$

where $\mathbf{res}(t)$ *is a nonzero residue which comes from the approximation of the state* $\mathbf{x}(t)$*. The* Petrov-Galerkin *conditions are then enforced by forcing the residue* $\mathbf{res}(t)$ *to be orthogonal to a subspace* $\mathcal{W}$*. By denoting* $W \in \mathbb{R}^{n\times r}$ *a basis of* $\mathcal{W}$ *and left-multiplying (3.8) by* W^T*, one obtains*

$$W^TV\dot{\hat{\mathbf{x}}}(t) = W^TAV\hat{\mathbf{x}}(t) + W^TB\mathbf{u}(t) + \underbrace{W^T\mathbf{res}(t)}_{=0}.$$

If W^TV is nonsingular, the reduced-order differential equation becomes

$$\dot{\hat{\mathbf{x}}} = \left(W^TV\right)^{-1}W^TAV\hat{\mathbf{x}}(t) + \left(W^TV\right)^{-1}W^TB\mathbf{u}(t),$$

and the associated output $\hat{\mathbf{y}}(t)$ is given by

$$\hat{\mathbf{y}} = CV\hat{\mathbf{x}}(t) + D\mathbf{u}(t).$$

If $\mathcal{W} = \mathcal{S}$, i.e. *$W = V$, then one talks about* orthogonal projection. *Otherwise, one talks about* oblique projection. *See [Saad, 2003] for further information on projection.*

The projectors W and V can then be built so that the projected model $\hat{\mathbf{H}} = (\hat{A}, \hat{B}, \hat{C}, \hat{D})$, where $\hat{A} = \left(W^TV\right)^{-1}W^TAV$, $\hat{B} = \left(W^TV\right)^{-1}W^TB$, $\hat{C} = CV$ and $\hat{D} = D$, has some specific properties. In this section, they are constructed as the basis of some Krylov subspaces to ensure interpolation conditions.

3.2.1 Moment matching problem

The moments of a transfer function are presented in Definition 6 and the problem of model approximation by moment matching is introduced in Problem 2.

Definition 6 (Moments of a transfer function)**.** *Let us consider a model $\mathbf{H}$, its transfer function $H(s) = C(sI_n - A)^{-1}B + D \in \mathbb{C}^{n_y \times n_u}$ can be decomposed through a Laurent series expansion around a given shift point $\sigma \in \mathbb{C}$ as,*

$$H(s) = \sum_{i=0}^{\infty} \eta_i(\sigma)\frac{(s-\sigma)^i}{i!},$$

where $\eta_i(\sigma) \in \mathbb{C}^{n_y \times n_u}$, $(i \in \mathbb{N})$ is the i-th moment of $H(s)$ at σ associated with the model and is defined as,

$$\eta_i(\sigma) = (-1)^i \frac{d^iH(s)}{ds^i}\Big|_{s=\sigma}$$

Problem 2 (Moment matching problem)**.** *Given a n-th order model $\mathbf{H}$ whose transfer function $H(s)$ is decomposed at $\sigma \in \mathbb{C}$ as*

$$H(s) = \sum_{i=0}^{\infty} \eta_i(\sigma)\frac{(s-\sigma)^i}{i!},$$

the moment matching problem consists in finding a reduced-order model $\hat{\mathbf{H}}$ which first r moments $\hat{\eta}_i(\sigma)$ at σ satisfy,

$$\hat{\eta}_i(\sigma) = \eta_i(\sigma),\ i = 1, \ldots, r.$$

Depending on the point σ around which the moments have to be matched, the problem have different names :

- if $\sigma = 0$, it is called a *Padé approximation* and the associated moments to match are given by
$$\eta_0(0) = -CA^{-1}B + D, \quad \text{and} \quad \eta_i(0) = (-1)^{(i+1)}CA^{-(i+1)}B,\ i > 0.$$
- if $\sigma = \infty$, then it is a *partial realisation* problem. In that case, the moments are called the *Markov parameters* and are given by
$$\eta_0(\infty) = D, \quad \text{and} \quad \eta_i(\infty) = CA^{i-1}B,\ i > 0. \tag{3.9}$$

- otherwise it is a *rational interpolation* problem.

The simplest approach to address Problem 2 consists in computing explicitly the moments of the initial model $\mathbf{H}$ and finding the reduced-order model that matches the first r ones [Grimme, 1997, chap.2]. However, computing the moments is ill-conditioned because it implies to evaluate successive power of the matrix A (see for instance equation (3.9)). Thus this method is not reliable. A more suitable approach consists in implicitly matching the moments of the full-order model by projecting it onto particular Krylov subspaces.

3.2.2 Implicit moment matching : the SISO case

An in-depth study of the methods presented in this section can be found in [Grimme, 1997]. The first way to use Krylov subspaces to perform rational interpolation consists in projecting the full-order model $\mathbf{H}$ on the Krylov subspace $\mathcal{K}_r\left((\sigma I_n - A)^{-1}, (\sigma I_n - A)^{-1}B\right)$. Indeed, as stated in Theorem 5, this enables the reduced-order model to match the first r moments of $\mathbf{H}$ at $\sigma \in \mathbb{C}$. It is called *one-sided moment matching* since only one Krylov subspace is involved.

Theorem 5 (One-sided moment matching)**.** *Let us consider a SISO LTI dynamical model* $\mathbf{H} := (A, B, C, D)$ *and an interpolation point* $\sigma \in \mathbb{C}$ *such that* $(\sigma I_n - A)$ *is nonsingular. If* $V \in \mathbb{C}^{n\times r}$ *is a full-rank matrix that columns span the Krylov subspace* $\mathcal{K}_r\left((\sigma I_n - A)^{-1}, (\sigma I_n - A)^{-1}B\right)$, i.e.

$$\mathcal{K}_r\left((\sigma I_n - A)^{-1}, (\sigma I_n - A)^{-1}B\right) \subseteq \mathcal{V} = \mathbf{span}\,(V)\,,$$

then the reduced-order model $\hat{\mathbf{H}} = \left(\left(V^TV\right)^{-1}V^TAV, \left(V^TV\right)^{-1}V^TB, CV, D\right)$ *obtained by projection matches the first* r *moments of* $\mathbf{H}$ *at* σ, i.e.

$$\hat{\eta}_i(\sigma) = \eta_i(\sigma),$$

for $i = 1, \ldots, r$.

To efficiently create a basis $V \in \mathbb{R}^{n\times r}$ for the Krylov subspace used in Theorem 5, the *Arnoldi algorithm* can be used. Its basic version is presented in Algorithm 3. This procedure enables to construct an orthonormal basis for a Krylov subspace $\mathcal{K}_r(A, \mathbf{v_1})$ without explicitly computing the successive power of the matrix A.

Algorithm 3 Arnoldi Algorithm

Require: $A \in \mathbb{C}^{n\times n}$, $\mathbf{v_1} \in \mathbb{C}^{n\times 1}$ with $\|\mathbf{v_1}\|_2 = 1$, and $r \in \mathbb{N}_+^*$.

for $i = 1, \ldots, r$ **do**
  Compute $\mathbf{w} = A\mathbf{v_i}$
  for $j = 1, \ldots, i$ **do**
    $h_{j,i} = \mathbf{v_j}^H\mathbf{w}$
  end for
  Set $\mathbf{w}_\perp = \mathbf{w} - \sum_{j=1}^{i} h_{j,i}\mathbf{v_j}$
  Set $h_{i+1,i} = \|\mathbf{w}_\perp\|_2$
  if $h_{i+1,i} = 0$ **then**
    Stop
  end if
  Set $\mathbf{v_{i+1}} = \mathbf{w}_\perp / h_{i+1,i}$
end for

At each iteration of Algorithm 3 a new vector $\mathbf{w}$ is created by multiplying the matrix A with the Arnoldi vector $\mathbf{v_i}$, and by applying a simple Gram-Schmidt procedure, this vector

$\mathbf{w}$ is then orthogonalised with respect to the previous Arnoldi vectors $\mathbf{v_j}$, $j = 1, \ldots, i$. The new Arnoldi vector $\mathbf{v_{i+1}}$ if finally obtained by normalising the orthogonal vector $\mathbf{w}_\perp$. The orthogonalisation step is very important, so a more evolved orthogonalisation procedure than the simple Gram-Schmidt one can be used. For instance, the modified Gram-Schmidt or the Householder procedures [Saad, 2003, chap. 1] are more adapted from a numerical point of view.

Remark 9. *A much more detailed presentation of the Arnoldi procedure and its properties can be found in [Saad, 2003, chap.6].*

By using the Arnoldi procedure with $A \leftarrow (\sigma I_n - A)^{-1}$ and $\mathbf{v_1} \leftarrow (\sigma I_n - A)^{-1} B$ then the resulting matrix $V \in \mathbb{C}^{n \times r}$ is a orthonormal basis for the Krylov subspace

$$\mathcal{K}_r\left((\sigma I_n - A)^{-1}, (\sigma I_n - A)^{-1} B\right).$$

Note that since the Arnoldi procedure directly creates an orthonormal basis V, the reduced-order model is obtained directly as $\hat{\mathbf{H}} = \left(V^T A V, V^T B, CV, D\right)$.

In Theorem 5, an orthogonal projection is performed, but better results can be obtained by using an oblique projection as presented in Theorem 6 (see Remark 8). Indeed, twice the number of moments of $\mathbf{H}$ are matched with a two-sided approach compared to the one-sided approach.

Theorem 6 (Two-sided moment matching)**.** *Let us consider a SISO LTI dynamical model $\mathbf{H} := (A, B, C, D)$ and an interpolation point $\sigma \in \mathbb{C}$ such that $(\sigma I_n - A)$ is nonsingular. If $V \in \mathbb{C}^{n \times r}$ and $W \in \mathbb{C}^{n \times r}$ are full-rank matrices that columns span the following Krylov subspaces,*

$$\begin{aligned} \mathcal{K}_r\left((\sigma I_n - A)^{-1}, (\sigma I_n - A)^{-1} B\right) &\subseteq \mathcal{V} &&= \mathbf{span}\,(V), \\ \mathcal{K}_r\left((\sigma I_n - A)^{-T}, (\sigma I_n - A)^{-T} C^T\right) &\subseteq \mathcal{W} &&= \mathbf{span}\,(W), \end{aligned}$$

then the reduced-order model $\hat{\mathbf{H}} = \left(\left(W^T V\right)^{-1} W^T A V, \left(W^T V\right)^{-1} W^T B, CV, D\right)$ obtained by projection matches the first $2r$ moments of $\mathbf{H}$ at σ, i.e.

$$\hat{\eta}_i(\sigma) = \eta_i(\sigma),$$

for $i = 1, \ldots, 2r$.

Similarly to the basis $V \in \mathbb{C}^{n \times r}$, $W \in \mathbb{C}^{n \times r}$ can be efficiently computed by using the Arnoldi procedure with $A \leftarrow (\sigma I_n - A)^{-T}$ and $\mathbf{v_1} \leftarrow (\sigma I_n - A)^{-T} C^T$. Alternatively, the Lanczos method can be used to directly compute both V and W in the same time. It is computationally cheaper than the Arnoldi procedure but more complex to implement (see [Saad, 2003, chap.6] or [Antoulas, 2005, chap.10]). The advantage of the two-sided moment matching approach over the one-sided one is illustrated in Example 8.

Example 8. *To compare the one-sided and two-sided moment matching approaches, the LAH model is reduced to an order 2 with $\sigma = 0$ with both techniques. The frequency responses of the error models are plotted in Figure 3.5. As expected, the two-sided method leads to better results than the one-sided method. Since the interpolation is done at $\sigma = 0$, the error is smaller at low frequency than at high frequency.*

Note that by explicitly computing the moments of the different models, we can check that the one-sided technique matches the first two moments whereas the two-sided approach matches the first four moments.

Both Theorem 5 and Theorem 6 only enable to match the moments of the full-order model $\mathbf{H}$ at one single point $\sigma \in \mathbb{C}$. For a moment-matching at several points, the generalized Krylov subspaces, which are the union of different Krylov subspaces at different points $\sigma_i \in \mathbb{C}$, $i = 1, \ldots, n_\sigma$, must be considered. The result is presented in Theorem 7.

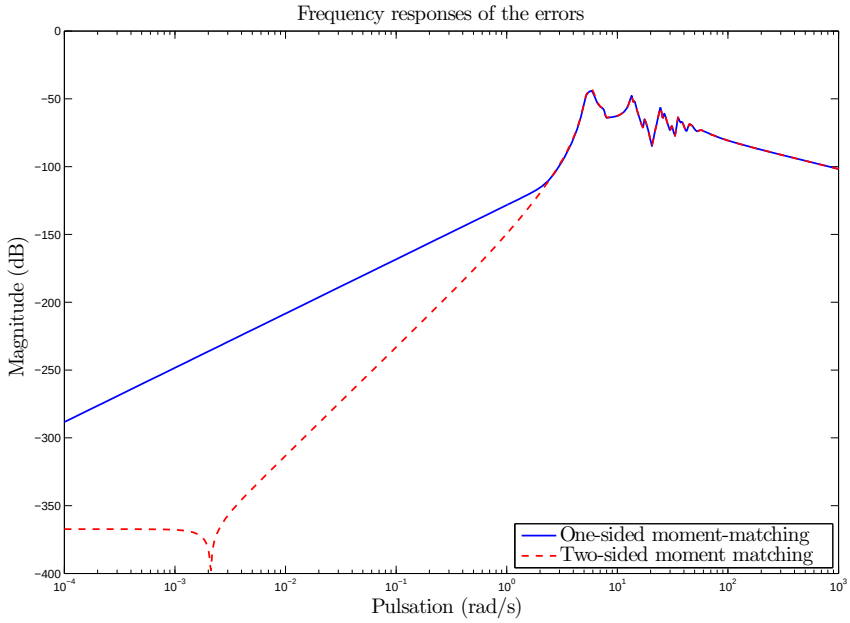


Figure 3.5: Comparison of one-sided and two-sided moment matching on the LAH model which is approximated to an order 2 at $\sigma = 0$.

Theorem 7 (Multiple points two-sided moment matching)**.** *Let us consider a SISO LTI dynamical model* $\mathbf{H} := (A, B, C, D)$ *and* n_σ *interpolation points* $\{\sigma_1, \ldots, \sigma_{n_\sigma}\} \in \mathbb{C}$ *such that each* $(\sigma_k I_n - A)$*,* $k = 1, \ldots, n_\sigma$ *is nonsingular. If* $V \in \mathbb{C}^{n \times r}$ *and* $W \in \mathbb{C}^{n \times r}$ *are full-rank matrices that columns spans the following generalized Krylov subspaces*

$$\begin{aligned} \bigcup_{k=1}^{n_\sigma} \mathcal{K}_{r_k}\left((\sigma_k I_n - A)^{-1}, (\sigma_k I_n - A)^{-1} B\right) &\subseteq \mathcal{V} &&= \mathbf{span}\,(V)\,, \\ \bigcup_{k=1}^{n_\sigma} \mathcal{K}_{r_k}\left((\sigma_k I_n - A)^{-T}, (\sigma_k I_n - A)^{-T} C^T\right) &\subseteq \mathcal{W} &&= \mathbf{span}\,(W)\,, \end{aligned}$$

then the reduced-order model $\hat{\mathbf{H}} = \left(\left(W^T V\right)^{-1} W^T A V, \left(W^T V\right)^{-1} W^T B, CV, D\right)$ *obtained by projection matches the first* $2r_k$ *moments of* $\mathbf{H}$ *in each* σ_k*,* i.e. *for* $k = 1, \ldots, n_\sigma$ *and* $i = 1, \ldots, 2r_k$*,*

$$\hat{\eta}_i(\sigma_k) = \eta_i(\sigma_k).$$

All the moment-matching theorems presented above consider complex interpolation points $\sigma_i \in \mathbb{C}$. However, to obtain a real realisation for the reduced-order model, the interpolation points must either be real or closed under conjugation. In the latter case, the columns of the matrix V and W are also closed under conjugation, thus they can be replaced by two real matrices which span the same subspace by separating the real and imaginary parts of each complex pair. For instance, consider

$$V = \begin{bmatrix} \mathbf{v_1} & \mathbf{v_1}^* \end{bmatrix} \in \mathbb{C}^{n \times 2}.$$

V spans the same subspace as

$$V_{\mathbb{R}} = \begin{bmatrix} \mathbf{Re}\,(\mathbf{v_1}) & \mathbf{Im}\,(\mathbf{v_1}) \end{bmatrix} \in \mathbb{R}^{n \times 2}.$$

This is exploited in Example 9 where the multiple points two-sided moment matching approach is compared to the single point one.

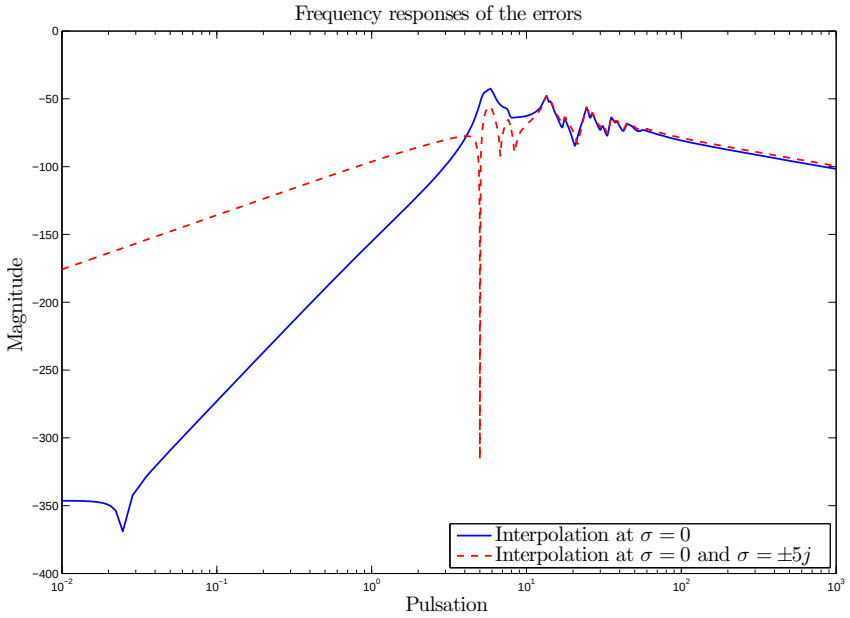


Figure 3.6: Comparison of the 3[rd]-order approximations obtained by interpolation of the LAH model at $\sigma = \{0\}$ and at $\sigma = \{0, 5j, -5j\}$.

Example 9. *This example illustrates the rational interpolation at multiple points. The LAH model is reduced to an order* 3 *by a two-sided moment matching at* $\sigma = \{0\}$ *and at* $\sigma = \{0, 5j, -5j\}$*. The reduced model obtained with a Padé approximation, denoted* $\hat{\mathbf{H}}_{padé}$*, matches the first* 6 *moments of the full-order model* $\mathbf{H}$ *at* 0 *while the one obtained with a rational interpolation at* $\sigma = \{0, 5j, -5j\}$*, denoted* $\hat{\mathbf{H}}_{rat}$*, matches the first* 2 *moments of* $\mathbf{H}$ *at each interpolation points.*

This can be seen in Figure 3.6 where the frequency responses of the error are plotted. Indeed, for low frequency, $\hat{\mathbf{H}}_{padé}$ *leads to a better approximation than* $\hat{\mathbf{H}}_{rat}$ *whereas at* $\omega = 5$ *rad/s (which corresponds to the complex point* $5j$*), the error* $H - \hat{H}_{rat}$ *falls drastically.*

All the results presented in this section can theoretically be generalised to MIMO models. However, in the latter case, the available algorithms, such as the block Arnoldi procedure [Saad, 2003, chap.6], are very tricky to implement and does not necessarily lead to satisfactory results. Instead, the *tangential interpolation* framework proposed in [Gallivan et al., 2004a] enables to elegantly and naturally address the MIMO case.

3.2.3 Tangential interpolation

An in-depth study of the tangential interpolation framework is done in [Gallivan et al., 2004a], here, only the result involved in the optimal model approximation is presented. Theorem 8 shows how a bi-tangential interpolation of a LTI model can be achieved by projection onto a particular subspace. It is very similar to the previous generalised Krylov subspaces excepted that in this case, tangential directions are involved.

Theorem 8 (Bi-tangential interpolation [Gallivan et al., 2004a])**.** *Let us consider a MIMO LTI dynamical model* $\mathbf{H}$ *with* n_u *inputs and* n_y *outputs. Let* V*,* $W \in \mathbb{C}^{n \times r}$ *be full rank matrices and let* $\sigma_i \in \mathbb{C}$*,* $\hat{\mathbf{b}}_\mathbf{i} \in \mathbb{C}^{1 \times n_u}$ *and* $\hat{\mathbf{c}}_\mathbf{i} \in \mathbb{C}^{1 \times n_y}$*,* $i = 1, \ldots, n_\sigma$ *be a given set of interpolation points and*

associated right and left tangential directions, respectively. Suppose that the points σ_i *are such that* $(\sigma_i I_n - A)$ *is nonsingular. If for* $i = 1, \ldots, n_\sigma$,

$$\begin{array}{rcl} (\sigma_i I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{i}^T & \subseteq & \mathbf{span}\,(V)\,, \\ (\sigma_i I_n - A)^{-T} C^T \hat{\mathbf{c}}_\mathbf{i}^T & \subseteq & \mathbf{span}\,(W)\,, \end{array} \tag{3.10}$$

then the reduced order model $\hat{\mathbf{H}} = \left(\left(W^T V\right)^{-1} W^T A V, \left(W^T V\right)^{-1} W^T B, CV, D \right)$ *satisfies the following bi-tangential interpolation equations,*

$$\begin{array}{rcl} \hat{\mathbf{c}}_\mathbf{i} H(\sigma_i) & = & \hat{\mathbf{c}}_\mathbf{i} \hat{H}(\sigma_i), \\ H(\sigma_i) \hat{\mathbf{b}}_\mathbf{i}^T & = & \hat{H}(\sigma_i) \hat{\mathbf{b}}_\mathbf{i}^T, \\ \hat{\mathbf{c}}_\mathbf{i} H'(\sigma_i) \hat{\mathbf{b}}_\mathbf{i}^T & = & \hat{\mathbf{c}}_\mathbf{i} \hat{H}'(\sigma_i) \hat{\mathbf{b}}_\mathbf{i}^T. \end{array}$$

for $i = 1, \ldots, n_\sigma$.

It is easy to check that a model satisfying one of the condition in equation (3.10) tangentially interpolates the initial model at σ_i from the left or the right. Indeed Let us consider r interpolation points $\sigma_i \in \mathbb{C}$ such that $(\sigma_i I_n - A)$ is non-singular and r tangential directions $\hat{\mathbf{b}}_\mathbf{i}$. Suppose that $V \in \mathbb{C}^{n \times r}$ is given as

$$V = \left[\begin{array}{ccc} (\sigma_1 I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{1}^T & \ldots & (\sigma_r I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{r}^T \end{array} \right].$$

Let us consider the matrix $Z \in \mathbb{C}^{n \times r}$ defined as

$$Z^T = \left(W^T (\sigma_i I_n - A) V\right)^{-1} W^T \left(\sigma_i I_n - A\right).$$

It is clear that $Z^T V = I_r$. The transfer function of the large-scale model evaluated at σ_i in the direction $\hat{\mathbf{b}}_\mathbf{i}$ is given by

$$H(\sigma_i) \hat{\mathbf{b}}_\mathbf{i}^T = C(\sigma_i I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{i}^T.$$

Since for $i = 1, \ldots, r$, $(\sigma_i I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{i}^T$ is a column of V, then by using Lemma 1, one obtains

$$\begin{array}{rcl} H(\sigma_1) \hat{\mathbf{b}}_\mathbf{i}^T & = & CVZ^T (\sigma_i I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{i}^T \\ & = & CV \left(W^T (\sigma_i I_n - A) V\right)^{-1} W^T B \hat{\mathbf{b}}_\mathbf{i}^T \\ & = & \hat{C} \left(\sigma_i I_r - \hat{A}\right)^{-1} \hat{B} \hat{\mathbf{b}}_\mathbf{i}^T \\ & = & \hat{H}(\sigma_i) \hat{\mathbf{b}}_\mathbf{i}^T. \end{array}$$

The proof of the whole Theorem is more complex and the reader should refer to [Gallivan et al., 2004a] for further information. The projection matrices that enable to fulfil those interpolation conditions can either be built as before with iterative procedures but they can also be obtained by solving some specific Sylvester equations [Gallivan et al., 2004b] or by contructing a Loewner matrix [Mayo and Antoulas, 2007; Ionita, 2013]. An illustration of the tangential interpolation is presented in Example 10.

Lemma 1 (See [Gallivan et al., 2004a])**.** *Let* $V \in \mathbb{C}^{n \times r}$*, if the vector* v *belongs to the column span of the matrix* V*. Then, for any matrix* $Z \in \mathbb{C}^{n \times r}$ *such that* $Z^T V = I_r$*,*

$$v = V Z^T v.$$

Example 10 (Illustration of the tangential interpolation)**.** *In this example, the ISS model, which has* 3 *inputs and* 3 *outputs, is considered. It is tangentially interpolated at the points* σ_1 *and* σ_2 *in the left directions* $\hat{\mathbf{c}}_\mathbf{1}$*,* $\hat{\mathbf{c}}_\mathbf{2}$ *and in the right directions* $\hat{\mathbf{b}}_\mathbf{1}^T$*,* $\hat{\mathbf{b}}_\mathbf{2}^T$ *with*

$$\sigma_1 = 0.8j, \quad \hat{\mathbf{b}}_\mathbf{1} = \hat{\mathbf{c}}_\mathbf{1} = \left[\begin{array}{ccc} j & 0 & 0 \end{array} \right],$$

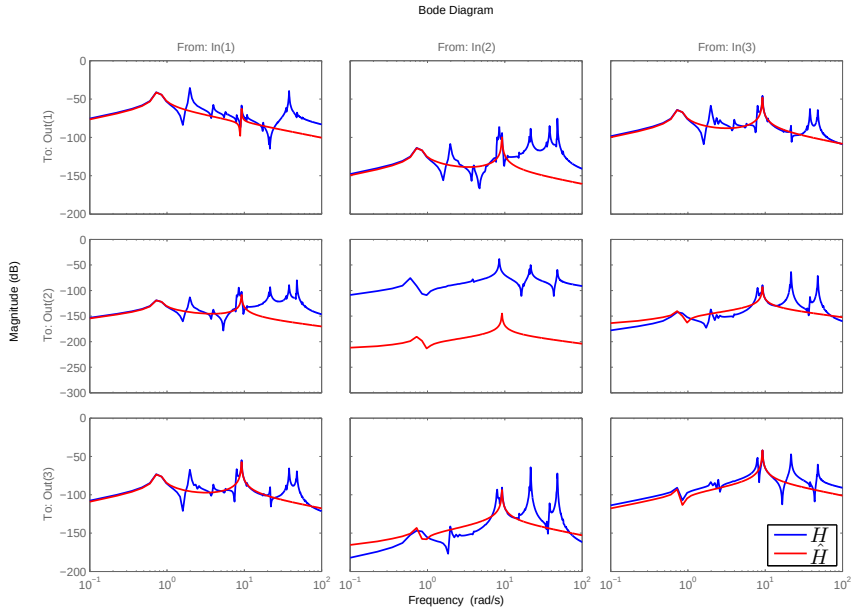


Figure 3.7: Frequency responses of the large-scale model $\mathbf{H}$ and reduced-order one $\hat{\mathbf{H}}$ obtained by tangential interpolation.

and

$$\sigma_2 = 10j, \quad \hat{\mathbf{b}}_\mathbf{2} = \hat{\mathbf{c}}_\mathbf{2} = \begin{bmatrix} 0 & 0 & j \end{bmatrix}.$$

In order to obtain a real reduced-order model $\hat{\mathbf{H}}$*, the complex conjugates of those points are considered too, thus leading to a* 4*-th order for* $\hat{\mathbf{H}}$*. With this choice of interpolation points and directions, the reduced-order transfer matrix* $\hat{H}(s)$ *should interpolate the initial transfer function matrix* $H(s)$ *according to a specific pattern. In particular, the first line and the first column of the matrix transfer* $\hat{H}(s)$ *should interpolate the first line and the first column of* $H(s)$ *at* σ_1 *and the last line and the last column of* $\hat{H}(s)$ *should interpolate the corresponding transfers of* $H(s)$ *at* σ_2*. This pattern can be represented by the following table where each entry indicates the point at which the corresponding transfers of* $\hat{H}(s)$ *and* $H(s)$ *should be matched,*

σ_1	σ_1	σ_1 *and* σ_2
σ_1	$-$	σ_2
σ_1 *and* σ_2	σ_2	σ_2

The frequency responses of both models plotted in Figure 3.7 illustrate this pattern. Indeed, one can observe for instance that the transfer from the second input to the second output is not interpolated at all. For the other interpolation points, the pattern appears more clearly when the magnitude of the error between the two transfer functions is computed at σ_1*,*

$$|H(\sigma_1) - \hat{H}(\sigma_1)| = 10^{-4} \begin{bmatrix} \mathbf{0.0000} & \mathbf{0.0000} & \mathbf{0.0000} \\ \mathbf{0.0000} & 0.1048 & 0.0003 \\ \mathbf{0.0000} & 0.0003 & 0.0581 \end{bmatrix},$$

and at σ_2*,*

$$|H(\sigma_2) - \hat{H}(\sigma_2)| = 10^{-3} \begin{bmatrix} 0.0692 & 0.0010 & \mathbf{0.0000} \\ 0.0004 & 0.3353 & \mathbf{0.0000} \\ \mathbf{0.0000} & \mathbf{0.0000} & \mathbf{0.0000} \end{bmatrix}.$$

Note that the maximum value among the terms in bold is around the machine precision.

Conclusion

In this chapter, two standard model approximation methods based on truncation have firstly been recalled. On the one hand, the modal truncation, which consists in keeping some relevant modes of the large-scale model, is conceptually simple and widespread among engineers which have a good knowledge of the physical systems below the models. On the other hand, the balanced truncation, which is based on more systematic input-to-output energy transfer considerations, is often considered as the *golden standard* of model approximation and is one of the most popular model approximation method. A lot of researches are still conducted on this method, especially to extend it to other kinds of models, such as LPV models, or to ease its numerical computation through the development of efficient sparse solvers for large-scale Lyapunov equations.

Then, a completely different framework based on the projection of the large-scale model onto specific Krylov subspaces has been introduced. It enables to implicitly match several moments of a dynamical model at several interpolation points. In practice, this framework relies on iterative algorithms that are perfectly suited for very large-scale models. The choice of the interpolation points is of main interest since it determines the properties of the reduced-order model, but it has not been discussed yet. This is done in the next chapter where the stationary points of the optimal $\mathcal{H}_2$ model approximation problem are presented as tangential interpolation conditions at the mirror images of the reduced-order model poles.

Chapter 4

Optimal $\mathcal{H}_2$ model approximation

In this chapter, the general model approximation problem formulated in the introduction (see Problem 1) is refined in Problem 3 by using the $\mathcal{H}_2$-norm to evaluate the quality of the reduced-order model leading to the so-called *optimal $\mathcal{H}_2$ model approximation problem*.

Problem 3 (optimal $\mathcal{H}_2$ model approximation problem)**.** *Given a n-th order, asymptotically stable and strictly proper continuous LTI dynamical model $\mathbf{H}$, the problem of optimal model reduction in the $\mathcal{H}_2$ norm consists in finding an asymptotically stable and strictly proper r-th order model $\hat{\mathbf{H}}$ (with $r \ll n$) which minimises the $\mathcal{H}_2$-norm of the approximation error,* i.e.

$$\hat{H} = \arg\min_{G \in \mathcal{H}_2} \underbrace{\|H - G\|_{\mathcal{H}_2}^2}_{\mathcal{J}_{\mathcal{H}_2}}.$$

Problem 3 has been extensively studied over the years, from a theoretical point of view in [Baratchart, 1986], where the existence of a solution as well as some other theoretical properties has been proven and also from a more practical point of view leading to several approaches to address it :

- approaches based on the projection of the large-scale model onto specific Krylov subspaces [Gugercin et al., 2008; Gugercin, 2007],
- approaches grounded on non-linear optimisation procedures have been developed in the continuous case [Yan and Lam, 1999; Beattie and Gugercin, 2009] and in the discrete case [Baratchart et al., 1991; Fulcheri and Olivi, 1998; Marmorat et al., 2002][1],
- hybrid methods coupling both interpolation through Loewner matrices and optimisation [Beattie and Gugercin, 2012].

Here, the approaches based on projection presented in [Gugercin et al., 2008; Gugercin, 2007] and the optimisation procedure [Beattie and Gugercin, 2009] are described. In Section 4.1, the first-order optimality conditions of the optimal $\mathcal{H}_2$ approximation problem are introduced. Then, the various algorithms are described in Section 4.2.

Contents

4.1 First order optimality conditions 48
4.1.1 $\mathcal{H}_2$ approximation error 48
4.1.2 Formulation of the first-order optimality conditions 49
4.2 Algorithms for optimal $\mathcal{H}_2$ approximation 54
4.2.1 Iterative Rational Krylov Algorithm (IRKA) 54
4.2.2 Iterative SVD Rational Krylov Algorithm (ISRKA) 58
4.2.3 Optimisation algorithm for optimal $\mathcal{H}_2$ model approximation 59

[1] The method is available in the software RARL2 [Marmorat and Olivi, 2002] at `http://www-sop.inria.fr/apics/RARL2/rarl2.html`.

4.1 First order optimality conditions

Minimizing $\mathcal{J}_{\mathcal{H}_2}$ is a non-linear and non-convex problem (see Example 11), thus finding a global minimum is a complex task. Instead, finding a local minimum is way more tractable. To this goal, first-order necessary optimality conditions have been derived :

- as interpolation conditions between the transfer functions of the full-order and reduced-order models for SISO systems with semi-simples poles only in [Meier and Luenberger, 1967],
- in a state-space form which involves the gramians of the error model in [Wilson, 1974],
- more recently in [Gugercin et al., 2008], tangential interpolation conditions have been formulated thus generalising the interpolation conditions of [Meier and Luenberger, 1967] to MIMO systems. These tangential interpolation conditions have also been presented in a very comprehensive way in [Van Dooren et al., 2008b]. Finally, they have been extended to the case of models with high order poles in [Van Dooren et al., 2010][2].

All the different formulations of the first-order optimality conditions have been shown to be equivalent and the links that exist between them has been highlighted in different ways in [Hyland and Bernstein, 1985; Gugercin et al., 2008; Van Dooren et al., 2008a].

One approach to obtain the first-order optimality conditions as interpolation conditions is described below. In particular, in Section 4.1.1, the approximation error between the large-scale and reduced-order model is formulated through the poles-residues expression of the $\mathcal{H}_2$-norm and in Section 4.1.2, the stationary points of the optimal $\mathcal{H}_2$ approximation problem are characterised.

4.1.1 $\mathcal{H}_2$ approximation error

As presented in Section 2.2.1, the $\mathcal{H}_2$-norm can be computed either using the gramians of the model or using its poles and associated residues. Since it directly leads to the interpolation conditions, the latter formulation is used. However, an equivalence can be made with the gramian formulation as shown in [Van Dooren et al., 2008b].

Let us consider asymptotically stable and strictly proper models $\mathbf{H} = (A, B, C)$ and $\hat{\mathbf{H}} = (\hat{A}, \hat{B}, \hat{C}$ of order n and r, respectively. Let us assume that both models have semi-simple poles only, *i.e.* A and $\hat{A}$ are both diagonalisable.

Under these assumptions, their transfer functions $H(s)$ and $\hat{H}(s)$ can be written as

$$H(s) = \sum_{i=1}^{n} \frac{\Phi_i}{s + \lambda_i}, \text{ and } \quad \hat{H}(s) = \sum_{k=1}^{r} \frac{\hat{\Phi}_k}{s + \hat{\lambda}_k}, \tag{4.1}$$

where λ_i, $\hat{\lambda}_k \in \mathbb{C}$, and Φ_i, $\hat{\Phi}_k \in \mathbb{C}^{n_y \times n_u}$ $(i = 1, \ldots, n$ and $k = 1, \ldots, r)$ are the poles and associated residues of the transfer functions H and $\hat{H}$, respectively.

For the model $\hat{\mathbf{H}}$ as expressed in equation (4.1) to be of order r, the residues $\hat{\Phi}_k$ must be of rank one. To explicitly show this constraint, the residues can be expressed as an outer product of vectors, *i.e.* for $k = 1, \ldots, r$,

$$\hat{\Phi}_k = \hat{\mathbf{c}}_\mathbf{k}^T \hat{\mathbf{b}}_\mathbf{k},$$

where $\hat{\mathbf{c}}_\mathbf{k} \in \mathbb{C}^{1 \times n_y}$ and $\hat{\mathbf{b}}_\mathbf{k} \in \mathbb{C}^{1 \times n_u}$. As mentioned in Section 2.1.1, the vectors $\hat{\mathbf{c}}_\mathbf{k}$ and $\hat{\mathbf{b}}_\mathbf{k}$ are linked to the state-space representation of $\hat{\mathbf{H}}$, indeed

$$\hat{\mathbf{c}}_\mathbf{k}^T = \hat{C}\hat{X}\mathbf{e}_\mathbf{k}, \text{ and } \hat{\mathbf{b}}_\mathbf{k} = \mathbf{e}_\mathbf{k}{}^T \hat{X}^{-1} \hat{B},$$

[2] A previous version of this work with additional examples is available on arXiv [Van Dooren et al., 2008a].

where $\hat{X}$ is the matrix which columns are the right eigenvectors of $\hat{A}$.

Let us also consider the $(n+r)$-th order error model $\tilde{\mathbf{H}} = \mathbf{H} - \hat{\mathbf{H}}$. The approximation error $\mathcal{J}_{\mathcal{H}_2}$ between $\mathbf{H}$ and $\hat{\mathbf{H}}$ is the square of the $\mathcal{H}_2$-norm of $\tilde{\mathbf{H}}$ and can be expressed as

$$\|\tilde{H}\|_{\mathcal{H}_2}^2 = \mathcal{J}_{\mathcal{H}_2} = \sum_{i=1}^{n+r} \mathbf{tr}\left(\tilde{\Phi}_i \tilde{H}(-\tilde{\lambda}_i)^T\right), \tag{4.2}$$

where $\tilde{\lambda}_i \in \mathbb{C}$ and $\tilde{\Phi}_i \in \mathbb{C}^{n_y \times n_u}$ are the poles and residues of the transfer function $\tilde{H}$ of the error model.

The poles and residues of $\tilde{H}$ are in fact composed of the poles and residues of the large-scale and reduced-order transfer functions H and $\hat{H}$. Let us assume that the elements are ordered as follows

$$\tilde{\Phi}_i = \begin{cases} \Phi_i & i = 1, \dots, n \\ -\hat{\Phi}_i & i = n+1, \dots, n+r, \end{cases}$$

and

$$\tilde{\lambda}_i = \begin{cases} \lambda_i & i = 1, \dots, n \\ \hat{\lambda}_i & i = n+1, \dots, n+r. \end{cases}$$

Equation (4.2) can then be reformulated as

$$\mathcal{J}_{\mathcal{H}_2} = \sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i \left(H(-\lambda_i) - \hat{H}(-\lambda_i)\right)^T\right) - \sum_{k=1}^{r} \mathbf{tr}\left(\hat{\Phi}_k \left(H(-\hat{\lambda}_k) - \hat{H}(-\hat{\lambda}_k)\right)^T\right).$$

By expanding and separating all the sums, the approximation error $\mathcal{J}_{\mathcal{H}_2}$ can finally be written as

$$\mathcal{J}_{\mathcal{H}_2} = \underbrace{\sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i H(-\lambda_i)^T\right)}_{\|H\|_{\mathcal{H}_2}^2} + \underbrace{\sum_{k=1}^{r} \mathbf{tr}\left(\hat{\Phi}_k \hat{H}(-\hat{\lambda}_k)^T\right)}_{\|\hat{H}\|_{\mathcal{H}_2}^2} - 2\sum_{i=1}^{r} \mathbf{tr}\left(H(-\hat{\lambda}_i)\hat{\Phi}_i^T\right). \tag{4.3}$$

The $\mathcal{H}_2$ approximation error is composed of the $\mathcal{H}_2$-norms of the large-scale and reduced-order models and of some cross terms between the two models.

4.1.2 Formulation of the first-order optimality conditions

The reduced-order model $\hat{\mathbf{H}}$ is represented by its poles $\hat{\lambda}_k$ and its residues decomposed in two vectors $\hat{\mathbf{c}}_\mathbf{k}$ and $\hat{\mathbf{b}}_\mathbf{k}$ $(k = 1, \dots, r)$. These parameters appear explicitly by expanding the transfer function of the reduced-order model $\hat{H}$ and the residues $\hat{\Phi}_k$ in equation (4.3)

$$\mathcal{J}_{\mathcal{H}_2} = \|H\|_{\mathcal{H}_2}^2 - \sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\hat{\mathbf{c}}_\mathbf{k}\hat{\mathbf{c}}_\mathbf{i}^T\hat{\mathbf{b}}_\mathbf{i}\hat{\mathbf{b}}_\mathbf{k}^T}{\hat{\lambda}_i + \hat{\lambda}_k} - 2\sum_{i=1}^{r} \hat{\mathbf{c}}_\mathbf{i} H(-\hat{\lambda}_i)\hat{\mathbf{b}}_\mathbf{i}^T. \tag{4.4}$$

By differentiating equation (4.4) with respect to $\hat{\lambda}_m$, $\hat{\mathbf{c}}_\mathbf{m}$ and $\hat{\mathbf{b}}_\mathbf{m}$ $(m = 1, \ldots, r)$, one obtains

$$\begin{aligned}
\frac{\partial \mathcal{J}_{\mathcal{H}_2}}{\partial \hat{\lambda}_m} &= -2\hat{\mathbf{c}}_\mathbf{m} \sum_{k=1}^{r} \frac{\hat{\mathbf{c}}_\mathbf{k}^T \hat{\mathbf{b}}_\mathbf{k}}{\left(\hat{\lambda}_m + \hat{\lambda}_k\right)^2} \hat{\mathbf{b}}_\mathbf{m}^T + 2\hat{\mathbf{c}}_\mathbf{m} \sum_{i=1}^{n} \frac{\Phi_i}{\left(\hat{\lambda}_m + \lambda_i\right)^2} \hat{\mathbf{b}}_\mathbf{m}^T \\
&= -2\hat{\mathbf{c}}_\mathbf{m} \left(\hat{H}'(-\hat{\lambda}_m) - H'(-\hat{\lambda}_m)\right) \hat{\mathbf{b}}_\mathbf{m}^T. \\
\frac{\partial \mathcal{J}_{\mathcal{H}_2}}{\partial \hat{\mathbf{c}}_\mathbf{m}} &= -2 \sum_{k=1}^{r} \frac{\hat{\mathbf{c}}_\mathbf{k}^T \hat{\mathbf{b}}_\mathbf{k}}{\hat{\lambda}_m + \hat{\lambda}_k} \hat{\mathbf{b}}_\mathbf{m}^T - 2H(-\hat{\lambda}_m) \hat{\mathbf{b}}_\mathbf{m}^T \\
&= 2\left(\hat{H}(-\hat{\lambda}_m) - H(-\hat{\lambda}_m)\right) \hat{\mathbf{b}}_\mathbf{m}^T. \\
\frac{\partial \mathcal{J}_{\mathcal{H}_2}}{\partial \hat{\mathbf{b}}_\mathbf{m}} &= -2 \sum_{k=1}^{r} \frac{\hat{\mathbf{b}}_\mathbf{k}^T \hat{\mathbf{c}}_\mathbf{k}}{\hat{\lambda}_m + \hat{\lambda}_k} \hat{\mathbf{c}}_\mathbf{m}^T - 2H(-\hat{\lambda}_m) \hat{\mathbf{c}}_\mathbf{m}^T \\
&= 2\left(\hat{H}(-\hat{\lambda}_m) - H(-\hat{\lambda}_m)\right)^T \hat{\mathbf{c}}_\mathbf{m}^T.
\end{aligned}$$

Finally, by setting $\frac{\partial \mathcal{J}_{\mathcal{H}_2}}{\partial \hat{\lambda}_m} = 0$, $\frac{\partial \mathcal{J}_{\mathcal{H}_2}}{\partial \hat{\mathbf{c}}_\mathbf{m}} = 0$ and $\frac{\partial \mathcal{J}_{\mathcal{H}_2}}{\partial \hat{\mathbf{b}}_\mathbf{m}} = 0$, the stationary points of Problem 3 can be characterised as presented in Theorem 9. See Remark 10 for some comments about the specific case of SISO models and refer to Remark 11 for further information about the more general case of models with high order poles and for some remarks about the parametrisation of this optimisation problem.

Theorem 9 (First-order optimality conditions for the $\mathcal{H}_2$ model reduction problem)**.** *Let us consider a n-th order asymptotically stable and strictly proper LTI dynamical model* $\mathbf{H}$ *whose transfer function is* $H(s) \in \mathbb{C}^{n_y \times n_u}$ *and* $\hat{\mathbf{H}}$ *a r-th order asymptotically stable and strictly proper model with semi-simple poles whose transfer function is* $\hat{H}(s) = \hat{C}(sI_r - \hat{A})^{-1} B_r \in \mathbb{C}^{n_y \times n_u}$*. If* $\hat{\mathbf{H}}$ *solves Problem 3, then, for* $i = 1, \ldots, r$*,*

$$\begin{aligned}
\hat{\mathbf{c}}_\mathbf{i} H(-\hat{\lambda}_i) &= \hat{\mathbf{c}}_\mathbf{i} \hat{H}(-\hat{\lambda}_i), \\
H(-\hat{\lambda}_i) \hat{\mathbf{b}}_\mathbf{i}^T &= \hat{H}(-\hat{\lambda}_i) \hat{\mathbf{b}}_\mathbf{i}^T, \\
\hat{\mathbf{c}}_\mathbf{i} H'(-\hat{\lambda}_i) \hat{\mathbf{b}}_\mathbf{i}^T &= \hat{\mathbf{c}}_\mathbf{i} \hat{H}'(-\hat{\lambda}_i) \hat{\mathbf{b}}_\mathbf{i}^T.
\end{aligned} \tag{4.5}$$

where $\hat{\lambda}_i \in \mathbb{C}$ *are the poles of the reduced-order model and* $\hat{\mathbf{c}}_\mathbf{i} \in \mathbb{C}^{1 \times n_y}$*,* $\hat{\mathbf{b}}_\mathbf{i} \in \mathbb{C}^{1 \times n_u}$*,* $(i = 1, \ldots, r)$ *are tangential directions given as,*

$$\begin{bmatrix} \hat{\mathbf{c}}_\mathbf{1}^T & \ldots & \hat{\mathbf{c}}_\mathbf{r}^T \end{bmatrix} = \hat{C}\hat{X},$$

and

$$\begin{bmatrix} \hat{\mathbf{b}}_\mathbf{1}^T & \ldots & \hat{\mathbf{b}}_\mathbf{r}^T \end{bmatrix}^T = \hat{X}^{-1}\hat{B},$$

where $\hat{X}$ *is the matrix which columns are the right eigenvectors of* $\hat{A}$*.*

Theorem 9 asserts that a solution of the optimal $\mathcal{H}_2$ model approximation problem must be a bi-tangential Hermite interpolant of the large-scale model $\mathbf{H}$ at the opposite of the reduced-order model poles. In [Van Dooren et al., 2010], it is proved that these first-order optimality conditions impose $r\,(n_y + n_u)$ interpolations conditions which match the number of free parameters required to represent the reduced-order model expressed as

$$\hat{H}(s) = \sum_{i=1}^{r} \frac{\hat{\mathbf{c}}_\mathbf{i}^T \hat{\mathbf{b}}_\mathbf{i}}{s - \hat{\lambda}_i},$$

with $||\hat{\mathbf{c}}_\mathbf{i}|| = 1$ for $i = 1, \ldots, r$.

The reduced-order model poles $\hat{\lambda}_i$ appear in a non-linear way in the approximation error $\mathcal{J}_{\mathcal{H}_2}$ (4.4) while each $\hat{\mathbf{c}}_\mathbf{i}$ and $\hat{\mathbf{b}}_\mathbf{i}$ appear in a quadratic way. In the SISO case, the residues can

be uniquely determined in an optimal way for fixed poles (see Remark 10) but in the MIMO case, determining the optimal residues for fixed poles require an iterative optimisation algorithm itself. In [Beattie and Gugercin, 2012], the authors have suggested to use an alternating least square method for achieving what they call *residue correction*.

Note that if either the $\hat{\mathbf{c}}_\mathbf{i}$ or the $\hat{\mathbf{b}}_\mathbf{i}$ are chosen so that one of the two first set of interpolation conditions in equation (4.5) is fulfilled then

$$\sum_{i=1}^{r} \hat{\mathbf{c}}_\mathbf{i} H(-\hat{\lambda}_i)\hat{\mathbf{b}}_\mathbf{i}^T = \sum_{i=1}^{r} \hat{\mathbf{c}}_\mathbf{i} \hat{H}(-\hat{\lambda}_i)\hat{\mathbf{b}}_\mathbf{i}^T = \|\hat{H}\|_{\mathcal{H}_2}^2,$$

thus, the $\mathcal{H}_2$ norm of the reduced-order model simplifies with the cross term in equation (4.4) and the approximation error becomes

$$\mathcal{J}_{\mathcal{H}_2} = \|H\|_{\mathcal{H}_2}^2 - \|\hat{H}\|_{\mathcal{H}_2}^2.$$

Remark 10 (First-order optimality conditions for SISO models)**.** *In the SISO case, the tangential interpolation conditions expressed in Theorem 9 become the following rational interpolation conditions :*

$$\begin{array}{rcl} H(-\hat{\lambda}_i) & = & \hat{H}(-\hat{\lambda}_i), \\ H'(-\hat{\lambda}_i) & = & \hat{H}'(-\hat{\lambda}_i) \end{array} \tag{4.6}$$

for $i = 1, \ldots, r$. In other words, the reduced-order model has to match the first two moments of the large-scale model at the mirror images of the reduced-order model poles.

In [Gugercin et al., 2008], it is shown that if $\hat{H}(s)$ interpolates $H(s)$ at the opposite of its poles, i.e. *if it satisfies the first condition in equation (4.6), then the reduced-order model $\hat{\mathbf{H}}$ is guaranteed to be an optimal $\mathcal{H}_2$ approximation of $\mathbf{H}$ among all the reduced-order models that have the same poles as $\hat{\mathbf{H}}$.*

This can also be seen from equation (4.3), indeed, in the SISO case, the scalar residues $\hat{\phi}_i \in \mathbb{C}$ are involved in a quadratic way in the approximation error. First let us rewrite the approximation error as

$$\mathcal{J}_{\mathcal{H}_2} = \|H\|_{\mathcal{H}_2}^2 - \sum_{k=1}^{r}\sum_{i=1}^{r} \frac{\hat{\phi}_k \hat{\phi}_i}{\hat{\lambda}_k + \hat{\lambda}_i} - 2\sum_{i=1}^{r} H(-\hat{\lambda}_i)\hat{\phi}_i.$$

Notice that each complex couple $\{\hat{\lambda}_k, \hat{\phi}_k\}$ appears with its complex conjugate somewhere else in the sum. Hence, by re-ordering the indexes, the double sum can also be written as,

$$\sum_{k=1}^{r}\sum_{i=1}^{r} \frac{\hat{\phi}_k \hat{\phi}_i}{\hat{\lambda}_k + \hat{\lambda}_i} = \sum_{l=1}^{r}\sum_{i=1}^{r} \frac{\hat{\phi}_l^* \hat{\phi}_i}{\hat{\lambda}_l^* + \hat{\lambda}_i}.$$

Finally, the approximation error becomes

$$\mathcal{J}_{\mathcal{H}_2} = \begin{bmatrix} \hat{\phi}_1^* & \ldots & \hat{\phi}_r^* \end{bmatrix} M \begin{bmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_r \end{bmatrix} + 2h^T \begin{bmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_r \end{bmatrix} + \|H\|_{\mathcal{H}_2}^2, \tag{4.7}$$

where the hermitian matrix M is defined as

$$M = -\begin{bmatrix} \frac{1}{2\mathbf{Re}(\hat{\lambda}_1)} & \cdots & \frac{1}{\hat{\lambda}_1^* + \hat{\lambda}_r} \\ \vdots & \ddots & \vdots \\ \frac{1}{\hat{\lambda}_r^* + \hat{\lambda}_1} & \cdots & \frac{1}{2\mathbf{Re}(\hat{\lambda}_r)} \end{bmatrix} \in \mathbb{C}^{r\times r},$$

and

$$h = -\begin{bmatrix} H(-\hat{\lambda}_1) & \ldots & H(-\hat{\lambda}_r) \end{bmatrix}^T \in \mathbb{C}^{r\times 1}.$$

The second-order term of the quadratic function (4.7) is the $\mathcal{H}_2$-norm of the reduced-order model. By denoting $\mathbf{v} = \begin{bmatrix} \hat{\phi}_1 & \dots & \hat{\phi}_r \end{bmatrix}^T \in \mathbb{C}^{r\times 1}$,

$$\mathbf{v}^H M \mathbf{v} = \|\hat{H}\|_{\mathcal{H}_2}^2 \geq 0,$$

which means that M is positive semidefinite. Hence the quadratic function $\mathcal{J}_{\mathcal{H}_2}(\mathbf{v})$ of equation (4.7) is convex which makes the necessary conditions for optimality also sufficient to find a global minimum of $\mathcal{J}_{\mathcal{H}_2}$ with respect to the residues.

To differentiate the first-order optimality conditions, let us undo the change of indexes and write the second-order term as

$$\mathbf{v}^H M \mathbf{v} = \mathbf{v}^T L \mathbf{v},$$

where the symmetric matrix L is given as,

$$L = -\begin{bmatrix} \frac{1}{2\hat{\lambda}_1} & \cdots & \frac{1}{\hat{\lambda}_1+\hat{\lambda}_r} \\ \vdots & \ddots & \vdots \\ \frac{1}{\hat{\lambda}_r+\hat{\lambda}_1} & \cdots & \frac{1}{2\hat{\lambda}_r} \end{bmatrix} \in \mathbb{C}^{r\times r}.$$

Then, differentiating the approximation error,

$$\mathcal{J}_{\mathcal{H}_2} = \mathbf{v}^T L \mathbf{v} + 2h^T \mathbf{v} + \|H\|_{\mathcal{H}_2}^2,$$

with respect to $\mathbf{v}$ yields the following optimality condition[3]

$$2L\mathbf{v} + 2h = 0. \tag{4.8}$$

Note that L is a Cauchy matrix and that solving the linear system (4.8) is equivalent to finding a rational function with simple poles $\hat{\lambda}_i$ which has fixed values (given by $-h$) at the points $-\hat{\lambda}_i$. Hence one retrieves the first part of the interpolation conditions (4.6) expressed under another form. This is illustrated in Example 11.

Example 11 (Non-convexity of the optimal $\mathcal{H}_2$ approximation problem - code available in Appendix C.2)**.** *To illustrate the non-convexity of Problem 3 with respect to the poles of the reduced-order model, the LAH model is approximated by a second order model $\hat{\mathbf{H}}$ whose transfer function is*

$$\hat{H}(s) = \frac{\hat{\phi}}{s-\hat{\lambda}} + \frac{\hat{\phi}^*}{s-\hat{\lambda}^*}$$

where the pole $\hat{\lambda} \in \mathbb{C}$ is varying and the residue $\hat{\phi} \in \mathbb{C}$ is determined in an optimal way for each value of the pole.

As explained in Remark 10, if the poles λ of the reduced-order model are fixed, then its residues ϕ can be uniquely determined in an optimal way by fulfilling the following interpolation conditions[4] *:*

$$H(-\hat{\lambda}) = \hat{H}(-\hat{\lambda}) \text{ and } H(-\hat{\lambda}^*) = \hat{H}(-\hat{\lambda}^*), \tag{4.9}$$

which is equivalent to solve

$$\begin{bmatrix} -\frac{1}{2\hat{\lambda}} & -\frac{1}{2\mathbf{Re}(\hat{\lambda})} \\ -\frac{1}{2\mathbf{Re}(\hat{\lambda})} & \frac{1}{2\hat{\lambda}^*} \end{bmatrix} \begin{bmatrix} \hat{\phi} \\ \hat{\phi}^* \end{bmatrix} = \begin{bmatrix} H(-\hat{\lambda}) \\ H(-\hat{\lambda}^*) \end{bmatrix}.$$

[3] See [Brandwood, 1983] or Section 8.2.1 for the differentiation of non-analytic functions.

[4] A similar example has been provided in [Van Dooren et al., 2008a] where $\hat{\mathbf{H}}$ is a second-order MIMO model. In that case, the residue $\hat{\phi}$ is decomposed in two vectors as $\phi = \hat{c}^T\hat{b}$ and those vectors correspond to the dominant singular vectors of $H(-\lambda)^T$.

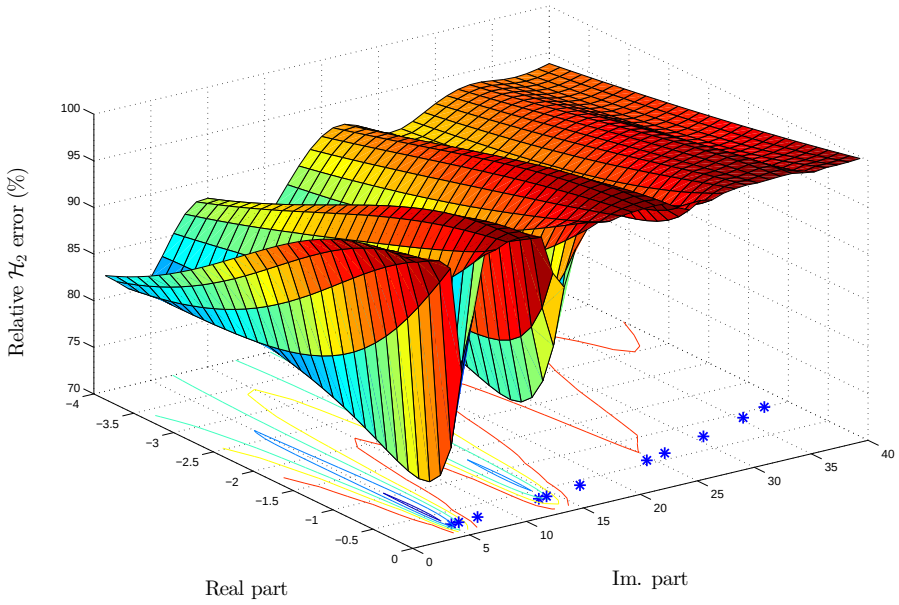


Figure 4.1: $\mathcal{H}_2$-norm of the approximation error between the LAH model and a second order model with respect to the poles of the reduced-order model. The blue stars are the poles of the large-scale model contained in this part of the complex plane.

Since the interpolation conditions (4.9) *are satisfied, the approximation error becomes*

$$\|H - \hat{H}\|_{\mathcal{H}_2}^2 = \|H\|_{\mathcal{H}_2}^2 - \|\hat{H}\|_{\mathcal{H}_2}^2,$$

where $\|\hat{H}\|_{\mathcal{H}_2}^2$ *can be expressed as,*

$$\|\hat{H}\|_{\mathcal{H}_2}^2 = -\frac{|\hat{\phi}|^2}{\mathbf{Re}\left(\hat{\lambda}\right)} - \mathbf{Re}\left(\frac{\hat{\phi}}{\hat{\lambda}}\right).$$

The resulting relative $\mathcal{H}_2$ *approximation error* $\frac{\|H-\hat{H}\|_{\mathcal{H}_2}}{\|H\|_{\mathcal{H}_2}}$ *is plotted in Figure 4.1. One can observe that the approximation error has several local minima. The poles that yield these minima seem to have imaginary parts close to the frequencies of the main dynamics of the model (see Figure 1.5).*

It is also interesting to note that the approximation error varies much more along the imaginary direction than along the real direction in which the error forms valleys. With another model such as the ISS one which has very poorly damped dynamics, those valleys are actually thinner.

Remark 11 (Case of model with high order poles and parametrisation of the reduced-order model)**.** *The result of Theorem 9 has been obtained by assuming some elements about the structure of the reduced-order model. Indeed, the reduced-order model is assumed to have semi-simple poles only. These assumptions can be alleviated by considering the Jordan decomposition of the transfer function as explained in [Van Dooren et al., 2010], in that case, the tangential directions involved in Theorem 9 become functions of the variable* s*.*

Besides, the authors show that choosing the parametrisation given in equation (4.1) *for the reduced-order model can lead to numerical issues if the reduced-order model comes close to a*

model with high order poles. Indeed, near a model with high order poles, some eigenvectors of $\hat{A}$ might become linearly dependent thus making the interpolation conditions ill-conditionned.

To avoid this issue, the authors in [Van Dooren et al., 2010] recommend to work with a non-minimal parametrisation of the reduced-order model such a its state-space representation since the $\mathcal{H}_2$-norm of the error is smooth with respect to $\hat{A}$, $\hat{B}$ and $\hat{C}$.

In [Fulcheri and Olivi, 1998; Marmorat et al., 2002], a clever parametrisation of the reduced-order model state-space representation based on its Douglas-Shapiro-Shield factorisation (see for instance [Fuhrmann, 1994]) is used. This enables to have a minimal (in terms of the number of parameters) and smooth representation that directly embeds some properties like stability.

In the next section, some of the methods that enable to fulfil these first-order optimality conditions are presented.

4.2 Algorithms for optimal $\mathcal{H}_2$ approximation

As it has been mentioned before, several approaches are now available to address the problem of optimal $\mathcal{H}_2$ approximation. In this section, the three approaches that have mainly inspired the methods developed during this study are presented :

- the MIMO version of the *Iterative Rational Krylov Algorithm* (**IRKA**) [Gugercin et al., 2008] in Section 4.2.1,
- the MIMO version of the *Iterative SVD-Rational Krylov Algorithm* (**ISRKA**) [Gugercin, 2007; Poussot-Vassal, 2011] in Section 4.2.2,
- the optimisation procedure based on the poles-residues formulation of the reduced-order model presented in [Beattie and Gugercin, 2009] in Section 4.2.3.

4.2.1 Iterative Rational Krylov Algorithm (IRKA)

This method has been proposed in [Gugercin et al., 2006] for SISO models based on the rational interpolation framework developed in [Grimme, 1997] (see Section 3.2.2). The MIMO extension has then been suggested in [Gugercin et al., 2008] based on the tangential interpolation framework developed in [Gallivan et al., 2004a] (see Section 3.2.3). It is based on the projection of the initial large-scale model on suitable Krylov subspaces in order to find a reduced-order model that fulfil the first-order optimality conditions presented in Theorem 9. The method, called *Iterative Rational Krylov Algorithm* (**IRKA**), is recalled thereafter in Algorithm 4. Note that a similar iterative interpolation scheme had been developed in [Krajewski et al., 1995] based on the frequency-domain representation of the models in terms of numerator and denominator of the associated transfer functions.

In steps 1 and 2, the initial projectors V and W are constructed from the initial interpolation parameters. Then, until the interpolation points do not vary anymore, a reduced-order model is built by projection of the large-scale one in step 5. The eigenvalues and right eigenvectors of this reduced-order model are computed at step 6 and used in steps 7 and 8 in order to build the next interpolation points and directions which lead the new projectors in steps 9 and 10.

At each iteration i and based on the Theorem 8 concerning tangential interpolation, the reduced-order model $\hat{\mathbf{H}}_i$ (step 5) tangentially interpolates the large-scale model at the opposite of the previous reduced-order model eigenvalues, *i.e.*

$$\begin{aligned}
\hat{\mathbf{c}}_{\mathbf{k}}^{(i-1)} H(-\hat{\lambda}_k^{(i-1)}) &= \hat{\mathbf{c}}_{\mathbf{k}}^{(i-1)} \hat{H}_i(-\hat{\lambda}_k^{(i-1)}), \\
H(-\hat{\lambda}_k^{(i-1)}) \hat{\mathbf{b}}_{\mathbf{k}}^{(i-1)T} &= \hat{H}_i(-\hat{\lambda}_k^{(i-1)}) \hat{\mathbf{b}}_{\mathbf{k}}^{(i-1)T}, \\
\hat{\mathbf{c}}_{\mathbf{k}}^{(i-1)} H'(-\hat{\lambda}_k^{(i-1)}) \hat{\mathbf{b}}_{\mathbf{k}}^{(i-1)T} &= \hat{\mathbf{c}}_{\mathbf{k}}^{(i-1)} \hat{H}_i'(-\hat{\lambda}_k^{(i-1)}) \hat{\mathbf{b}}_{\mathbf{k}}^{(i-1)T}.
\end{aligned}$$

Algorithm 4 Iterative Rational Krylov Algorithm (IRKA)

Require: $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n_u}$, $C \in \mathbb{R}^{n_y\times n}$, initial interpolation points $\sigma_i^{(0)} \in \mathbb{C}$ and initial tangential directions $\hat{\mathbf{c}}_\mathbf{i} \in \mathbb{C}^{1\times n_y}$ and $\hat{\mathbf{b}}_\mathbf{i} \in \mathbb{C}^{1\times n_u}$ $(i = 1, \ldots, r)$.

1: Construct,

$$\begin{aligned} V &= \begin{bmatrix} (\sigma_1^{(0)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{1}^T & \ldots & (\sigma_r^{(0)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{r}^T \end{bmatrix} \\ W &= \begin{bmatrix} (\sigma_1^{(0)} I_n - A)^{-T} C^T \hat{\mathbf{c}}_\mathbf{1}^T & \ldots & (\sigma_r^{(0)} I_n - A)^{-T} C^T \hat{\mathbf{c}}_\mathbf{r}^T \end{bmatrix} \end{aligned}$$

2: Compute $W \leftarrow W(V^T W)^{-1}$

3: **while** $\max\limits_{k=1,\ldots,r} |\sigma_k^{(i)} - \sigma_k^{(i-1)}| > \varepsilon$ **do**

4: $\quad i \leftarrow i + 1$

5: $\quad$ Project the large-scale model as

$$\hat{A} = W^T A V,\ \hat{B} = W^T B \text{ and } \hat{C} = CV.$$

6: $\quad$ Compute the eigenvalues and eigenvectors of the reduced-order model,*i.e.*

$$\hat{A}\hat{X} = \Delta\hat{X},$$

$\quad$ where $\Delta = \mathbf{diag}\left(\hat{\lambda}_1, \ldots, \hat{\lambda}_r\right)$.

7: $\quad$ Compute the new tangential directions $\hat{\mathbf{b}}_\mathbf{k}$ and $\hat{\mathbf{c}}_\mathbf{k}$ $(k = 1, \ldots, r)$ as

$$\begin{aligned} \begin{bmatrix} \hat{\mathbf{b}}_\mathbf{1}^T & \ldots & \hat{\mathbf{b}}_\mathbf{r}^T \end{bmatrix}^T &= \hat{X}^{-1}\hat{B}, \\ \begin{bmatrix} \hat{\mathbf{c}}_\mathbf{1}^T & \ldots & \hat{\mathbf{c}}_\mathbf{r}^T \end{bmatrix} &= \hat{C}\hat{X}. \end{aligned}$$

8: $\quad$ Set $\sigma_k^{(i)} = -\lambda_k(\hat{A})$, $k = 1, \ldots, r$.

9: $\quad$ Construct,

$$\begin{aligned} V &= \begin{bmatrix} (\sigma_1^{(i)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{1}^T & \ldots & (\sigma_r^{(i)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{r}^T \end{bmatrix} \\ W &= \begin{bmatrix} (\sigma_1^{(i)} I_n - A)^{-T} C^T \hat{\mathbf{c}}_\mathbf{1}^T & \ldots & (\sigma_r^{(i)} I_n - A)^{-T} C^T \hat{\mathbf{c}}_\mathbf{r}^T \end{bmatrix} \end{aligned}$$

10: $\quad$ Compute $W \leftarrow W(V^T W)^{-1}$

11: **end while**

12: Construct $\hat{\mathbf{H}} = \left(W^T A V, W^T B, CV, D\right)$

Yet, the poles of the current reduced-order model $\hat{\mathbf{H}}_i$ might be different from those of the previous reduced-order model, which means that the above interpolation conditions are not the optimality conditions. At convergence though, since the interpolation points do not vary anymore, the eigenvalues of consecutive reduced-order models do not vary either, thus the first-order optimality conditions are fulfilled.

This algorithm can be seen as a fixed-point iteration where the fixed points are the stationary points of the optimal $\mathcal{H}_2$ approximation problem. In fact, the interpolation points selection strategy $\sigma_k^{(i)} \leftarrow \lambda_k(\hat{A})$ is a relaxation of a Newton iteration proposed in [Gugercin et al., 2008] for the SISO case. Indeed, let σ denotes the set of interpolations points $\{\sigma_1, \ldots, \sigma_r\}$ and $\lambda(\sigma)$ the poles of the reduced-order model $\hat{\mathbf{H}}$ whose transfer function interpolates the large-scale model transfer function and its derivative at $\{\sigma_1, \ldots, \sigma_r\}$. Then, finding a reduced-order model that satisfies the SISO $\mathcal{H}_2$ first-order optimality conditions can be reduced to the root finding problem

$\lambda(\sigma) + \sigma = 0$. To this goal, the authors have proposed the following Newton iteration

$$\sigma^{(k+1)} = \sigma^{(k)} - (I_r + J)^{-1} \left(\sigma^{(k)} + \lambda\left(\sigma^{(k)}\right)\right), \tag{4.10}$$

where J is the Jacobian of $\lambda(\sigma)$ with respect to σ. In the neighbourhood of a stationary point, the entries of the Jacobian matrix tend to vanish and setting $J = 0$ in equation (4.10) leads to the iteration strategy of Algorithm 4. With this Newton iteration, the algorithm will converge if the initial point is close enough to a fixed-point.

The convergence of the **IRKA** with the relaxed iteration scheme $\sigma_k^{(i)} \leftarrow \lambda_k(\hat{A})$ has been proved only for the case of symmetric models in [Flagg et al., 2012]. Indeed, for these models, the method is convergent towards a fixed point and this point is a local minimiser of Problem 3.

If the algorithm converges, the resulting reduced-order model satisfies the first-order optimality conditions but it is not necessarily a minimum. To ensure that the reduced-order model is a local minimum, the method proposed in [Beattie and Gugercin, 2007] can be used. It consists in coupling the **IRKA** with an optimal interpolation point selection strategy ensuring that consecutive reduced-order models yield a decreasing $\mathcal{H}_2$ approximation error.

The following additional remarks can be addressed concerning this algorithm :

- At each iteration, the algorithm requires to solve $2r$ large-scale linear systems. These steps are the most expensive in terms of computational costs. However, efficient methods that exploit sparsity exist to address this problem even in very large-scale settings ([Saad, 2003]) and inexact solves can also be considered [Beattie et al., 2012]. This makes **IRKA** suitable for the approximation of *very* large-scale models as illustrated in Example 12.

- There is no standard way to select the initial interpolation parameters and they can be chosen randomly. When the poles of the large-scale model are available, the initial interpolation points can be set as the mirror images of the poles associated with the largest residues of the transfer function.

Example 12 (Illustration of **IRKA** on a very-large scale model)**.** *In this example, the small 3D Micropyros Thruster model available in the* Oberwolfach benchmark collection *[Korvink and Rudnyi, 2005] is considered*[5]*. This model has* 20360 *states,* 7 *outputs and* 1 *input, it is reduced to an order* 10 *with the* ***IRKA****.*

With 10 *initial real interpolation points logarithmically spaced between* 10^{-2} *and* 10^{2}*, the algorithm converges (with the tolerance* $\epsilon = 10^{-2}$*) in* 24 *iterations which took about* 20 *minutes. The frequency responses of both models are plotted in figure 4.2. The size of the large-scale model prevents from computing exactly the norm of the approximation error on a standard computer, hence, to estimate the error, the mean error between the two models at several frequencies,* i.e.

$$E = \frac{\sum_{i=1}^{n_\omega} ||H(j\omega_i) - \hat{H}(j\omega_i)||_F}{\sum_{i=1}^{n_\omega} ||H(j\omega_i)||_F},$$

is computed for $n_\omega = 70$ *logarithmically spaced frequency points between* 10^{-3} *and* 10^{3}*. In that case* $E = 0.01\%$ *which is coherent with the frequency responses of the models that match well at low frequency where the high gains are located. One can also see that the first* 4 *transfers are better matched than the last ones. This comes from the fact that the former have on average a slightly higher gain than the latter.*

Unlike approximation methods based on the Singular Value Decomposition such as the **BT**, the **IRKA** does not guarantee the stability of the reduced-order model. This drawback can be alleviated by modifying the construction of one of the two projectors thus leading to the *Iterative SVD Rational Krylov Algorithm* (**ISRKA**) presented in the next section.

[5]The model is available at `http://simulation.uni-freiburg.de/downloads/benchmark/Thruster%20%2838847%29`.

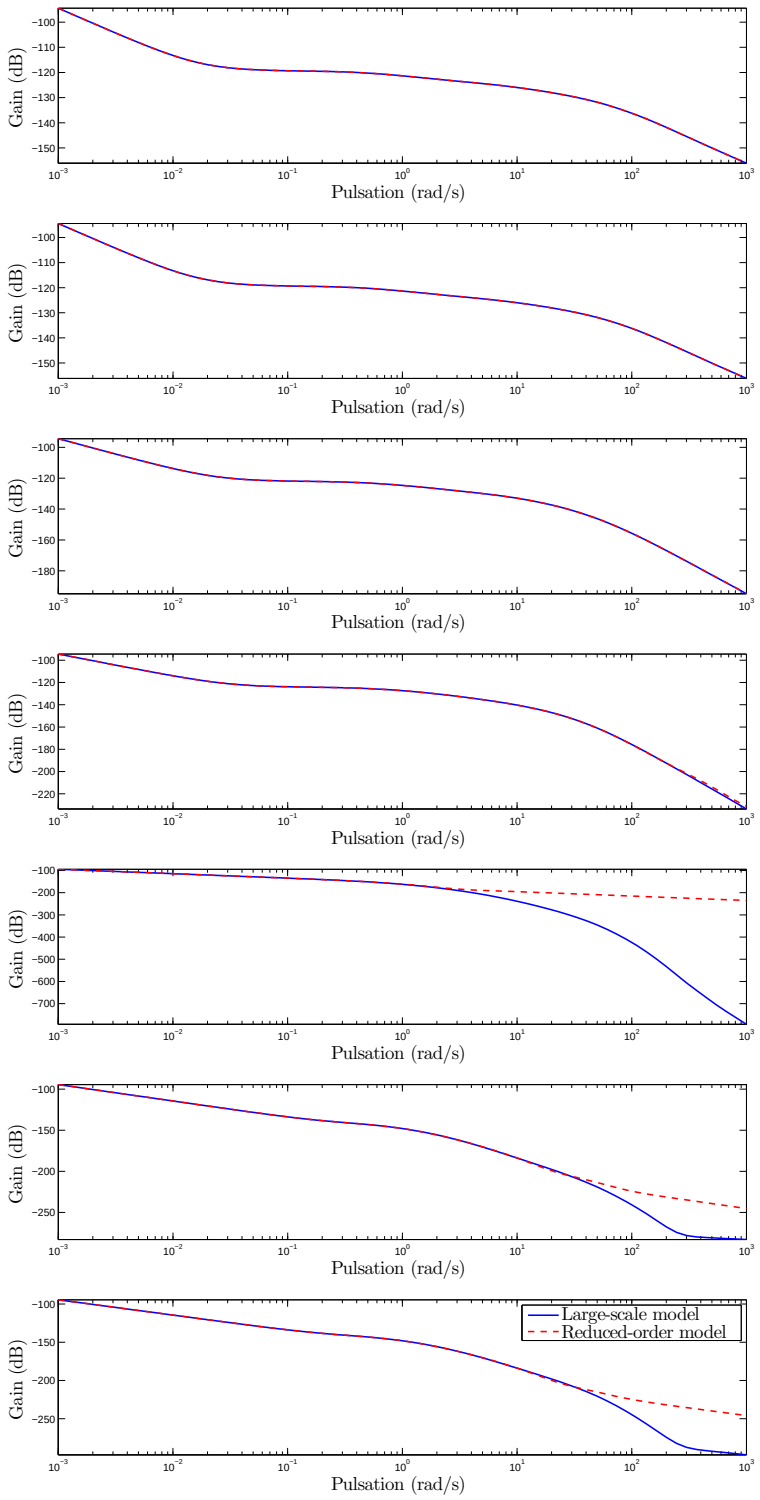


Figure 4.2: Frequency responses of the 3D Micropyros Thruster model ($n = 20360$) and its approximation to an order 10 with the **IRKA**.

4.2.2 Iterative SVD Rational Krylov Algorithm (ISRKA)

This modification of **IRKA** has been proposed in [Gugercin, 2007] for the SISO, MISO and SIMO cases and in [Poussot-Vassal, 2011] for the MIMO case (see Remark 12 for some comments on the nomenclature of the algorithm). It is aimed at ensuring the stability of the reduced-order model by using one gramian to build one of the projectors. Similar uses of the gramian can be found in [de Villemagne and Skelton, 1987] and [Yousuff et al., 1985]. The **ISRKA** is quickly presented in Algorithm 5. The only differences with the **IRKA** are the following :

- at step 2, the observability gramian $\mathcal{Q}$ of the large-scale model is computed,
- it is used to build the projector $W \in \mathbb{R}^{n\times r}$ at step 3 and 10 so that $W^TV = I_r$.

Note that the controllability gramian $\mathcal{P}$ could be used instead of $\mathcal{Q}$, in that case, the left projector W is replaced by

$$W = \begin{bmatrix} (\sigma_1^{(i)} I_n - A)^{-T} C^T \hat{\mathbf{c}}_\mathbf{1}^T & \dots & (\sigma_r^{(i)} I_n - A)^{-T} C^T \hat{\mathbf{c}}_\mathbf{r}^T \end{bmatrix}$$

at each iteration and the right projector V is built as

$$V = \mathcal{P}W\left(W^T\mathcal{P}W\right)^{-1}.$$

Using the controllability or observability gramian in the algorithm yields similar properties, only the side of the tangential interpolation changes. These properties are enunciated below by using the observability gramian.

Remark 12 (About the nomenclature of the algorithm)**.** *In the SISO, MISO and SIMO cases, the algorithm performs a rational interpolation of the large-scale model. In the MIMO case however, a tangential interpolation is performed. That is why, the MIMO version of the algorithm which relies on tangential interpolation is also called* Iterative SVD-Tangential Interpolation Algorithm *(**ISTIA**) in [Poussot-Vassal, 2011]. This name is also used in the MORE Toolbox [Poussot-Vassal and Vuillemin, 2012].*

Algorithm 5 Iterative SVD Rational Krylov Algorithm (**ISRKA**)

Require: $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n_u}$, $C \in \mathbb{R}^{n_y\times n}$, initial interpolation points $\sigma_i^{(0)} \in \mathbb{C}$ and initial tangential directions $\hat{\mathbf{b}}_\mathbf{i} \in \mathbb{C}^{1\times n_u}$ $(i = 1, \dots, r)$.

1: Construct, $V = \begin{bmatrix} (\sigma_1^{(0)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{1}^T & \dots & (\sigma_r^{(0)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{r}^T \end{bmatrix}$
2: Solve $\mathcal{Q}A + A^T\mathcal{Q} + C^TC = 0$ in $\mathcal{Q}$
3: Compute $W = \mathcal{Q}V(V^T\mathcal{Q}V)^{-1}$
4: **while** $max_{k=1,\dots,r}|\sigma_k^{(i)} - \sigma_k^{(i-1)}| > \varepsilon$ **do**
5: $\quad i \leftarrow i+1$, $\hat{A} = W^TAV$, $\hat{B} = W^TB$
6: $\quad$ Solve the eigenvalues problem $\hat{A}X = \Delta X$
7: $\quad$ Compute $\begin{bmatrix} \hat{\mathbf{b}}_\mathbf{1}^T & \dots & \hat{\mathbf{b}}_\mathbf{r}^T \end{bmatrix}^T = \hat{X}^{-1}\hat{B}$
8: $\quad$ Set $\sigma^{(i)} = -\lambda(\hat{A})$
9: $\quad$ Construct, $V = \begin{bmatrix} (\sigma_1^{(i)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{1}^T & \dots & (\sigma_r^{(i)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{r}^T \end{bmatrix}$
10: $\quad$ Compute $W = \mathcal{Q}V(V^T\mathcal{Q}V)^{-1}$
11: **end while**
12: Construct $\hat{\mathbf{H}} = \left(W^TAV, W^TB, CV, D\right)$

This modification leads to several differences concerning the properties of **ISRKA** in comparison to the **IRKA** :

- unlike the **IRKA**, if the initial large-scale model is asymptotically stable, then, the reduced-order model produced by the **ISRKA** is guaranteed to be asymptotically stable. The main idea of the proof consists in projecting the large-scale Lyapunov equation. Indeed, let us assume that the large-scale model $\mathbf{H}$ is in a basis in which $\mathcal{Q} = I_n$, then $W = V$ and the observability Lyapunov equation becomes

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

  By projecting this equation with $W^T = V^T$ and V, one obtains

$$\hat{A}^T + \hat{A} + \hat{C}^T \hat{C} = 0,$$

  which implies the stability of the reduced-order model through inertia results [Ostrowski and Schneider, 1962]. The asymptotic stability is then proved by contradiction (see [Gugercin, 2007]).

- Since the **ISRKA** only uses one Krylov subspace instead of two, the reduced-order model obtained with Algorithm 5 only interpolates the large-scale model from the right, *i.e.* at each iteration i, for $k = 1, \ldots, r$,

$$H(\sigma_k^{(i-1)})\hat{\mathbf{b}}_{\mathbf{k}}^{(i-1)^T} = \hat{H}_i(\sigma_k^{(i-1)})\hat{\mathbf{b}}_{\mathbf{k}}^{(i-1)^T}.$$

  This implies, that at convergence, the reduced-order model fulfils only one of the three first-order optimality conditions which yields weaker optimality properties than the **IRKA**. In the SISO case, the reduced-order model obtained with the **ISRKA** is the best (in the $\mathcal{H}_2$ sense) among all other models sharing the same poles. This is also valid in the SIMO case and in the MISO case if the controllability gramian is used. In the MIMO case however, the **ISRKA** leads to reduced-order models that are the best among models that have the same poles and the same left (or right) tangential directions.

- The **ISRKA** requires to solve r large-scale linear systems at each iteration instead of $2r$ but one gramian of the large-scale model needs also to be computed. Hence the method might not be tractable in very large-scale settings similarly to the **BT**. However, as with the latter method, a low-rank approximation of the gramian can be used (see Section 3.1.2). In that case, and just like with the **BT**, the stability is no longer guaranteed.

  A detailed cost comparison of the **ISRKA** with the **BT** is presented in [Gugercin, 2007]. The **ISRKA** is expected to be cheaper than the **BT** when exact gramians are used but more expensive when approximate gramians are used.

Similarly to the **IRKA**, *(i)* a Newton iteration can be used instead of the simple interpolation point selection of step 8, *(ii)* the convergence has not been proved in theory but does not seem to be an issue in practice and *(iii)* the initial interpolation parameters can be selected in various different ways.

4.2.3 Optimisation algorithm for optimal $\mathcal{H}_2$ model approximation

The method introduced in this section has been proposed in [Beattie and Gugercin, 2009]. Unlike the **IRKA** or the **ISRKA**, it is not based on projection but on a more direct optimisation approach.

The method is based on the poles-residues decomposition of the reduced-order model transfer function. Indeed, let $\hat{\lambda}_k$ and $\hat{\Phi}_k = \hat{\phi}_k \hat{\mathbf{c}}_{\mathbf{k}}^T \hat{\mathbf{b}}_{\mathbf{k}}$[6] (with $\hat{\phi}_k > 0$ and $||\hat{\mathbf{c}}_{\mathbf{k}}|| = ||\hat{\mathbf{b}}_{\mathbf{k}}|| = 1$), $k = 1, \ldots, r$

[6] For notations consistency with previous Sections, the vectors involved in the decomposition of the residues $\hat{\Phi}_k$ are line vectors, this differs from the notation in Beattie and Gugercin [2009] where column vectors are used.

denote the (semi-simple) poles and associated residues of the reduced-order model, then, the transfer function $\hat{H}(s)$ can be written as

$$\hat{H}(s) = \sum_{k=1}^{r} \frac{\hat{\phi}_k}{s - \hat{\lambda}_k} \hat{\mathbf{c}}_\mathbf{k}^T \hat{\mathbf{b}}_\mathbf{k}.$$

This representation varies slightly from what has been presented in Section 2.1.1 since the directions $\hat{\mathbf{c}}_\mathbf{k}$ and $\hat{\mathbf{b}}_\mathbf{k}$ of the residues are separated from their norm $\hat{\phi}_k$. By imposing the norm constraints $||\hat{\mathbf{c}}_\mathbf{k}|| = ||\hat{\mathbf{b}}_\mathbf{k}|| = 1$ for $k = 1, \ldots, r$, this representation is minimal. The approximation error $\mathcal{J}_{\mathcal{H}_2}$ between the large scale and reduced-order model induced by this formulation is very close to the one given in equation (4.4).

With this representation of the reduced-order model $\hat{\mathbf{H}}$, the optimal $\mathcal{H}_2$ approximation problem (Problem 3) consists in finding the reduced-order model poles $\hat{\lambda}_k$ and residues $\{\hat{\phi}_k, \hat{\mathbf{c}}_\mathbf{k}, \hat{\mathbf{b}}_\mathbf{k}\}$ $(k = 1, \ldots, r)$ that minimises the $\mathcal{H}_2$-norm of the approximation error.

More specifically, the authors have gathered the optimisation parameters in two overlapping sets : the degrees of freedom corresponding to the reduced residues $\{\hat{\phi}, \hat{\mathbf{c}}, \hat{\mathbf{b}}\}$ and the degrees of freedom corresponding to the reduced poles $\{\hat{\phi}, \hat{\lambda}\}$. This structure is developed in order to decouple the optimisation of the residues from the optimisation of the poles because the latter is *harder* than the former. This also enables to decrease the number of optimisation parameters simultaneously involved in the optimisation procedure. A conceptually similar division of the optimisation parameters is also used in [Marmorat et al., 2002] but based on a completely different representation of the reduced-order model as mentioned in Remark 11.

The optimisation scheme used by the authors is in two folds :

- Firstly, a trust-region[7] step in applied to improve the reduced poles location for *for fixed residues*. It consists in approximating the local behaviour of the approximation error $\mathcal{J}_{\mathcal{H}_2}$ by a quadratic function formed with the gradient $\nabla_{\{\hat{\phi},\hat{\lambda}\}}\mathcal{J}_{\mathcal{H}_2}$ and the Hessian $\nabla^2_{\{\hat{\phi},\hat{\lambda}\}}\mathcal{J}_{\mathcal{H}_2}$ of the error with respect to the poles degrees of freedom. Some constraints on the poles location are also added here so that the reduced-order model stays stable.

- Secondly, the residues degrees of freedom $\{\hat{\phi}, \hat{\mathbf{c}}, \hat{\mathbf{b}}\}$ are adjusted through a corrective step for fixed poles $\hat{\lambda}$. This can be done easily in the SISO case (see Remark 10) but require a optimisation procedure in the MIMO case (see for instance [Beattie and Gugercin, 2012]).

At each iteration, this method yields a stable reduced-order model which achieves a lower $\mathcal{H}_2$ approximation error than the previous one. If in addition, upon convergence, the Hessian is positive definite, then it has led to a local minimiser of Problem 3. Unlike the **IRKA** and the **ISRKA** (as presented in the previous sections), this approach offers more control over the error and more guarantees concerning the reduced-order model. These advantages are inherent to the optimisation framework.

The computation of the approximation error $\mathcal{J}_{\mathcal{H}_2}$ requires to evaluate the large-scale model transfer function at the opposite of the reduced-order model poles. Its gradient and Hessian (with respect to the poles degrees of freedom) require in addition the value of the large-scale transfer function first and second derivatives at these points, respectively. Depending on the number of trial in the trust-region step, a large number of evaluations of the objective function might be required. Hence, the computational burden of this method might be higher than for the **IRKA** depending on the number of iterations before convergence of both methods.

[7] An in-depth description of optimisation methods is out of the scope of this thesis, for further details on these techniques, we refer to standard textbooks such as [Nocedal and Wright, 1999].

Conclusion

In this chapter, the $\mathcal{H}_2$ optimal approximation problem has been introduced together with some of the numerous methods available to address it. Among those methods, the **IRKA** is probably the most suited for very large-scale model approximation. Its performances (and simplicity) make it increasingly popular and it has been extended to the approximation of other types of models such as bi-linear or descriptor models [Benner and Breiten, 2012; Gugercin et al., 2013]. However, for medium-scale model approximation, the methods based on optimisation may be more adequate since they offer more guarantee on the reduced-order model.

In some cases, it is more relevant to preserve the behaviour of the large-scale model over a bounded frequency interval only. Then, the $\mathcal{H}_2$-norm and the associated optimal model approximation methods are not completely relevant anymore and other approaches have to be considered. In the next chapter, the problem of approximation over a bounded frequency interval is introduced together with some existing methods to address it.

Chapter 5

Frequency weighted and frequency-limited model approximation

In this chapter, some methods aimed at approximating a large-scale model, so that its behaviour over a bounded frequency interval is reproduced, are briefly presented. Here, these methods are classified into two groups

- the approaches based on the use of explicit filters which are referred here as *frequency-weighted* methods and
- those that do not involve explicitly any filter, called here *frequency-limited* methods.

The former class of methods is presented in Section 5.1 while the latter is described in Section 5.2.

Contents

5.1 Frequency weighted model approximation 63
5.1.1 Frequency weighted balanced truncation 63
5.1.2 Frequency weighted $\mathcal{H}_2$ model approximation 65
5.2 Frequency-limited model approximation 67
5.2.1 Frequency-limited balanced truncation 67
5.2.2 Gramian-based $\mathcal{H}_{2,\Omega}$ optimal model approximation 69

5.1 Frequency weighted model approximation

In this section, the methods based on the use of frequency filters are presented. Firstly, the frequency-weighted extension of the balanced truncation is presented in Section 5.1.1. Then, in Section 5.1.2, the idea leading to the extension of the **IRKA** (see Section 4.2.1) to the optimal frequency-weighted $\mathcal{H}_2$ model approximation problem is introduced.

5.1.1 Frequency weighted balanced truncation

The underlying principle of frequency-weighted model truncation methods is the same as for the balanced truncation (see Section 3.1.2), *i.e.* the large-scale model is balanced with respect to some positive definite matrices of interest and the states corresponding to the smallest singular values are discarded. The difference between the various balanced truncation methods lies in the choice of the matrices with respect to which the model is balanced. Here, *frequency-weighted gramians are considered.*

Let us consider a large-scale model $\mathbf{H}$ and input and output filters $\mathbf{W_i}$ and $\mathbf{W_o}$, respectively. The objective of frequency-weighted balanced truncation methods is to find a reduced-order model $\hat{\mathbf{H}}$ so that the filtered frequency error

$$\|W_o(H - \hat{H})W_i\|_{\mathcal{H}_\infty}, \tag{5.1}$$

is small.

Note that the construction of the input and output weightings $\mathbf{W_i}$ and $\mathbf{W_o}$ is a problem in itself which is not covered here. Instead, one directly assumes that the realisations

$$\mathbf{W_i} := \left(\begin{array}{c|c} A_{W_i} & B_{W_i} \\ \hline C_{W_i} & D_{W_i} \end{array}\right) \quad \text{and} \quad \mathbf{W_o} := \left(\begin{array}{c|c} A_{W_o} & B_{W_o} \\ \hline C_{W_o} & D_{W_o} \end{array}\right)$$

are available. Based on these realisations, and assuming that there is no pole-zero cancellation, one can build the minimal realisations of $H_i(s) = H(s)W_i(s)$ and $H_o(s) = W_o(s)H(s)$ as

$$\mathbf{H_i} = \left(\begin{array}{c|c} A_i & B_i \\ \hline C_i & D_i \end{array}\right) = \left(\begin{array}{cc|c} A & BC_{W_i} & BD_{W_i} \\ 0 & A_{W_i} & B_{W_i} \\ \hline C & 0 & DD_{W_i} \end{array}\right), \tag{5.2}$$

and

$$\mathbf{H_o} = \left(\begin{array}{c|c} A_o & B_o \\ \hline C_o & D_o \end{array}\right) = \left(\begin{array}{cc|c} A & 0 & B \\ B_{W_o}C & A_{W_o} & 0 \\ \hline D_{W_o}C & C_{W_o} & D_{W_o}D \end{array}\right). \tag{5.3}$$

Let us also consider the matrices $\mathcal{Q}_o$ and $\mathcal{P}_i$ which are the solutions of

$$A_i\mathcal{P}_i + \mathcal{P}_iA_i^T + B_iB_i^T = 0 \quad \text{and} \quad A_o^T\mathcal{Q}_o + \mathcal{Q}_oA_o + C_o^TC_o = 0,$$

and which can be decomposed as,

$$\mathcal{Q}_o = \begin{bmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12} \\ \mathcal{Q}_{12}^T & \mathcal{Q}_{22} \end{bmatrix} \quad \text{and} \quad \mathcal{P}_i = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{12}^T & \mathcal{P}_{22} \end{bmatrix}.$$

$\mathcal{Q}_{11}$ and $\mathcal{P}_{11}$ are called the frequency-weighted observability and reachability gramians (see [Antoulas, 2005, Sec. 7.6] for further information). Based on these gramians, different balancing strategies can be considered,

- in [Enns, 1984], the authors have suggested to balance the model $\mathbf{H}$ such that $\mathcal{Q}_{11}$ and $\mathcal{P}_{11}$ are simultaneously diagonalised. The reduced-order model $\hat{\mathbf{H}}$ is then obtained by discarding the states associated with the smallest frequency-weighted Hankel singular values. This method does not offer any guarantee concerning the stability of the reduced-order model when two-sided weighting is considered.

  In [Kim et al., 1995], an error bound on the weighted $\mathcal{H}_\infty$-norm of the approximation error (5.1) is derived when $\hat{\mathbf{H}}$ is asymptotically stable.

- In [Lin and Chiu, 1990], the balancing method is based on the simultaneous diagonalisation of the Schur complements of $\mathcal{P}_i$ and $\mathcal{Q}_o$, *i.e.* $\mathcal{P}_{11} - \mathcal{P}_{12}{\mathcal{P}_{22}}^{-1}\mathcal{P}_{12}^T$ and $\mathcal{Q}_{11} - \mathcal{Q}_{12}{\mathcal{Q}_{22}}^{-1}\mathcal{Q}_{12}^T$. With this approach, the stability of the reduced-order model is guaranteed when the realisations (5.2) and (5.3) are minimal. An upper bound on (5.1) can also be derived with this approach.

- In [Wang et al., 1999], different quantities are built for the balancing, in particular, let us consider[1]

$$\begin{aligned} X_B &:= BC_{W_i}\mathcal{P}_{12} + \mathcal{P}_{12}^TC_{W_i}^TB^T + BD_{W_i}D_{W_i}^TB^T \\ X_C &:= \mathcal{Q}_{12}B_{W_o}C + C^TB_{W_o}^T\mathcal{Q}_{12}^T + C^TD_{W_o}^TD_{W_o}C. \end{aligned}$$

  If these two matrices are positive semi-definite, then the reduced-order model obtained by Enn's frequency weighted balanced truncation is guaranteed to be stable. Otherwise, let us consider the eigen-decompositions of X_B and X_C, *i.e.*

$$\begin{aligned} X_B &= M\mathbf{diag}\left(\lambda_1, \ldots, \lambda_n\right)M^T, \\ X_C &= N\mathbf{diag}\left(\delta_1, \ldots, \delta_n\right)N^T, \end{aligned}$$

[1] Those are the last terms in the Lyapunov equations associated with the frequency-weighted gramians $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$.

and define

$$\begin{aligned}\tilde{B} &= M\,\mathbf{diag}\left(|\lambda_1|^{\frac{1}{2}},\ldots,|\lambda_n|^{\frac{1}{2}}\right),\\ \tilde{C} &= \mathbf{diag}\left(|\delta_1|^{\frac{1}{2}},\ldots,|\delta_n|^{\frac{1}{2}}\right)N^T.\end{aligned}$$

The frequency-weighted gramians $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{Q}}$ considered in this approach are then obtained by solving

$$A\tilde{\mathcal{P}}+\tilde{\mathcal{P}}A^T+\tilde{B}\tilde{B}^T=0 \quad \text{and} \quad A^T\tilde{\mathcal{Q}}+\tilde{\mathcal{Q}}A+\tilde{C}^T\tilde{C}=0.$$

Based on this approach, the stability of the reduced-order model $\hat{\mathbf{H}}$ is guaranteed. If in addition $rank(\begin{bmatrix} B & \tilde{B}\end{bmatrix}) = rank(\tilde{B})$ and $rank(\begin{bmatrix} C^T & \tilde{C}^T\end{bmatrix}) = rank(\tilde{C}^T)$ then $\hat{\mathbf{H}}$ is asymptotically stable and there is a bound on the weighted $\mathcal{H}_\infty$ error (5.1) given by

$$\|W_o(H-\hat{H})W_i\|_{\mathcal{H}_\infty} \leq 2\|W_oL\|_{\mathcal{H}_\infty}\|KW_i\|_{\mathcal{H}_\infty}\sum_{i=r+1}^{n}\tilde{\sigma}_i,$$

where $K=\tilde{B}^TB$ and $L=C\tilde{C}^T$ and $\tilde{\sigma}_i$ are the frequency-weighted singular values associated with $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{Q}}$.

5.1.2 Frequency weighted $\mathcal{H}_2$ model approximation

In this section, the extension of the **IRKA** (see Section 4.2.1) to frequency-weighted $\mathcal{H}_2$ model approximation is presented. The first step towards this extension has been done in [Anić et al., 2013] where the first-order optimality conditions for the optimal frequency-weighted $\mathcal{H}_2$ model approximation problem have been formulated for SISO models as interpolation conditions. An heuristic approach based on the **IRKA** has also been proposed inspired by these interpolation conditions. Then in [Breiten et al., 2014], the conditions have been extended to MIMO models and the Krylov subspaces involved in **IRKA** have been modified accordingly. For sake of simplicity here, only the SISO case is considered.

The frequency-weighted $\mathcal{H}_2$-norm, denoted $\mathcal{H}_{2,W}$-norm here, of an asymptotically stable model $\mathbf{H}$ is defined as

$$\|H\|_{\mathcal{H}_{2,W}} := \|HW\|_{\mathcal{H}_2},$$

where $W(s)\in\mathcal{H}_\infty$ is the transfer function associated with a weighting model $\mathbf{W}$. The problem considered in [Breiten et al., 2014] is then to find an asymptotically stable reduced-order model $\hat{\mathbf{H}}$ of order r which minimises the frequency-weighted $\mathcal{H}_2$ approximation error with a large-scale model $\mathbf{H}$, *i.e.*

$$\hat{H}=\arg\min_{rank(G)=r}\|H-G\|_{\mathcal{H}_{2,W}}. \tag{5.4}$$

Let us assume here that $\mathbf{H}=(A,B,C,D)$, $\hat{\mathbf{H}}=(\hat{A},\hat{B},\hat{C},\hat{D})$ and $\mathbf{W}=(A_W,B_W,C_W,D_W)$ have semi-simple poles only and that their associated transfer functions are given as,

$$H(s)=\sum_{i=1}^{n}\frac{\phi_i}{s-\lambda_i}+D,\quad \hat{H}(s)=\sum_{i=1}^{r}\frac{\hat{\phi}_i}{s-\hat{\lambda}_i}+\hat{D}\quad\text{and}\quad W(s)=\sum_{i=1}^{p}\frac{\psi_i}{s-\gamma_i}+D_W. \tag{5.5}$$

where $\{\lambda_i,\phi_i\}$, $\{\hat{\lambda}_k,\hat{\phi}_k\}$ and $\{\gamma_l,\psi_l\}$ are the poles and associated residues of $\mathbf{H}$, $\hat{\mathbf{H}}$ and $\mathbf{W}$, respectively.

First-order optimality conditions for the $\mathcal{H}_{2,W}$ approximation problem

Let us define, for an asymptotically stable model $\mathbf{H}$ and a filter $\mathbf{W}$ such that $DD_W=0$,

$$\mathscr{F}[H](s)=H(s)W(s)W(-s)+\sum_{k=1}^{p}H(-\gamma_k)W(-\gamma_k)\frac{\psi_k}{s+\gamma_k}.$$

If the reduced-order model $\hat{\mathbf{H}}$ is a solution to (5.4), then, for $k = 1, \ldots, r$,

$$\begin{array}{rcl} \mathscr{F}[H](-\hat{\lambda}_k) & = & \mathscr{F}[\hat{H}](-\hat{\lambda}_k), \\ \mathscr{F}'[H](-\hat{\lambda}_k) & = & \mathscr{F}'[\hat{H}](-\hat{\lambda}_k), \end{array} \tag{5.6}$$

and in addition, for all $\mathbf{n} \in Ker(D_W^T)$

$$\left(\int_{-\infty}^{\infty} \mathscr{F}[H](j\omega)d\omega\right)\mathbf{n} = \left(\int_{-\infty}^{\infty} \mathscr{F}[\hat{H}](j\omega)d\omega\right)\mathbf{n}$$

For further information on $\mathscr{F}[.](s)$ and on these first-order optimality conditions, refer to [Breiten et al., 2014].

Use of IRKA for $\mathcal{H}_{2,W}$ model approximation

The first-order optimality conditions (5.6) are similar to the optimality conditions arising in $\mathcal{H}_2$ model approximation (see Section 4.1) which suggest that an iterative scheme of the **IRKA** can be used for the $\mathcal{H}_{2,W}$ model approximation problem as well. Two approaches have been considered,

- based on the notations of equation (5.5), the $\mathcal{H}_{2,W}$-norm of the approximation error between $\mathbf{H}$ and $\hat{\mathbf{H}}$ can be written as,

$$\begin{array}{rcl} \|H - \hat{H}\|_{\mathcal{H}_{2,W}}^2 & = & \displaystyle\sum_{i=1}^{n} \left(H(-\lambda_i) - \hat{H}(-\lambda_i)\right) W(-\lambda_i)W(\lambda_i)\phi_i \\ & & \displaystyle + \sum_{k=1}^{r} \left(\hat{H}(-\hat{\lambda}_k) - H(-\hat{\lambda}_k)\right) W(-\hat{\lambda}_k)W(\hat{\lambda}_k)\hat{\phi}_k \\ & & \displaystyle + \sum_{l=1}^{p} \left(H(-\gamma_l) - \hat{H}(-\gamma_l)\right) W(-\gamma_l)\left(H(\gamma_l) - \hat{H}(\gamma_l)\right)\psi_l. \end{array} \tag{5.7}$$

  In [Anić et al., 2013], a specific choice of interpolation points is used in **IRKA** in order to make some elements of the error (5.7) vanish. In particular, one of the projector in **IRKA** is built so that at convergence $\hat{H}(-\hat{\lambda}_k) = H(-\hat{\lambda}_k)$ (which makes the second sum vanish) while the other projector is built with fixed interpolation points so that $\hat{H}$ interpolates H at the mirror images of the poles λ_i and γ_l associated with the highest residues ϕ_i and ψ_l (which makes some elements of the first and third sums in (5.7) vanish).

  However, with this approach, the first-order optimality conditions (5.6) are not fulfilled.

- In [Breiten et al., 2014], it is proven that $\mathscr{F}[H](s)$ has a realisation given by

$$\mathscr{F}[H] = \left(\begin{array}{c|c} A_{\mathscr{F}} & B_{\mathscr{F}} \\ \hline C_{\mathscr{F}} & D_{\mathscr{F}} \end{array}\right) = \left(\begin{array}{cc|c} A & BC_W & ZC_W^T + BD_W D_W^T \\ 0 & A_W & \mathcal{P}_W C_W^T + B_W D_W^T \\ \hline C & DC_W & 0 \end{array}\right),$$

  where $\mathcal{P}_W$ is the reachability gramian associated with $\mathbf{W}$ and Z is the solution of

$$AZ + ZA_W^T + B(C_W\mathcal{P}_W + D_W B_W^T) = 0.$$

  Based on this state-space representation, the construction of the projectors W and V in **IRKA** has been modified so that at convergence, if $Ran(Z) \subset Ran(V)$, then the reduced-order model $\hat{\mathbf{H}}$ fulfils the first-order optimality conditions (5.6).

  The assumption $Ran(Z) \subset Ran(V)$ is not necessarily fulfilled by the algorithm proposed in [Breiten et al., 2014] which is why it is called *Near Optimal Weighted Interpolation*.

Indeed, an error might appear depending on the difference $Z - V\hat{Z}$ where $\hat{Z}$ is the solution of

$$\hat{A}\hat{Z} + \hat{Z}A_W^T + \hat{B}(C_W\mathcal{P}_W + D_W B_W^T) = 0,$$

and can be viewed as the Petrov-Galerkin approximation of Z. The interpolation error induced by $Z - V\hat{Z}$ is expected to vanish as r increases.

5.2 Frequency-limited model approximation

In this section, model approximation methods that do not require the use of filters are presented. In particular, in Section 5.2.1, the Frequency-Limited Balanced Truncation (**FL-BT**) is described and in Section 5.2.2, an optimal $\mathcal{H}_{2,\Omega}$ model approximation method is presented.

5.2.1 Frequency-limited balanced truncation

In [Gawronski and Juang, 1990], the authors have proposed to use the restriction of the frequency definition of the gramians, referred here as the frequency-limited gramians (see Section 2.1.2), to balance the initial large-scale model.

For the frequency interval $\Omega = [\omega_1, \omega_2]$, the frequency-limited observability and reachability gramians $\mathcal{Q}_\Omega$ and $\mathcal{P}_\Omega$, respectively, can be obtained, similarly to the infinite gramians, by solving two Lyapunov equations,

$$\begin{aligned} A\mathcal{P}_\Omega + \mathcal{P}_\Omega A^T + W_c(\Omega) &= 0, \\ A^T\mathcal{Q}_\Omega + \mathcal{Q}_\Omega A + W_o(\Omega) &= 0, \end{aligned} \tag{5.8}$$

where $W_c(\Omega) = W_c(\omega_2) - W_c(\omega_1)$ and $W_o(\Omega) = W_o(\omega_2) - W_o(\omega_1)$ with

$$W_c(\omega) = S(\omega)BB^T + BB^T S^*(\omega) \quad \text{and} \quad W_o(\omega) = S^*(\omega)C^TC + C^TCS(\omega),$$

and where $S(\omega) = \frac{j}{2\pi}\mathbf{logm}((A + j\omega I_n)(A - j\omega I_n)^{-1})$.

By definition (see Section 2.1.2), the frequency-limited gramians $\mathcal{Q}_\Omega$ and $\mathcal{P}_\Omega$ are positive semi-definite. Hence, they can be balanced (*i.e.* , simultaneously diagonalised) so that in this basis, they are both equal and diagonal,

$$\mathcal{Q}_\Omega = \mathcal{P}_\Omega = \mathbf{diag}\left(\sigma_1^\Omega, \ldots, \sigma_n^\Omega\right),$$

where σ_i^Ω are what may be called here the frequency-limited Hankel singular values (see Example 3 for an illustration of those singular values).

The reduced-order model is then obtained by discarding the states of the large-scale model associated with the smallest frequency-limited Hankel singular values. This approach can be shown (see [Gugercin and Antoulas, 2004]) to be equivalent to the frequency-weighted balanced truncation method presented in Section 5.1.1 considered with perfect filters. The advantage here is that no explicit weight is involved.

Note that since $W_c(\Omega)$ and $W_o(\Omega)$ are not necessarily positive semi-definite, the reduced-order model obtained with the **FL-BT** is not guaranteed to be stable (see Remark 6). In addition, there is no upper bound on the $\mathcal{H}_\infty$-norm of the approximation error. Both issues have been alleviated by slightly modifying the method.

Modification of the FL-BT to ensure stability preservation

In [Gugercin and Antoulas, 2004], the authors have proposed a modification of the **FL-BT** in order to ensure the stability of the reduced-order model. The idea is similar to the modification of the frequency-weighted balanced truncation method from [Enns, 1984] proposed in [Wang et al., 1999] and introduced in Section 5.1.1.

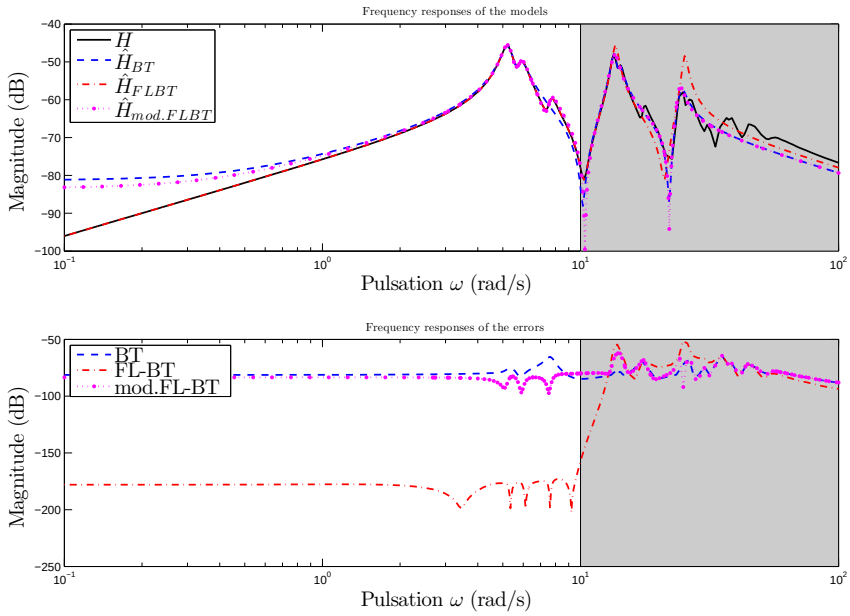


Figure 5.1: Comparison of the **FL-BT** with the **BT** and the **modified FL-BT** by approximation of the LAH model to and order $r = 10$ over $\Omega = [0, 10]$.

It consists in replacing the last terms involved in the Lyapunov equations (5.8), $W_c(\Omega)$ and $W_o(\Omega)$, by positive semi-definite matrices $\tilde{W}_c(\Omega)$ and $\tilde{W}_o(\Omega)$ which are constructed by taking the absolute values of the eigenvalues of $W_c(\Omega)$ and $W_o(\Omega)$.

With this modification, the reduced-order model is guaranteed to be stable and with similar assumptions as in Section 5.1.1, the asymptotic stability can be proven as well and the existence of a bound on the $\mathcal{H}_\infty$-norm of the approximation error and on the frequency-weighted $\mathcal{H}_\infty$-norm of the error (5.1).

Numerical illustration and comparison of the FL-BT

The **FL-BT** is compared to the **BT** and the **modified FL-BT** on a simple case. The LAH model is reduced to an order $r = 10$ over $\Omega = [0, 10]$ with the three methods. The frequency responses of each models as well as the frequency responses of the errors are plotted in Figure 5.1.

One can observe that in the considered frequency band Ω, the error corresponding to the model obtained with **FL-BT** is clearly inferior to the error obtained with the **BT** or the **modified FL-BT**. This is confirmed by $\mathcal{H}_{2,\Omega}$ norm of the errors, indeed,

$$\begin{aligned} \frac{\|H - \hat{H}_{BT}\|_{\mathcal{H}_{2,\Omega}}}{\|H\|_{\mathcal{H}_{2,\Omega}}} &= 10.40\% \\ \frac{\|H - \hat{H}_{mod.FLBT}\|_{\mathcal{H}_{2,\Omega}}}{\|H\|_{\mathcal{H}_{2,\Omega}}} &= 4.10\% \\ \frac{\|H - \hat{H}_{FLBT}\|_{\mathcal{H}_{2,\Omega}}}{\|H\|_{\mathcal{H}_{2,\Omega}}} &= 1.15 \times 10^{-4}\%. \end{aligned}$$

On the frequency responses, the efficiency of the **modified FL-BT** with respect to the **BT** is not obvious but it appears more clearly through the $\mathcal{H}_{2,\Omega}$-norms of the errors. Even if it does not

reach the performance of the **FL-BT**, it is still more efficient that the **BT** when the $\mathcal{H}_{2,\Omega}$-norm is considered.

5.2.2 Gramian-based $\mathcal{H}_{2,\Omega}$ optimal model approximation

The approach presented here has been developed in [Petersson, 2013; Petersson and Löfberg, 2014] in parallel of this thesis. It is an optimisation method aimed at finding a reduced-order model which minimises the $\mathcal{H}_{2,\Omega}$-norm of the approximation error where the latter is expressed with the gramian-based formulation of the norm (see Section 2.2.2). To be more specific, let us consider the n-th order large-scale model $\mathbf{H} := (A, B, C, D)$, the objective of the method is to find a r-th order ($r \ll n$) model $\hat{\mathbf{H}} := (\hat{A}, \hat{B}, \hat{C}, \hat{D})$ which minimises the $\mathcal{H}_{2,\Omega}$-norm of the approximation error between $\mathbf{H}$ and $\hat{\mathbf{H}}$, *i.e.*

$$\hat{H} = \arg\min_{G} \|E\|^2_{\mathcal{H}_{2,\Omega}} = \arg\min_{G} \|H - G\|^2_{\mathcal{H}_{2,\Omega}}, \tag{5.9}$$

where $\Omega = [0, \omega]$. Note that due to Property 1 from Section 2.2.2, more complex frequency intervals can readily be considered.

For simplicity, let us assume here that both models $\mathbf{H}$ and $\hat{\mathbf{H}}$ are strictly proper, *i.e.* $D = \hat{D} = 0$, the realisation of the error model $\mathbf{E} = \mathbf{H} - \hat{\mathbf{H}}$ is then given by

$$\mathbf{E} := \left(\begin{array}{c|c} A_E & B_E \\ \hline C_E & 0 \end{array}\right) = \left(\begin{array}{cc|c} A & 0 & B \\ 0 & \hat{A} & \hat{B} \\ \hline C & -\hat{C} & 0 \end{array}\right) \tag{5.10}$$

The frequency-limited $\mathcal{H}_{2,\Omega}$-norm of the approximation error can be expressed through the gramian-based formulation of the norm (see Section 2.2.2) as

$$\|E\|^2_{\mathcal{H}_{2,\Omega}} = \mathbf{tr}\left(C_E \mathcal{P}_{E,\omega} C_E^T\right) = \mathbf{tr}\left(B_E^T \mathcal{Q}_{E,\omega} B_E\right), \tag{5.11}$$

where $\mathcal{P}_{E,\omega}$ and $\mathcal{Q}_{E,\omega}$ are the frequency-limited gramians associated with $\mathbf{E}$ considered over $\Omega = [0, \omega]$. They are solutions of the following Lyapunov equations,

$$\begin{array}{rcl} A_E \mathcal{P}_{E,\omega} + \mathcal{P}_{E,\omega} A_E^T + W_c(\omega) & = & 0 \\ A_E^T \mathcal{Q}_{E,\omega} + \mathcal{Q}_{E,\omega} A_E + W_o(\omega) & = & 0, \end{array} \tag{5.12}$$

where

$$\begin{array}{rcl} W_c(\omega) & = & S_E(\omega) B_E B_E^T + B_E B_E^T S_E(\omega)^H, \\ W_o(\omega) & = & S_E(\omega)^H C_E^T C_E + C_E^T C_E S_E(\omega), \end{array}$$

with

$$S_E(\omega) = \frac{j}{2\pi} \mathbf{logm}\left(\left(A_E + j\omega I_{n+r}\right)\left(A_E - j\omega I_{n+r}\right)^{-1}\right). \tag{5.13}$$

For asymptotically stable models, (5.13) can be simplified as,

$$S_E(\omega) = \mathbf{Re}\left(\frac{j}{\pi} \mathbf{logm}\left(-A_E - j\omega I_{n+r}\right)\right).$$

Note that due to the structure of the error model (5.10),

$$S_{E,\omega} = \left[\begin{array}{cc} S(\omega) & 0 \\ 0 & \hat{S}(\omega) \end{array}\right].$$

where $S(\omega) = \mathbf{Re}(\frac{j}{\pi}\mathbf{logm}(-A - j\omega I_n))$ and $\hat{S}(\omega) = \mathbf{Re}(\frac{j}{\pi}\mathbf{logm}(-\hat{A} - j\omega I_r))$. Similarly, by partitioning the gramians as

$$\mathcal{P}_{E,\omega} = \left[\begin{array}{cc} \mathcal{P}_\omega & \mathcal{P}_{12,\omega} \\ \mathcal{P}_{12,\omega}^T & \hat{\mathcal{P}}_\omega \end{array}\right] \quad \text{and} \quad \mathcal{Q}_{E,\omega} = \left[\begin{array}{cc} \mathcal{Q}_\omega & \mathcal{Q}_{12,\omega} \\ \mathcal{Q}_{12,\omega}^T & \hat{\mathcal{Q}}_\omega \end{array}\right],$$

the Lyapunov equations (5.12) can be decomposed as

$$\begin{aligned} A\mathcal{P}_\omega + \mathcal{P}_\omega A^T + S(\omega)BB^T + BB^T S(\omega)^H &= 0, \\ A\mathcal{P}_{12,\omega} + \mathcal{P}_{12,\omega}\hat{A}^T + S(\omega)B\hat{B}^T + B\hat{B}^T\hat{S}(\omega)^H &= 0, \\ \hat{A}\hat{\mathcal{P}}_\omega + \hat{\mathcal{P}}_\omega\hat{A}^T + \hat{S}(\omega)\hat{B}\hat{B}^T + \hat{B}\hat{B}^T\hat{S}(\omega)^H &= 0, \end{aligned}$$

and

$$\begin{aligned} A^T\mathcal{Q}_\omega + \mathcal{Q}_\omega A + S(\omega)^H C^T C + C^T C S(\omega) &= 0, \\ A^T\mathcal{Q}_{12,\omega} + \mathcal{Q}_{12,\omega}\hat{A} - \hat{S}(\omega)^H C^T\hat{C} - C^T\hat{C}S(\omega) &= 0, \\ \hat{A}^T\hat{\mathcal{Q}}_\omega + \hat{\mathcal{Q}}_\omega\hat{A} + \hat{S}(\omega)^H\hat{C}^T\hat{C} + \hat{C}^T\hat{C}\hat{S}(\omega) &= 0. \end{aligned}$$

Finally, the approximation error (5.11) can be re-written as

$$\begin{aligned} \|E\|^2_{\mathcal{H}_{2,\Omega}} &= \mathbf{tr}\left(B^T\mathcal{Q}_\omega B + 2B^T\mathcal{Q}_{12,\omega}\hat{B} + \hat{B}^T\hat{\mathcal{Q}}_\omega\hat{B}\right) \\ &= \mathbf{tr}\left(C\mathcal{P}_\omega C^T - 2C\mathcal{P}_{12,\omega}\hat{C}^T + \hat{C}\hat{\mathcal{P}}_\omega\hat{C}^T\right). \end{aligned} \tag{5.14}$$

Based on (5.14), the $\mathcal{H}_{2,\Omega}$ approximation error can be differentiated with respect to the parameters of the reduced-order model $\hat{A}$, $\hat{B}$ and $\hat{C}$ which yields

$$\begin{aligned} \frac{\partial\|E\|^2_{\mathcal{H}_{2,\Omega}}}{\partial\hat{A}} &= 2\left(\mathcal{Q}^T_{12,\omega}\mathcal{P}_{12,\omega} + \hat{\mathcal{Q}}_\omega\hat{\mathcal{P}}_\omega\right) - 2W \\ \frac{\partial\|E\|^2_{\mathcal{H}_{2,\Omega}}}{\partial\hat{B}} &= 2\left(\hat{\mathcal{Q}}_\omega\hat{B} + \mathcal{Q}_{12,\omega}B\right) \\ \frac{\partial\|E\|^2_{\mathcal{H}_{2,\Omega}}}{\partial\hat{C}} &= 2\left(\hat{C}\hat{\mathcal{P}}_\omega - C\mathcal{P}_{12,\omega}\right) \end{aligned} \tag{5.15}$$

where

$$\begin{aligned} W &= \mathbf{Re}\left(\frac{j}{\pi}L\left(-\hat{A} - j\omega I_r, V\right)\right)^T, \\ V &= \hat{C}^T\hat{C}\hat{\mathcal{P}}_\omega - \hat{C}^T C\mathcal{P}_{12,\omega}, \end{aligned}$$

$L(.,.)$ being the Fréchet derivative of the matrix logarithm [Higham, 2008, chap. 11]. The first-order optimality conditions obtained from (5.15) are the generalisation to the $\mathcal{H}_{2,\Omega}$ case of the first-order optimality conditions derived in [Wilson, 1974] for the optimal $\mathcal{H}_2$ model approximation problem.

The approximation error (5.14) and its gradient (5.15) can then be used in an optimisation algorithm to solve the optimal $\mathcal{H}_{2,\Omega}$ approximation problem (5.9). The advantage of working with the gramian formulation of the approximation error in comparison to the optimal $\mathcal{H}_2$ approximation method presented in Section 4.2.3 or the method developed in this thesis in Chapter 9, lies in the fact that here, no limitation comes from the order of the poles. In particular, $\mathbf{H}$ and $\hat{\mathbf{H}}$ can have poles of high order, it does not impact the approach.

On the computation of the matrix logarithm and its Fréchet derivative

One of the most efficient way to compute the matrix logarithms involved in $S(\omega)$ and in $\hat{S}(\omega)$ is to use the *inverse squaring and scaling method* which basic idea relies on the identity,

$$\mathbf{logm}(M) = 2^k\mathbf{logm}(M^{1/2^k}), \tag{5.16}$$

where M is a matrix with no eigenvalues on $\mathbb{R}_-$. Since $M^{1/2^k} \to I$ as $k \to \infty$, a Padé approximation can be used to evaluate the right-hand side of (5.16) to a sufficient degree of accuracy for some k sufficiently large. See [Higham, 2008, chap. 11] for further information and a description of a stable algorithm to evaluate the logarithm of a matrix.

The computation of the Fréchet derivative $L(M, X)$ of the matrix logarithm is done through a similar idea (see [Al-Mohy et al., 2013] for an in-depth description of the algorithm). Indeed, similarly to identity (5.16),

$$L(M, X) = 2^k L(M^{1/2^k}, X_k),$$

where $X_0 = X$ and the X_i $(i = 1, \ldots, k)$ are solutions of the Sylvester equations,

$$M^{1/2^i} X_i + X_i M^{1/2^i} = X_{i-1} \quad i = 1, \ldots, k.$$

Similarly to the logarithm, its Fréchet derivative $L(M, X)$ is then approximated by the Padé approximation of $2^k L(M^{1/2^k}, X_k)$.

Conclusion

In this chapter, some methods for the approximation of a large-scale dynamical model over a bounded frequency interval have been briefly presented. These methods have been classified here into two groups, the methods that rely on the use of frequency-filters called frequency-weighted model approximation methods and the ones that do not referred as frequency-limited model approximation methods.

The most popular frequency-weighted model approximation method is the **FW-BT** which extends the standard **BT** by defining frequency-weighted gramians. Depending on the chosen approach, the **FW-BT** can guarantee the asymptotic stability of the reduced-order model and gives an upper bound on the weighted $\mathcal{H}_\infty$ approximation error. The use of filters has also been considered to extend the **IRKA** in order to address the optimal frequency-weighted $\mathcal{H}_2$ model approximation problem. However, the resulting algorithm does not necessarily leads to an optimal reduced-order model due to some error appearing during the interpolation. These frequency-weighted model approximation methods are efficient but the design of the filter add a tuning parameter which can be avoided by using frequency-limited model approximation methods.

The first one of these methods is the **FL-BT** which extends the **BT** by using frequency-limited gramians to balance the large-scale model. This method is equivalent to one of the **FW-BT** considered with perfect filters and is generally extremely efficient in practice. The other approach is a non-linear optimisation algorithm aimed at solving the optimal $\mathcal{H}_{2,\Omega}$ approximation problem based on the gamian formulation of the norm. This method is of particular interest and a similar approach has been developed in this thesis based on the poles-residues of the $\mathcal{H}_{2,\Omega}$-norm.

This chapter concludes the part dedicated to the state of the art. The next part gathers the contributions of this thesis concerning the frequency-limited approximation of linear dynamical models.

Part III

Frequency-limited approximation of linear dynamical models

Chapter 6

Development of a first approach for frequency-limited model approximation

In this chapter, a first approach to address the problem of $\mathcal{H}_{2,\Omega}$ model approximation is presented. It is an empirical method based on the *Iterative SVD-Rational Krylov Algorithm* (**ISRKA**, see Section 4.2.2) and inspired by the *Frequency-Limited Balanced Truncation* (**FL-BT**, see Section 5.2.1).

In Section 6.1, the principle of the method is described and the algorithm presented. Then in Section 6.2, several numerical experiments are made in order to illustrate the behaviour of the algorithm and its performances.

This method has been presented and successfully applied on an industrial aircraft model in [Vuillemin et al., 2013a,b].

Contents

6.1 Modification of ISRKA 75
6.1.1 Presentation of the method 75
6.1.2 Properties 77
6.1.3 Numerical improvement of the method 78
6.2 Numerical illustrations 79
6.2.1 Comparison with the standard version of **ISRKA** 79
6.2.2 Impact of restart and error watching 83
6.2.3 Comparison with other methods 86

6.1 Modification of ISRKA

In Section 6.1.1, the principle of the method as well as the corresponding algorithm are presented. In Section 6.1.2, the properties of the method are discussed and in Section 6.1.3, some numerical improvements are proposed to enhance the method.

6.1.1 Presentation of the method

This method is based on the sub-optimal $\mathcal{H}_2$ model approximation method **ISRKA** (see Section 4.2.2) and has been inspired by the use of frequency-limited gramians in the **BT** which has led to the **FL-BT** (see Section 5.2.1). In particular, it consists in using a frequency-limited gramian instead of an infinite one in order to build one of the projector in **ISRKA**. The resulting algorithm is called *Frequency-Limited Iterative SVD Tangential Algorithm* (**FL-ISTIA**) and is presented in Algorithm 6. The algorithm has been integrated in the MORE Toolbox [Poussot-Vassal and Vuillemin, 2012] and an example of its use can be found in Appendix C.3.

Initial projectors are built from step 1 to step 3. In particular, V is built as a basis of a Krylov subspace and W is built thanks to the frequency-limited observability gramian $\mathcal{Q}_\Omega$ so that $W^TV = I_r$. Then, while the algorithm has not converged, a reduced-order model is

constructed by projection of the initial large-scale model in step 5. The eigenvalues and right eigenvectors of this reduced-order model are computed at step 6 and are used in steps 7 and 8 in order to build new interpolation points and tangential directions. New projectors are then created in steps 9 and 10 which will be used in the next iteration to build the reduced-order model (step 5). Once the algorithm has converged, the final reduced-order model is obtained by projection of the initial large-scale model (step 12).

Algorithm 6 Frequency-Limited Iterative SVD-Tangential Interpolation Algorithm (**FL-ISTIA**)

Require: $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n_u}$, $C \in \mathbb{R}^{n_y\times n}$, a frequency interval Ω, r initial interpolation points $\sigma_i^{(0)} \in \mathbb{C}$, r initial tangential directions $\hat{\mathbf{b}}_\mathbf{i}^{(0)} \in \mathbb{C}^{1\times n_u}$

1: Construct,
$$V = \left[(\sigma_1^{(0)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{1}^T, \ldots, (\sigma_r^{(0)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{r}^T\right]$$
2: Compute the frequency-limited observability gramian $\mathcal{Q}_\Omega$
3: Compute $W = \mathcal{Q}_\Omega V (V^T \mathcal{Q}_\Omega V)^{-1}$
4: **while** $\max_{k=1,\ldots,r} |\sigma_k^{(i)} - \sigma_k^{(i-1)}| > \varepsilon$ **do**
5: $\quad i \leftarrow i+1$, $\hat{A} = W^T A V$, $\hat{B} = W^T B$
6: $\quad$ Compute $\hat{A}X = \mathbf{diag}(\lambda(\hat{A}))X$
7: $\quad$ Compute $\left[\hat{\mathbf{b}}_\mathbf{1}^T, \ldots, \hat{\mathbf{b}}_\mathbf{r}^T\right]^T = X^{-1}\hat{B}$
8: $\quad$ Set $\sigma_k^{(i)} = -\lambda_k(\hat{A})$, $k = 1, \ldots, r$.
9: $\quad$ Construct,
$$V = \left[(\sigma_1^{(i)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{1}^T, \ldots, (\sigma_r^{(i)} I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{r}^T\right]$$
10: $\quad$ Compute $W = \mathcal{Q}_\Omega V (V^T \mathcal{Q}_\Omega V)^{-1}$
11: **end while**
12: Construct the reduced-order model by projection $\hat{\mathbf{H}} = (W^T A V, W^T B, C V)$

With reference to this algorithm, the following remarks can be addressed :

- The frequency-limited controllability gramian $\mathcal{P}_\Omega$ can be indifferently used instead of the observability one in the algorithm. In that case, the left projector W is built as a basis of a Krylov subspace, *i.e.*
$$W = \left[(\sigma_1^{(i)} I_n - A^T)^{-1} C^T \hat{\mathbf{c}}_\mathbf{1}^T, \ldots, (\sigma_r^{(i)} I_n - A^T)^{-1} C^T \hat{\mathbf{c}}_\mathbf{r}^T\right],$$
where $\left[\hat{\mathbf{c}}_\mathbf{1}^T, \ldots, \hat{\mathbf{c}}_\mathbf{r}^T\right] = \hat{C}X$. The controllability gramian is then involved in the construction of the right projector $V = \mathcal{P}_\Omega W \left(W^T \mathcal{P}_\Omega W\right)^{-1}$.
- The two main numerical costs of this algorithm come from the computation of the frequency-limited gramian at step 2 and the resolution of r linear systems (step 9) at each iterations. Since computing a frequency-limited gramian requires, in addition of solving a Lyapunov equation, to evaluate the logarithm of a matrix, the **FL-ISTIA** is more complex than the **ISRKA** and for the same reasons as those presented in [Gugercin, 2007], it is cheaper than the **FL-BT** depending on the number of iterations.
- The algorithms stops when the interpolation points do not evolve anymore (step 4). Other stopping criteria could be considered. One could for instance stops when the $\mathcal{H}_{2,\Omega}$-norm of the error does not evolve anymore (see Section 6.1.3 for some remarks about the computation of the approximation error).

- To alleviate the computational complexity of the algorithm in very large-scale settings, a low-rank approximation of the gramian could be used as it has been done in [Gugercin, 2007]. However, this has not been considered here.
- Note that as in other methods rooted on Krylov subspaces, to obtain real valued projectors V and W, the interpolation points must either be real or closed under complex conjugation. Indeed, in the latter case, if two vectors are complex conjugate $\mathbf{v_2} = \mathbf{v_1}^*$, then they span the same subspace as their real and imaginary part, *i.e.* $\mathbf{span}\{\mathbf{v_1}, \mathbf{v_2}\} = \mathbf{span}\{\mathbf{Re}(\mathbf{v_1}), \mathbf{Im}(\mathbf{v_1})\}$. This also implies that only one linear system has to be solved (step 1 and 9) for each pair of complex conjugate interpolation points, which reduces the computational complexity.
- In general, the initial points are linearly or logarithmically spaced points in the interval of interest. They can also be chosen as the opposite of the poles of the large-scale model which are associated with the r largest residues in the considered frequency interval.
- Direct feedthrough are not taken into account in this algorithm because it is based on projection and direct feedthrough are not modified by projection. A simple way to handle non strictly proper models is to set $\hat{D} = D$. However, this is not the optimal choice considering the gradient of the approximation error with respect to $\hat{D}$ (see Chapter 8).

6.1.2 Properties

The properties of **FL-ISTIA** concerning stability preservation and interpolation of the large-scale model are discussed below.

Stability

To prove the stability of the reduced-order models built by **ISRKA** [Gugercin, 2007], the Lyapunov equation of the initial large-scale model is projected thus leading to the reduced-order model Lyapunov equation. And, due to inertia results, the reduced-order model is proved to be stable (see Section 4.2.2).

Here however, the same reasoning cannot be applied. Indeed, let assume that the basis is such that $\mathcal{Q}_\Omega = I_n$, then, the large-scale Lyapunov equation is

$$A^T + A + W_o(\Omega) = 0.$$

By applying the projectors $W^T = V^T$ and V to the equation, one obtains

$$\hat{A}^T + \hat{A} + V^T W_o(\Omega) V = 0.$$

The last term of the reduced-order Lyapunov equation is not necessarily positive semi-definite, hence stability cannot be proved. In order to do so, the same modification that has been developed for ensuring the stability with the **FL-BT** [Gugercin and Antoulas, 2004] could be applied (see Section 5.2.1). However, in the case of **FL-BT**, we have observed that this modification also leads to a loss of performances in terms of approximation error in the considered frequency interval. Since instability has rarely been observed in the reduced-order model built by **FL-ISTIA**, this modification is not applied to the algorithm in the sequel.

Interpolation

The Krylov subspace used in **FL-ISTIA** (steps 1 and 9) is the same as the one used in **IRKA** or **ISRKA**. It implies that at convergence, the reduced-order model interpolates tangentially the initial large-scale model **H** from the left or the right (depending on the Krylov subspace

that has been chosen) at $-\hat{\lambda}_i$ for all $i = 1, \ldots, n$. For instance, if at convergence, $V = \left[(\hat{\lambda}_1 I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{1}^\mathbf{T}, \ldots, (\hat{\lambda}_r I_n - A)^{-1} B \hat{\mathbf{b}}_\mathbf{r}^\mathbf{T}\right]$, then for all $i = 1, \ldots, n$,

$$H(-\hat{\lambda}_i)\hat{\mathbf{b}}_\mathbf{i}^T = \hat{H}(-\hat{\lambda}_i)\hat{\mathbf{b}}_\mathbf{i}^T. \tag{6.1}$$

Note that equation (6.1) is one of the first-order optimality conditions for the optimal $\mathcal{H}_2$ approximation problem (see Theorem 9 in Section 4.1.2). For SISO models, just like **ISRKA**, this implies that the reduced-order model built by **FL-ISTIA** is the best (in the $\mathcal{H}_2$-sense) among all the models which share the same eigenvalues. The same property holds in SIMO and MISO cases as well (depending on which side the Krylov subspace is built). In MIMO cases however, the optimality property is weaker since only one part of the first-order optimality conditions is fulfilled. In all cases, fulfilling this optimality condition implies that at convergence, the $\mathcal{H}_2$-norm of the approximation error between the large-scale and reduced-order models is equal to the difference of their $\mathcal{H}_2$-norm (see Section 4.1), *i.e.*

$$\|H - \hat{H}\|_{\mathcal{H}_2}^2 = \|H\|_{\mathcal{H}_2}^2 - \|\hat{H}\|_{\mathcal{H}_2}^2. \tag{6.2}$$

Fulfilling $\mathcal{H}_2$ optimality conditions is not really relevant in the case of model approximation over a bounded frequency range and the choice of the Krylov subspace could be improved. This stands also for the shifts point selection strategy at step 8 which is a relaxation of a Newton's scheme based on the optimality conditions of the $\mathcal{H}_2$ problem. However, as it will appear later in Chapter 8, the optimality conditions for the $\mathcal{H}_{2,\Omega}$ approximation problem cannot be expressed as convenient interpolation conditions and thus choosing a relevant Krylov subspace is not trivial.

6.1.3 Numerical improvement of the method

Below, two modifications are proposed for improving the behaviour of **FL-ISTIA**.

Error watch

One of the main drawback of **FL-ISTIA** lies in the fact that it does not give any information about the approximation error between the large-scale model and the reduced-order one. This issue can be alleviated by computing the approximation error at each iteration and keeping the model which has yielded the smallest approximation error.

In order to keep the computation time to an acceptable level, the approximation error must not be computed directly, instead its decomposition,

$$\|H - \hat{H}\|_{\mathcal{H}_{2,\Omega}}^2 = \|H\|_{\mathcal{H}_{2,\Omega}}^2 + \|\hat{H}\|_{\mathcal{H}_{2,\Omega}}^2 - \frac{1}{\pi}\int_{-\omega}^{\omega} \mathbf{tr}\left(H(j\nu)\hat{H}(-j\nu)^T\right) d\nu,$$

can be exploited. Indeed, since $\|H\|_{\mathcal{H}_{2,\Omega}}^2$ does not depend of the reduced-order model and consequently, it does not need to be computed. The $\mathcal{H}_{2,\Omega}$-norm of the reduced-order model is easily computed since the order of the model is low. The cross term has to be computed at each iteration and involves the large-scale model. Note that the $\mathcal{H}_{2,\Omega}$-norm of the approximation error cannot in general be simplified as in equation (6.2) since the optimality conditions are not the same as in the $\mathcal{H}_2$ case (see Chapter 8).

Each term of the approximation error can either be computed by solving Lyapunov/Sylvester equations (see for instance [Petersson, 2013]) or through the poles-residues formulation of the approximation error (see Chapter 8). The advantage of the latter formulation lies in the fact that once the eigenvalue decomposition of each model is available, the computation of each term in the error is done through an explicit formulae. Its drawback is that both models must have semi-simple poles only.

Note that both formulations still require at least one initial computationally demanding task, thus this modification can only be applied to models of moderate size.

Restarting procedure

In order to make **FL-ISTIA** less sensitive to the initial shift point selection and to explore more possibilities, a restarting procedure can be used. It consists in launching several times the algorithm with different initial interpolation points and keep the best result. Note that it implies that the approximation error must be computed at least at the end of each launch. Hence this modification also, is dedicated to models of medium size for which this information can be computed.

The new set of initial interpolation points can be chosen in many different ways, but here a simple perturbation of the final interpolation points is considered. This is done by adding some random numbers to the real and complex parts of the points. Note that this perturbation must preserve the fact that the interpolation points come in complex conjugate pairs and that they have a positive real part.

6.2 Numerical illustrations

In this section, several numerical tests are conducted in order to highlight the advantages of the frequency-limited version of **ISRKA**. In particular, in Section 6.2.1, **FL-ISTIA** is compared to **ISRKA** in order to illustrate the benefit provided by the use of a frequency-limited gramian. Then in Section 6.2.2, the improvement brought by the numerical improvement are estimated. In Section 6.2.3, the method is compared to other methods such as the **FL-BT**.

6.2.1 Comparison with the standard version of ISRKA

In order to estimate the gain of performance induced by the use of a frequency-limited gramian, the following experimental procedure is used.

Experimental procedure

A given large-scale model $\mathbf{H}$ is approximated with **ISRKA** and **FL-ISTIA** for reduced orders $r = 2, 4, 6, \ldots, 20$ and for the frequency interval $\Omega = [0, \omega]$ where ω varies linearly from ω_{min} to ω_{max} (40 points) which depends on the considered model[1]. Both methods are used with the same parameters, in particular, the restarting procedure is disabled, the gramian used is the controllability one, the maximum number of iterations is set to 30, the tolerance ϵ is set to 10^{-3}, the error is not checked at each iteration and the initial shift points are selected as real points linearly spaced between $\min_\lambda |\mathbf{Im}(\lambda)|$ and $\max_\lambda |\mathbf{Im}(\lambda)|$ where λ are the poles of $\mathbf{H}$.

For each couple $\{r, \omega\}$, the $\mathcal{H}_{2,\Omega}$-norm of the approximation errors, *i.e.*

$$e_{ISRKA} = \|H - H_{ISRKA}\|_{\mathcal{H}_{2,\Omega}} \text{ and } e_{FLISTIA} = \|H - H_{FLISTIA}\|_{\mathcal{H}_{2,\Omega}},$$

are computed. The improvement brought by **FL-ISTIA** is then measured by the following quantity[2]

$$g = \frac{e_{ISRKA} - e_{FLISTIA}}{e_{ISRKA} + e_{FLISTIA}},$$

which varies between -1 when $e_{ISRKA} = 0$ to 1 when $e_{FLISTIA} = 0$. Note that given $k \in \mathbb{R}$, such that $e_{FLISTIA} = k e_{ISRKA}$ then $k = \frac{1-g}{1+g}$. Hence, the ratio between the two errors can be retrieved from g quite easily.

This procedure is applied on the three test models from $COMPL_eib$ presented in the introduction, *i.e.* the LAH, the CBM and the ISS models where

[1] ω_{min} is chosen so that the first main dynamic of the model belongs to Ω and ω_{max} is chosen so that all the dynamics belong to Ω.

[2] The standard relative error $(e_{ISRKA} - e_{FLISTIA})/e_{ISRKA}$ is not used because it is not symmetric and can lead to misleading statistics.

test model	LAH	CBM	ISS
$g \geq 0$ (**FL-ISTIA** better)	60.75	42	69.5
$g < 0$ (**ISRKA** better)	34.75	50.25	28.75
unstable	0	0.5	0
not converged	4.5	7.25	1.75

Table 6.1: Proportion of cases for which g is positive, strictly negative, for which **FL-ISTIA** has led to an unstable model or has not converged (%).

test model	LAH	CBM	ISS
mean	13.47	2.69	13.62
max	99.99	99.91	95.54
min	-5.06	-32.31	-31.19
std	28.74	20.33	25.81

Table 6.2: Statistical indicators for the gain g (in %) computed among the case where **FL-ISTIA** has converged and led to a stable model.

- for the LAH model, $\omega_{min} = 6$ and $\omega_{max} = 100$,
- for the CBM model, $\omega_{min} = 0.1$ and $\omega_{max} = 45$,
- and for the ISS model, $\omega_{min} = 1$ and $\omega_{max} = 100$.

In Table 6.1, the proportions of couples of approximation parameters $\{r, \omega\}$ for which the gain g is positive, strictly negative, or for which **FL-ISTIA** has not converged or converged towards an unstable model are reported.

In Figure 6.1, the spatial distribution of the sign of g is plotted in the plane formed by the reduction order r and the frequency bound ω. The blue crosses indicate that **FL-ISTIA** has not converged at that point, the black squares indicate that **FL-ISTIA** has led to an unstable model, the green (resp. red) points indicate that g is positive (resp. strictly negative).

In Table 6.2, some statistical indicators about g considered among all the couples of approximation parameters for which **FL-ISTIA** has converged and has led to a stable model are reported.

Finally, Figure 6.2 shows the relative approximation error obtained with both method against ω for the LAH model and for a fixed approximation order $r = 4$ (this case corresponds to the second column on the first plot of Figure 6.1).

Interpretation of the results

By looking a the mean values of the gain g provided in Table 6.2, one can see that on average, the use of a frequency-limited gramian instead of an infinite one enables to improve the performances in terms of $\mathcal{H}_{2,\Omega}$-norm of the approximation error. However, this greatly varies between the models, for instance, the mean of g is lower for the CBM than for the LAH and ISS models. A similar conclusion comes from Table 6.1, indeed, we can see that with the beam model, the reduced-order model is improved by **FL-ISTIA** less often than for the other models. This is not surprising considering the frequency response of the clamped beam model (see Figure 1.6 in the introduction). Indeed, the CBM model is mainly characterised by a poorly damped oscillatory mode at low frequency, the remaining oscillatory modes are almost insignificant. A $\mathcal{H}_2$ model approximation method such as **ISRKA** directly catches this low frequency dynamic thus making frequency-limited method not specifically relevant in that case. In fact it is quite easy to build examples on which the **FL-ISTIA** (or other frequency-limited model approximation

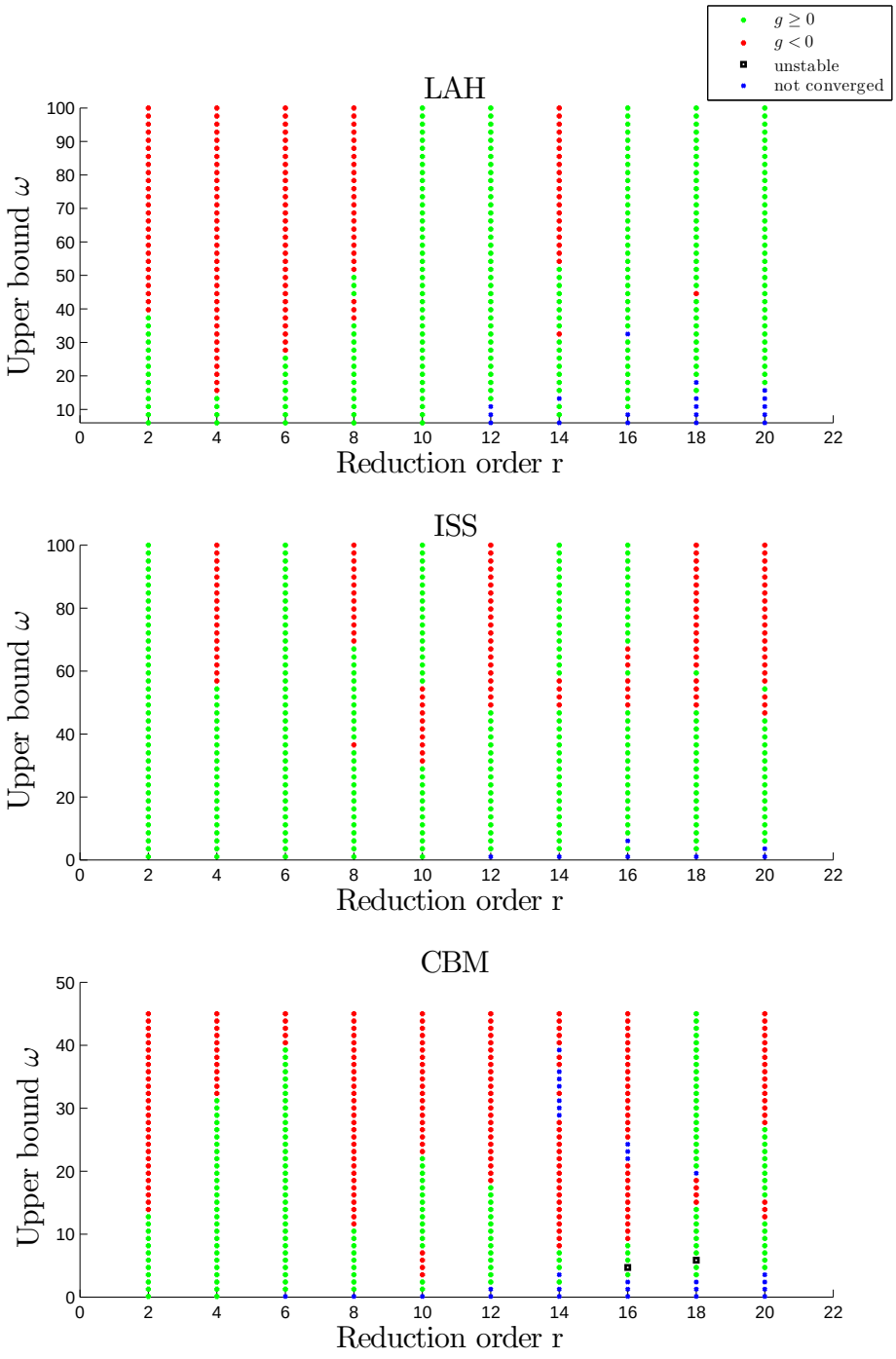


Figure 6.1: Sign of the gain g for each couple of approximation parameters $\{r, \omega\}$ for LAH (top), CBM (center) and ISS (bottom).

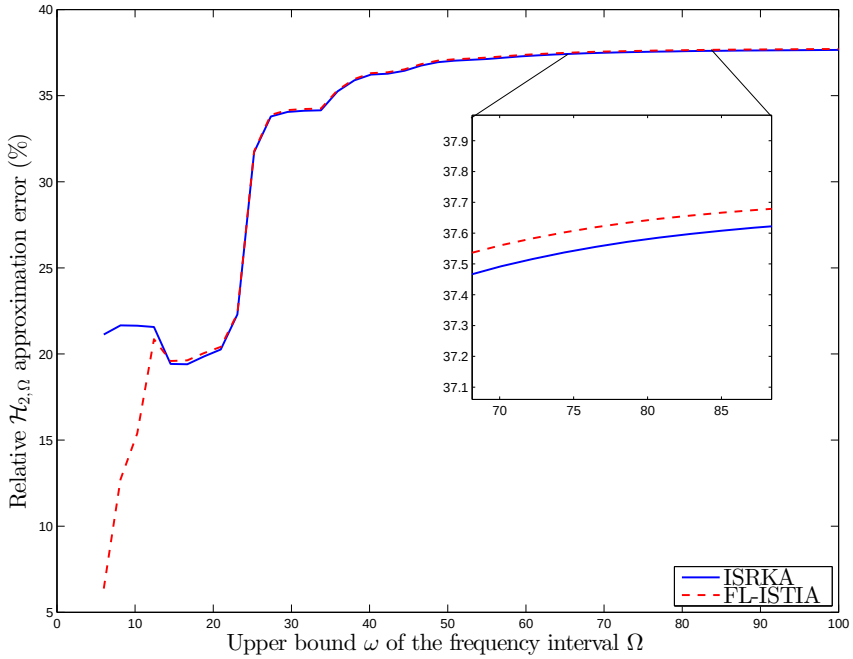


Figure 6.2: Relative $\mathcal{H}_{2,\Omega}$ approximation error with respect to ω for the LAH model and $r = 4$.

methods) is particularly relevant (see Example 13). Here, the ISS model is much more suited for frequency-limited model approximation over $\Omega = [0, \omega]$ than CBM.

When looking at the spatial distribution of the sign of g in Figure 6.1, two remarks can be made. Firstly, the cases for which the algorithm has not converged are mainly located at high reduction orders r and low frequency bounds ω (bottom right of the plots), *i.e.* small frequency interval Ω. This may be explained by the fact that the reduction order is too high compared to the number of meaningful dynamics in that small frequency interval. For instance, for the LAH model with $r = 10$ and $\Omega = [0, 6]$, the relative approximation error between the large-scale and reduced-order models,

$$\frac{\|H - H_{FLISTIA}\|_{\mathcal{H}_{2,\Omega}}}{\|H\|_{\mathcal{H}_{2,\Omega}}},$$

is already equal to $1.75 \times 10^{-4}\%$. Hence, increasing the approximation order is not necessarily relevant. Besides, the two unstable models obtained for the model CBM are also located at low frequency and high approximation orders.

Secondly, on average, the gain brought by **FL-ISTIA** seems to be more dominant for small values of ω as the red stars are mainly located in the top of the plots. This is also illustrated in Figure 6.2 where the difference between the approximation errors is particularly obvious for low values of ω.

Figure 6.2 also enables to mitigate the negative gain obtained by **FL-ISTIA** in some cases. The approximation errors obtained with **FL-ISTIA** are indeed higher than those obtained with **ISRKA** for $\omega > 10$ but the difference is not large.

Example 13 (Settings for which frequency-limited model approximation methods are particularly relevant)**.** *Let us consider a model* **H** *with two poorly damped modes corresponding to the*

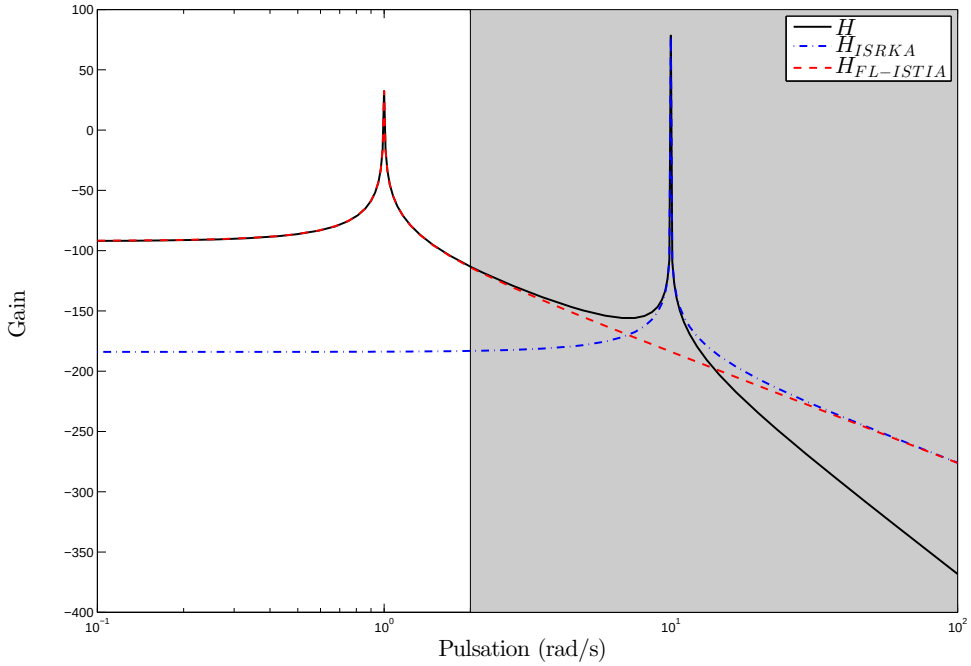


Figure 6.3: Frequency responses of the models H, H_{ISRKA} and $H_{FLISTIA}$.

eigenvalues $\lambda_1 = 10^{-3} + j$ *and* $\lambda_2 = 10^{-5} + 10j$ *and whose transfer function is*

$$H(s) = \frac{1}{(s^2 + 0.002s + 1)\,(s^2 + 2 \times 10^{-5}s + 100)}.$$

This 4-th order model is reduced to a second order model with ***ISRKA*** *and* ***FL-ISTIA*** *over* $\Omega = [0, 2]$. *The frequency responses of the three models are plotted in Figure 6.3.*

Since the high order mode has a higher gain than the low frequency one, ***ISRKA*** *catches it whereas* ***FL-ISTIA*** *catches the low frequency mode which is in the frequency interval* Ω. *The low frequency dynamic is discarded by* ***ISRKA*** *and the high frequency dynamic has no impact in the considered frequency interval, hence the approximation error obtained with* ***ISRKA*** *is extremely high :*

$$\frac{\|H - H_{ISRKA}\|_{\mathcal{H}_{2,\Omega}}}{\|H\|_{\mathcal{H}_{2,\Omega}}} \approx 100\%$$
$$\frac{\|H - H_{FLISTIA}\|_{\mathcal{H}_{2,\Omega}}}{\|H\|_{\mathcal{H}_{2,\Omega}}} = 0.05\%$$

If the considered frequency interval Ω *is chosen to embed the main dynamic, then both methods lead to the same reduced-order model.*

6.2.2 Impact of restart and error watching

In this section, the improvement brought by the restarting procedure and the error watch are estimated through similar tests to what have been used previously in Section 6.2.1.

test model	LAH	CBM	ISS
$g > 0$ (**modified FL-ISTIA** better)	91.75	86.25	67
$g = 0$ (**modified FL-ISTIA** equivalent)	4	8.25	31.5
modified FL-ISTIA unstable	0	0.25	0
modified FL-ISTIA has not converged	4.25	5.25	1.5

Table 6.3: Proportion of cases for which g is strictly positive, null or for which the **modified FL-ISTIA** has not converged or led to an unstable model (%).

test model	LAH	CBM	ISS
mean	4.17	4.01	4.43
max	84.68	82.59	40.85
min	0	0	0
std	10.22	7.56	8.05

Table 6.4: Statistical indicators for the gain g (in %) computed among the case where both instances of **FL-ISTIA** have converged and led to stable models.

Experimental procedure

The same set of models with the same set of approximation parameters than in Section 6.2.1 are used. But here, two instances of **FL-ISTIA** are compared, the first one has no restarting procedure enabled and no error watch while the second one has two restarts and the error is checked at each iteration. The second instance of **FL-ISTIA** may be referred as **modified FL-ISTIA** below. The gain g is here computed between the two instances of **FL-ISTIA**

$$g = \frac{e_{FLISTIA} - e_{mod.FLISTIA}}{e_{FLISTIA} + e_{mod.FLISTIA}},$$

where

$$e_{FLISTIA} = \|H - H_{FLISTIA}\|_{\mathcal{H}_{2,\Omega}} \text{ and } e_{mod.FLISTIA} = \|H - H_{mod.FLISTIA}\|_{\mathcal{H}_{2,\Omega}}.$$

In Table 6.3, the proportions of cases for which the gain g between **FL-ISTIA** and the **modified FL-ISTIA** is strictly positive, null, or for which the **modified FL-ISTIA** has not converged or has led to an unstable model are reported.

In Table 6.4, statistical indicators of g considered among all the cases for which both instances of **FL-ISTIA** have converged and have led to stable models are reported.

In Figure 6.4, the spatial distribution of the sign of the gain between the **ISRKA** and the **modified FL-ISTIA** is plotted.

Finally, in Table 6.5, the mean relative increase in computation times between the two instances of **FL-ISTIA**, *i.e.*

$$g_t = \frac{t_{modFLISTIA} - t_{FLISTIA}}{t_{FLISTIA}},$$

is reported.

test model	LAH	CBM	ISS
g_t	75	39	42

Table 6.5: Mean relative increase in computation times between **FL-ISTIA** and **modified FL-ISTIA** (%).

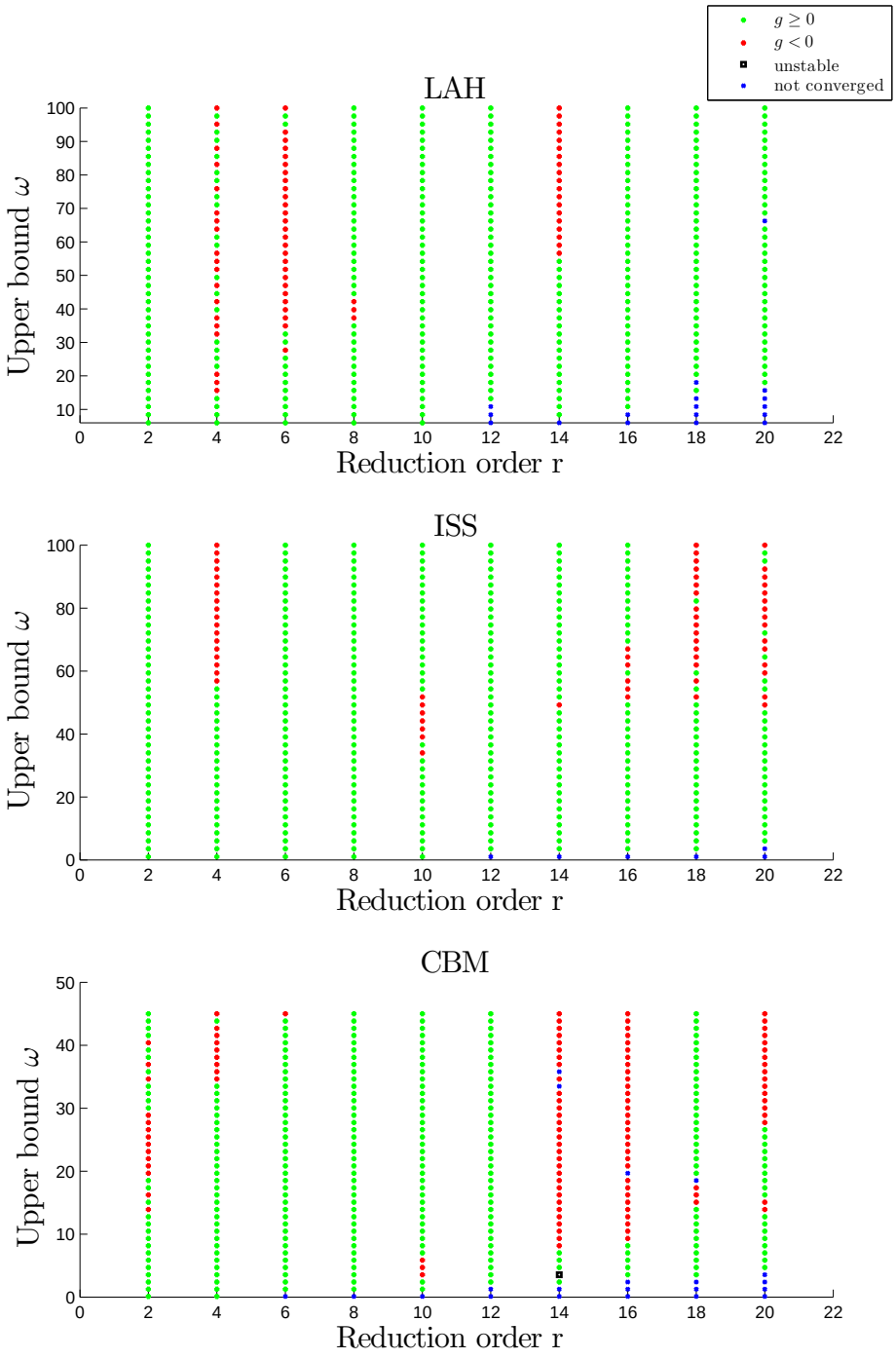


Figure 6.4: Sign of the gain g between **ISRKA** and the **modified FL-ISTIA** for each couple of approximation parameters $\{r, \omega\}$ for LAH (top), CBM (center) and ISS (bottom).

	LAH	CBM	ISS
mean	-2.60	13.50	2.29
max	39.62	92.78	57.73
min	-99.99	-98.12	-89.31
std	25.97	23.02	11.95

Table 6.6: Statistical indicators relative to the gain g_{BT} (in %) between the **FL-BT** and the **FL-ISTIA** considered in all the cases where both methods have converge towards a stable model.

Interpretation of the results

From Table 6.3, one can see that the modifications of **FL-ISTIA** improve the method in the majority of cases. In comparison to Table 6.1, it also shows that these modifications enable to slightly decreases the number of cases for which the algorithm does not converge or leads to a unstable model.

From Table 6.4, one can see that the gain brought by the modifications is on average of 4%, which means that on average $e_{mod.FLISTIA} = 0.92\, e_{FLISTIA}$. Obviously, it cannot be negative since in the worst case, the modifications do not impact the behaviour of the algorithm. Note that since the restart procedure involves random numbers, the result given by the modified **FL-ISTIA** might vary for the same inputs. However the variation is generally small, that is why we did not conduct any statistical study on this point.

Figure 6.4 enables to show the improvement brought by the modifications with respect to **ISRKA**. Indeed, in comparison from Figure 6.1, we can see that there are less red points (g negative).

This performance improvement comes at the cost of longer computation times as showed in Table 6.5. Yet, from an user's point of view, the difference is barely noticeable for models of size lower than 500 states.

6.2.3 Comparison with other methods

In this Section, the **FL-ISTIA** is compared to the **FL-BT** and a corrective example of a benchmark published in [Petersson, 2013] is also presented.

Comparison with the FL-BT

The algorithm **FL-BT** is compared with the **modified FL-ISTIA** with the same parameters as in Section 6.2.2 and on the same test models. The difference between the two methods is again measured by

$$g_{BT} = \frac{e_{FLBT} - e_{mod.FLISTIA}}{e_{FLBT} + e_{mod.FLISTIA}}.$$

In Table 6.6, statistical indicators relative to g_{BT} in the cases where both methods have converged and led to a stable model are reported.

From Table 6.6, one can see that with the ISS and CB models, **FL-ISTIA** performs better on average than the **FL-BT**. However, this is not the case on the LAH model. The **FL-BT** is particularly efficient on the LAH model because it manages to reproduce more accurately the differentiator behaviour of the model than **FL-ISTIA**.

Both methods and **ISRKA** are also used to approximate an industrial aircraft model ($n = 289$, $n_y = 4$ and $n_u = 3$) to an order $r = 12$ over the frequency interval $\Omega = [0, \omega]$. The relative $\mathcal{H}_{2,\Omega}$ approximation errors are potted in Figure 6.5. This model is very ill-conditioned and in fact, the **FL-BT** only produces unstable models even if the modification proposed in [Gugercin and Antoulas, 2004] (see Section 5.2.1) is used. Similarly, one can observe that the evolution of

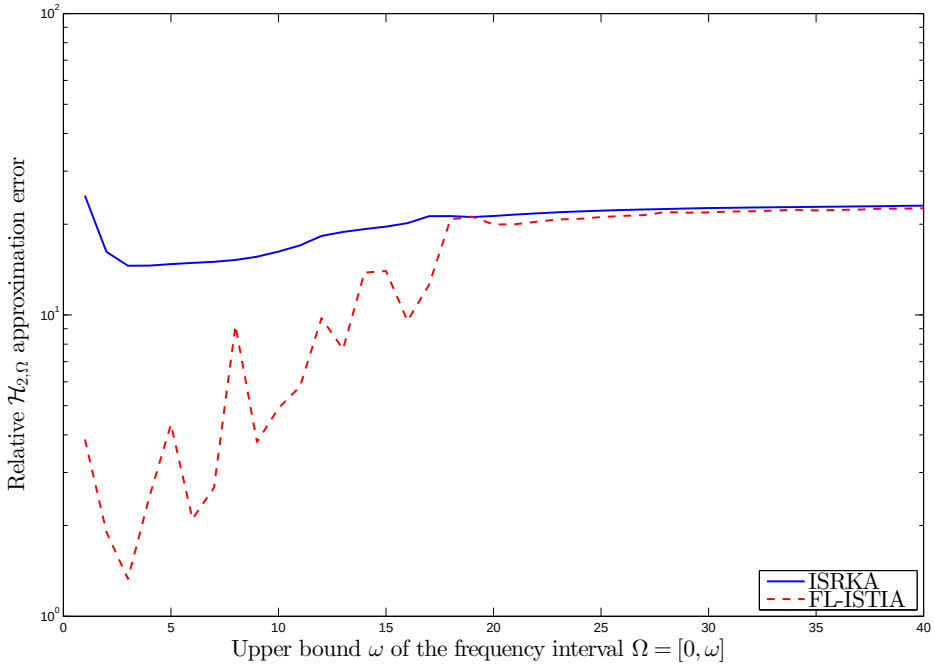


Figure 6.5: Relative $\mathcal{H}_{2,\Omega}$ approximation error obtained by approximation of an industrial aircraft model to an order $r = 12$ with **ISRKA**, **FL-ISTIA** with $r = 12$ and $\Omega = [0, \omega]$ for varying values of ω.

the approximation error corresponding to **FL-ISTIA** is chaotic but the reduced-order model is still stable and of acceptable accuracy.

These results can be explained by the fact that the ill-conditioning of the model prevents from solving accurately the Lyapunov equations which results in inaccurate frequency-limited gramians. Indeed, when the frequency limited gramians are re-injected in their corresponding Lyapunov equations, one can see that $\|A^T\mathcal{Q}_\Omega + \mathcal{Q}_\Omega A + W_o(\Omega)\|_2$ and $\|A\mathcal{P}_\Omega + \mathcal{P}_\Omega A^T + W_c(\Omega)\|_2$ are far from being null with this model. In the case of **FL-ISTIA**, this inaccuracy is compensated by the Krylov part of the method.

Corrective example

In [Petersson, 2013, chap.4], a comparison between several frequency-limited model approximation algorithms, among which **FL-ISTIA**, is presented. The Example 4.7 consists in reducing the CD player model [Leibfritz and Lipinski, 2003] to an order[3] $r = 12$ over $\Omega = [10, 1000]$.

In the presented results, the **FL-ISTIA** performed poorly. This is probably due to some problems present in the previous release of the MORE toolbox. Indeed, the error obtained in the current state of the algorithm differs as shown in Table 6.7 where the relative $\mathcal{H}_{2,\Omega}$ approximation errors between the large-scale model $\mathbf{H}$ and the reduced order model $\hat{\mathbf{H}}$, *i.e.*

$$\frac{\|H - \hat{H}\|_{\mathcal{H}_{2,\Omega}}}{\|H\|_{\mathcal{H}_{2,\Omega}}},$$

are reported for the **FL-BT**, the old version of **FL-ISTIA**, the new one and the methods

[3]The reduced order is not clearly indicated but is likely to be 12 due to the error produced by the **FL-BT**.

FL-BT	1.24e-03
old **FL-ISTIA**	8.23e-02
current **FL-ISTIA**	9.76e-04
FLH2NL	6.95e-04
WH2NL	8.94e-04

Table 6.7: Relative approximation errors obtained by approximation of the CDP model to an order $r = 12$ over $\Omega = [10, 1000]$ with different approximation methods

proposed in [Petersson, 2013]: FLH2NL and WH2NL. Note that the approximation with **FL-ISTIA** is considered without restart, without error watch, with a tolerance of 10^{-3}, with the observability gramian, and with initial shift points linearly spaced between 10 and 1000.

The error obtained with the current version of **FL-ISTIA** is far better than the old one, but still above the errors induced by FLH2NL and WH2NL, in that specific case.

Conclusion

In this Chapter, our first attempt at addressing the problem of model approximation over a bounded frequency range has been presented. The idea was to use frequency-limited gramians in place of infinite gramians in the algorithm **ISRKA**, the resulting algorithm is called **FL-ISTIA**. Numerical experiments have shown that **FL-ISTIA** is indeed more efficient than **ISRKA** when the $\mathcal{H}_{2,\Omega}$-norm of the approximation error is considered. Its performances are also comparable to those of the **FL-BT** and it seems that **FL-ISTIA** is numerically more robust. The **FL-ISTIA** could be improved by using a Krylov subspace to enforce interpolation conditions that are relevant for the $\mathcal{H}_{2,\Omega}$ approximation problem, but determining this Krylov subspace is not obvious and is still under investigation. In the sequel, the problem of approximation over a bounded frequency range is considered directly as an optimisation problem in terms of the $\mathcal{H}_{2,\Omega}$-norm.

Chapter 7

Formulation of the $\mathcal{H}_{2,\Omega}$-norm with the poles and residues of the transfer function

The poles-residues representation of a LTI model is not the most numerically robust representation that exists and it is dedicated (in order to be applicable) to models having semi-simple poles only. But the associated formulation of the $\mathcal{H}_2$-norm has led to convenient first-order optimality conditions in the context of optimal $\mathcal{H}_2$ model approximation (see Chapter 4). That is why in this chapter, this representation is used to derive a poles-residues formulation for the frequency-limited $\mathcal{H}_2$-norm for which only the gramian formulation existed (see Section 2.2).

Note that this poles-residues formulation of the norm is used in Chapters 8 and 9 to develop an algorithm that minimises the $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$-norm of the approximation error between a large-scale model and a reduced-order one. Hence, this Chapter plays a pivotal role in this manuscript.

In a first step in Section 7.1, the case of models with semi-simple poles only is addressed. The formulation is then generalised to models with high order poles in Section 7.2. This poles-residues formulation is finally used in Section 7.3 to derive two bounds on the $\mathcal{H}_\infty$-norm of LTI dynamical models which are of great interest in the context of model approximation (see Section 9.2.3 and Chapter 10).

The poles-residues formulations of the $\mathcal{H}_{2,\Omega}$-norm has been presented in [Vuillemin et al., 2012b] and in [Vuillemin et al., 2014c] for the models with semi-simple and high order poles, respectively, and the bounds on the $\mathcal{H}_\infty$-norm has been presented in [Vuillemin et al., 2014d].

Contents

7.1 Models with semi-simple poles only . . . 89
7.1.1 Preliminary results on complex functions . . . 90
7.1.2 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm . . . 91
7.1.3 Numerical illustration of the formulation . . . 93
7.2 Models with higher order poles . . . 94
7.2.1 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm for models with high order poles . . . 95
7.2.2 Special case 1 : n eigenvalues of multiplicity 1 (semi-simple case) . . . 97
7.2.3 Special case 2 : 1 eigenvalue of multiplicity n . . . 97
7.3 Upper bounds of the $\mathcal{H}_\infty$-norm . . . 98
7.3.1 Formulation of the bounds . . . 98
7.3.2 Computation of the bounds . . . 101
7.3.3 Experimental study of the bounds quality . . . 102
7.3.4 Construction of a frequency template . . . 105

7.1 Models with semi-simple poles only

The poles-residues (or spectral) formulation of the frequency-limited $\mathcal{H}_2$-norm for LTI models with semi-simple poles only involves the principal value of the inverse tangent function which

is introduced in Section 7.1.1. The main Theorem is then stated in Section 7.1.2. Finally, this formulation of the norm is compared in some numerical examples in Section 7.1.3.

7.1.1 Preliminary results on complex functions

The principal value[1] of the inverse tangent function relies on the principal value of the complex logarithm which relies itself on the principal value of the complex argument. Hence, the definitions of both functions are conditioned by the way the complex argument is defined, which is merely a convention, but an important one. In this work, we have chosen a widely used convention (see [Abramowitz and Stegun, 1964, chap. 4] or [Corless et al., 2000]) for the principal value of the complex argument. The argument $\mathbf{arg}(z)$ of a non-zero complex number $z \neq 0$ is chosen to lie in $]-\pi, \pi]$, *i.e.*

$$-\pi < \mathbf{arg}(z) \leq \pi,$$

for $z \neq 0$. This definition induces a branch cut along the negative real axis $(-\infty, 0]$ which is inherited by the functions based on this convention. In particular, the principal value of the complex logarithm, which is presented in Definition 7, has the same branch cut as the complex argument.

Definition 7 (Principal value of the complex logarithm)**.** *The principal value of the logarithm of z, denoted $\mathbf{log}(z)$, is defined for $z \neq 0$ as*

$$\mathbf{log}(z) = \mathbf{ln}\left(|z|\right) + j\mathbf{arg}(z),$$

where $\mathbf{ln}(x)$ is the natural logarithm of $x \in \mathbb{R}_+^*$.

There exist various definitions for the complex inverse tangent function. Here, the definition suggested in [Kahan, 1987] is considered (see Remark 13 for some remarks about alternative formulations). It is recalled in Definition 8. The function $\mathbf{atan}(z)$, $z \neq \pm j$ is single-valued and continuous excepted along two branch cuts $(-j\infty, -j]\cup[j, j\infty)$ (illustrated in Appendix B.1).

Definition 8 (Principal value of the complex inverse tangent)**.** *The principal value of the inverse tangent of z, denoted $\mathbf{atan}(z)$, is defined for $z \neq \pm j$ as*

$$\mathbf{atan}(z) = \frac{1}{2j}\left(\mathbf{log}(1+jz) - \mathbf{log}(1-jz)\right).$$

Remark 13 (About the definition of the principal value of the complex inverse tangent)**.** *The principal value of the complex inverse tangent can also be defined, for $z \neq \pm j$ as*

$$\frac{1}{2j}\mathbf{log}\left(\frac{1+jz}{1-jz}\right),$$

which yields the same branch cuts as Definition 8 but does not coincide everywhere in the complex plane. Indeed as shown in [Haber, 2012] (and illustrated in Appendix B.1),

$$\mathbf{log}\left(\frac{1+jz}{1-jz}\right) = \begin{cases} \pi + \mathbf{log}(1+jz) - \mathbf{log}(1-jz) & \text{if } z \in (-j\infty, -j[\\ \mathbf{log}(1+jz) - \mathbf{log}(1-jz) & \text{otherwise} \end{cases}.$$

In this study, the use of the complex inverse tangent function is merely a notational convenience that enables to compact the expression of the $\mathcal{H}_{2,\Omega}$-norm. The same results could have been obtained by sticking with the complex logarithm. Moreover the hypothesis considered in Theorem 10 on the value of ω ensures that one cannot be in a case where the two definitions are not equivalent.

[1] In this work, the principal values of complex functions are denoted by lower-case (and bold) names. The notation often varies in the literature, here it is the same as in [Abramowitz and Stegun, 1964, chap. 4] but differs, for instance, from the notation employed in the documentation of Mathematica.

Yet, generally speaking, special care must be taken when using **atan**(z) *in practice because the convention of the function might vary between different software. Indeed, Mathematica documentation explicitly mentions that Definition 8 is used*[2]*, but there is no indication in the documentations of Matlab®, Scilab and Octave.*

7.1.2 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm

Theorem 10 (Spectral expression of the $\mathcal{H}_{2,\Omega}$-norm)**.** *Given a frequency interval $\Omega = [0, \omega]$ and a continuous MIMO LTI dynamical model* $\mathbf{H}$ *of degree n. Suppose that the two following hypothesis are satisfied*

1. *The transfer function associated with* $\mathbf{H}$ *can be written as*

$$H(s) = \sum_{i=1}^{n} \frac{\Phi_i}{s - \lambda_i} + D \in \mathbb{C}^{n_y \times n_u},$$

   where $D \in \mathbb{R}^{n_y \times n_u}$, $\Phi_i \in \mathbb{C}^{n_y \times n_u}$ and $\lambda_i \in \mathbb{C}$ $(i = 1, \ldots, n)$.

2. *If there are indexes $\{j_1, \ldots, j_K\}$ such that $\{\lambda_{j_1}, \ldots, \lambda_{j_K}\}$ are purely imaginary, then the upper bound ω of the interval Ω is smaller than their modulus,* i.e. $\omega < \min\{|\lambda_{j_1}|, \ldots, |\lambda_{j_K}|\}$.

Then, the frequency-limited $\mathcal{H}_2$-norm of $\mathbf{H}$ *can be written as*

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{i,k} + \frac{\omega}{\pi}\mathbf{tr}\left(DD^T\right) - \frac{2}{\pi}\sum_{i=1}^{n}\mathbf{tr}\left(\Phi_i D^T\right)\mathbf{atan}\left(\frac{\omega}{\lambda_i}\right) \tag{7.1}$$

where

$$a_{i,k} = \begin{cases} \dfrac{2}{\pi}\mathbf{tr}\left(\dfrac{\Phi_i\Phi_k^T}{\lambda_i + \lambda_k}\right)\mathbf{atan}\left(\dfrac{\omega}{\lambda_i}\right) & \text{if } \lambda_i + \lambda_k \neq 0 \\ -\dfrac{1}{\pi}\mathbf{tr}\left(\dfrac{\omega\Phi_i\Phi_k^T}{\omega^2 + \lambda_i\lambda_i}\right) & \text{otherwise.} \end{cases}$$

Proof. Since A is diagonalisable, the transfer function $H(s)$ can be decomposed as (see Section 2.1.1)

$$H(s) = \sum_{i=1}^{n} \frac{\Phi_i}{s - \lambda_i} + D,$$

where $\Phi_i \in \mathbb{C}^{n_y \times n_u}$ is the residue associated with the eigenvalue λ_i. Using this decomposition in the definition of the $\mathcal{H}_{2,\Omega}$-norm (Definition 4) enables to divide it into several elements than can be integrated under some additional assumptions in particular cases. Indeed, if some poles are on the imaginary axis, the upper bound ω of the frequency interval Ω must be smaller than the smallest modulus of those imaginary poles. The full proof can be found in Appendix A.1. □

The two expressions for the $a_{i,k}$ ensure that (7.1) is continuous for every value of λ_i and λ_k. The second expression is the limit of the first one as $|\lambda_i + \lambda_k|$ tends towards 0. Note that $\lambda_i = -\lambda_k$ happens if the model has purely imaginary poles and can also happen if it has poles in the right half-plane located symmetrically to some stable poles with respect to the imaginary axis. Both cases do not occur if asimptotically stable models are considered and are therefore discarded in the sequel.

Note that if $\lambda_i + \lambda_k \neq 0 \ \forall\, i,\ k = 1, \ldots, n$, then equation (7.1) can be rewritten as

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \sum_{i=1}^{n} -\frac{2}{\pi}\mathbf{tr}\left(\Phi_i H(-\lambda_i)^T\right)\mathbf{atan}\left(\frac{\omega}{\lambda_i}\right) + \frac{\omega}{\pi}\mathbf{tr}\left(DD^T\right), \tag{7.2}$$

[2] See `http://mathworld.wolfram.com/InverseTangent.html`

which is close to the poles-residues expression of the $\mathcal{H}_2$-norm (see Section 2.1.1)

$$\|H\|_{\mathcal{H}_2}^2 = \sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i H(-\lambda_i)^T\right). \tag{7.3}$$

Apart from the non-proper terms, the difference between equations (7.2) and (7.3) comes from the coefficients $\mathbf{atan}(\frac{\omega}{\lambda_i})$ which acts as weightings on each contribution in the sum.

If the model $\mathbf{H}$ is strictly proper, then the limit of the $\mathcal{H}_{2,\Omega}$-norm of $\mathbf{H}$ as ω tends towards infinity is determined by the limits of $\mathbf{atan}(\frac{\omega}{\lambda_i})$. The latter are presented in Corollary 1 and the former in Theorem 11.

Corollary 1. *Given $\lambda \in \mathbb{C}^*$ and $\omega > 0$, the limits of $\mathbf{atan}\left(\frac{\omega}{\lambda}\right)$, where $\mathbf{atan}(z)$ is given by Definition 8, are*

$$\lim_{\omega\to\infty} \mathbf{atan}\left(\frac{\omega}{\lambda}\right) = \begin{cases} -\frac{\pi}{2} & \text{if } \mathbf{Re}\,(\lambda) < 0 \\ \frac{\pi}{2} & \text{if } \mathbf{Re}\,(\lambda) > 0 \\ \frac{\pi}{2} & \text{if } \mathbf{Re}\,(\lambda) = 0 \text{ and } \mathbf{Im}\,(\lambda) < 0 \\ -\frac{\pi}{2} & \text{if } \mathbf{Re}\,(\lambda) = 0 \text{ and } \mathbf{Im}\,(\lambda) > 0 \end{cases}.$$

Proof. See Appendix B.2. □

Theorem 11 (Limit as ω tends towards infinity)**.** *Let us consider a frequency interval $\Omega = [0, \omega]$ and a strictly proper model $\mathbf{H}$ with n_- poles λ_i^- in the open left half-plane and n_+ poles λ_k^+ in the open right-half plane that satisfies $\lambda_i^- + \lambda_k^+ \neq 0 \ \forall i, k$. The transfer function of $\mathbf{H}$ is written as*

$$H(s) = \sum_{i=1}^{n_-} \frac{\Phi_i^-}{s - \lambda_i^-} + \sum_{k=1}^{n_+} \frac{\Phi_k^+}{s - \lambda_k^+}.$$

Then, the limit of the frequency-limited $\mathcal{H}_2$-norm (see Remark 14 for some clarification about the notation for unstable models) as ω tends towards ∞ is given by

$$\lim_{\omega\to\infty} \|H\|_{\mathcal{H}_{2,\Omega}} = \sum_{i=1}^{n_-} \mathbf{tr}\left(\Phi_i^- H^T(-\lambda_i^-)\right) - \sum_{k=1}^{n_+} \mathbf{tr}\left(\Phi_k^+ H^T(-\lambda_k^+)\right).$$

Proof. Using Corollary 1 in equation (7.2) leads to the result. □

Remark 14 (On the notation $\mathcal{H}_{2,\Omega}$ for unstable models)**.** *Given the definition of the frequency-limited $\mathcal{H}_2$-norm considered in this thesis (see Definition 4), the $\mathcal{H}_{2,\Omega}$-norm of an unstable model is not necessarily infinite unlike the $\mathcal{H}_2$-norm.*

The notation $\mathcal{H}_{2,\Omega}$ is misleading in that case and it would have been more appropriate to talk about frequency-limited $\mathcal{L}_2$ norm, or $\mathcal{L}_{2,\Omega}$-norm, which matches Definition 4 and enables to encompass both the stable and unstable cases.

Yet, since the models considered in this thesis are mainly stable ones, the notation $\mathcal{H}_{2,\Omega}$-norm is used.

In the case of asymptotically stable models, Theorem 11 states that the frequency-limited $\mathcal{H}_2$-norm of $\mathbf{H}$ tends towards its $\mathcal{H}_2$-norm as ω tends towards ∞. For a model $\mathbf{H} = (A, B, C)$ which poles are located in the right half-plane, the $\mathcal{H}_{2,\Omega}$-norm of $\mathbf{H}$ tends towards the $\mathcal{H}_2$-norm of the model which realisation is $(-A, B, C)$. This is not surprising, since the $\mathcal{H}_{2,\Omega}$-norm measures the impact of the transfer function on the imaginary axis, whether the poles are located on the right or on the left of the axis does not modify this impact.

If the model $\mathbf{H}$ is not strictly proper or has poles on the imaginary axis, then the limit is infinite.

On the computation of the $\mathcal{H}_{2,\Omega}$-norm

To efficiently compute the $\mathcal{H}_{2,\Omega}$-norm (7.2), the residues Φ_i should not explicitly be computed. Indeed, by noting $X \in \mathbb{C}^{n\times n}$ the matrix which columns are the right eigenvectors of A, $Y = X^{-1}$, and $\mathbf{e_i}$ the canonical basis column vector, it turns out that

$$\begin{aligned}\mathbf{tr}\left(\Phi_i\Phi_k^T\right) &= \mathbf{tr}\left(CX\mathbf{e_i}\mathbf{e_i}^TYB(CX\mathbf{e_k}\mathbf{e_k}^TYB)^T\right)\\ &= \mathbf{e_k}^T\underbrace{(CX)^TCX}_{M_1}\mathbf{e_i}\mathbf{e_i}^T\underbrace{YB(YB)^T}_{M_2}\mathbf{e_k}.\end{aligned}$$

Hence

$$\mathbf{tr}\left(\Phi_i\Phi_k^T\right) = [M_1]_{k,i}\,[M_2]_{i,k}\,.$$

Since M_1 and M_2 are symmetric,

$$\mathbf{tr}\left(\Phi_i\Phi_k^T\right) = [M_1]_{i,k}\,[M_2]_{i,k}\,.$$

Consequently, by denoting L the matrix defined by

$$[L]_{i,k} = \frac{1}{\lambda_i+\lambda_k},$$

one obtains

$$\sum_{i=1}^{n}\sum_{k=1}^{n}\frac{\mathbf{tr}\left(\Phi_i\Phi_k^T\right)}{\lambda_i+\lambda_k}\mathbf{atan}\left(\frac{\omega}{\lambda_i}\right) = \mathbb{1}^T\left(M_1\odot M_2\odot L\right)\begin{bmatrix}\mathbf{atan}\left(\frac{\omega}{\lambda_1}\right)\\ \vdots\\ \mathbf{atan}\left(\frac{\omega}{\lambda_n}\right)\end{bmatrix}, \tag{7.4}$$

where $\mathbb{1}$ denotes the column vector filled with ones and $\odot$ the Hadamard product. Similarly, by denoting $\mathbf{diag}\,(M)$ the column vector formed with the diagonal of the matrix M,

$$\sum_{i=1}^{n}\mathbf{tr}\left(\Phi_iD^T\right)\mathbf{atan}\left(\frac{\omega}{\lambda_i}\right) = \mathbf{diag}\left(YBD^TCX\right)^T\begin{bmatrix}\mathbf{atan}\left(\frac{\omega}{\lambda_1}\right)\\ \vdots\\ \mathbf{atan}\left(\frac{\omega}{\lambda_n}\right)\end{bmatrix}.$$

Note that the matrices M_1, M_2 and L in equation (7.4) are symmetric and contain elements that are complex conjugate. Hence, by exploiting this structure, it is possible to decrease even further the number of operations required to computed the $\mathcal{H}_{2,\Omega}$-norm with this formulation.

7.1.3 Numerical illustration of the formulation

To compare the numerical efficiency of the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm, two tests are conducted :

- firstly, with $\Omega = [0,\infty]$, the norms of randomly generated models (with $n_y = n_u = 5$) which orders vary from 1 to 200 are computed with *(i)* the build-in routine `norm` of Matlab®, *(ii)* the gramian formulation of the norm[3] as presented in Section 2.1.2 and *(iii)* the poles-residues formulation presented above. For each order n, 50 models are generated and the average CPU times required for the computation of the norms are reported in Figure 7.1.

[3]Note that the eigen-decomposition of the matrix A is not used when solving the Lyapunov equation to determine the gramian because it would remove what makes this formulation interesting (the absence of hypothesis on the poles of the model).

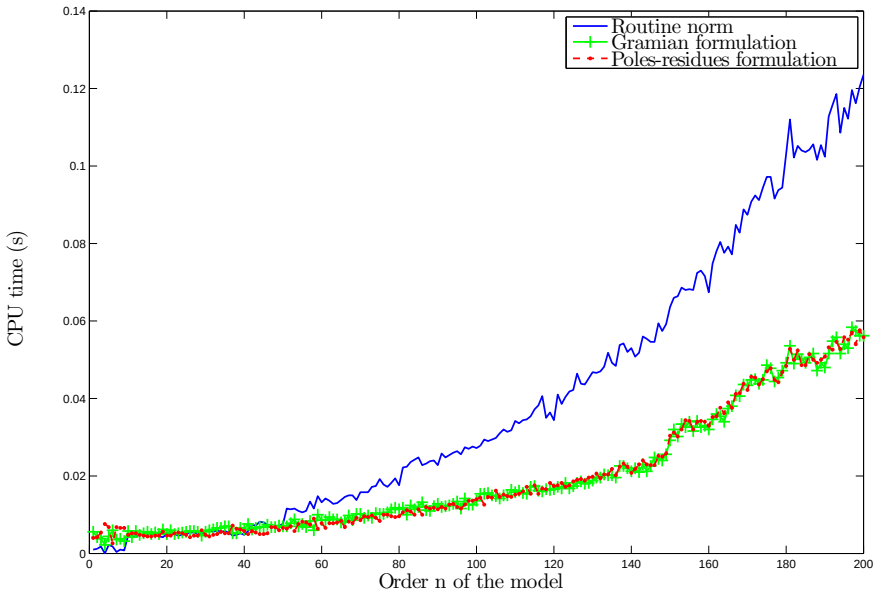


Figure 7.1: Average CPU times required to compute the $\mathcal{H}_2$-norm of 50×200 randomly generated models with the built-in function of Matlab®, the gramian formulation of the norm and the poles-residues formulation

- Secondly, a similar test is achieved with a frequency interval $\Omega = [0, 5]$ (in that case, there is not built-in function and both formulations are computed with the routine developed in the MORE Toolbox, see Appendix C.1). The computation times are presented in Figure 7.2.

In Figure 7.1, one can observe that the routine of Matlab® is seemingly slower than the two other methods but this is probably due to the fact that the built-in method is likely to include way more checks to avoid errors. This first example mainly illustrates the fact that both the poles-residues and gramian formulations are equivalent for $\Omega = [0, \infty]$ in term of computation times.

However, one can see in Figure 7.2 that when a frequency-limited gramian is involved, the poles-residues formulation tends to be faster. This can be explained by the fact that with a bounded frequency interval, a logarithm of matrix must be evaluated to create the last term of the Lyapunov equation (see equation (2.11)) which adds additional numerical cost for the gramian formulation. On the contrary, the complexity of the poles-residues formulation does not significantly vary with the frequency interval since it is explicit in ω and only involves elementary functions.

The poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm presented in Theorem 10 relies on the fact that the model has semi-simple poles only. This assumption can be alleviated and a similar expression can be formulated for models that have higher order poles.

7.2 Models with higher order poles

The poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm presented before in Theorem 10 is extended to models with higher order poles in Section 7.2.1. Then two special cases are addressed in Sections 7.2.2 and 7.2.3.

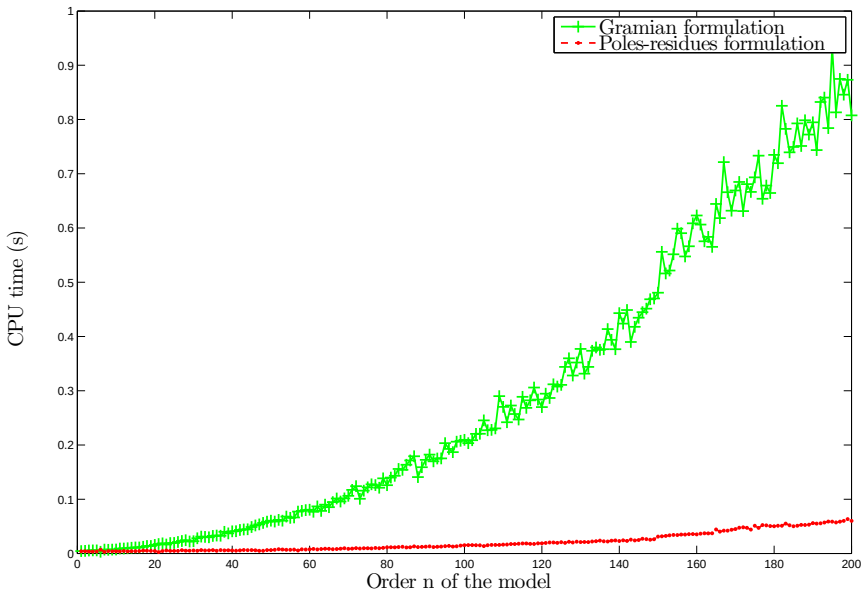


Figure 7.2: Average CPU times required to compute the $\mathcal{H}_{2,\Omega}$-norm of 50×200 randomly generated models with the gramian formulation of the norm and the poles-residues formulation for $\Omega = [0, 5]$.

7.2.1 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm for models with high order poles

The poles-residues expression of the $\mathcal{H}_{2,\Omega}$-norm for models with high order poles is presented in Theorem 12. For sake of simplicity, the Theorem is restricted to asymptotically stable and strictly proper models but these assumptions could be easily alleviated.

Theorem 12. *Given a n-th order asymptotically stable and strictly proper MIMO LTI dynamical model $\mathbf{H} := (A, B, C)$ whose transfer function is $H(s)$ and an interval $\Omega = [0, \omega]$ with $\omega > 0$. Let $\mathbf{H}$ have n_b Jordan blocks of size n_i $(i = 1, \ldots, n_b)$, each associated with the eigenvalue λ_i. Then the frequency-limited $\mathcal{H}_2$-norm of $\mathbf{H}$ is given by*

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \frac{j}{2\pi} \sum_{i=1}^{n_b} \sum_{k=1}^{n_b} \sum_{l=1}^{n_i} \sum_{m=1}^{n_k} \mathbf{tr}\left(\Phi_i^{(l)} \Phi_k^{(m)^T}\right) I_{lm}(\lambda_i, \lambda_k, \omega) \tag{7.5}$$

with

$$I_{lm}(\lambda_i, \lambda_k, \omega) = \sum_{p=1}^{l} r_{m,l-p}(\lambda_i, \lambda_k) W_{p-1}(j\omega, \lambda_i) + \sum_{q=1}^{m} r_{l,m-q}(\lambda_i, \lambda_k) W_{q-1}(j\omega, \lambda_k),$$

where

$$W_u(j\omega, \lambda) = \frac{1}{u!} \frac{d^u}{dy^u} \left(\mathbf{log}(-j\omega - y) - \mathbf{log}(j\omega - y)\right)\Big|_{y=\lambda},$$

and

$$r_{u,v}(\lambda_i, \lambda_k) = (-1)^{(u+v)} \begin{pmatrix} u+v-1 \\ v \end{pmatrix} \frac{1}{(\lambda_i + \lambda_k)^{u+v}}.$$

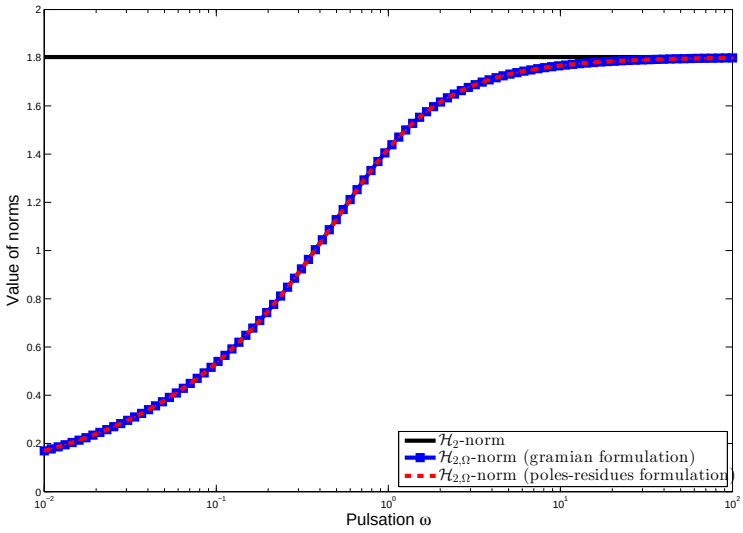


Figure 7.3: $\mathcal{H}_2$ and $\mathcal{H}_{2,\Omega}$ norms (with $\Omega = [0, \omega]$) of the model $\mathbf{H}$ for varying ω computed with the gramian and poles-residues formulations.

Proof. The idea is the same as in the semi-simple poles case excepted that the transfer function decomposition is different here. Indeed, as presented in Section 2.1.1, when a model has multiple poles, the transfer function can be decomposed as

$$H(s) = \sum_{i=1}^{n_b} \sum_{k=1}^{n_i} \frac{\Phi_i^{(k)}}{(s-\lambda_i)^k},$$

where n_b is the number of Jordan blocks and n_i $(i = 1, \ldots, n_b)$ the size of the i-th Jordan block and $\Phi_i^{(k)} \in \mathbb{C}^{n_y \times n_u}$ is the k-th residue associated with the i-th Jordan block (see Section 2.1.1). The full proof can be found in Appendix A.2. □

Similarly to the semi-simple poles case, the formulation consists in the weighted sum of cross multiplication among all the residues associated with each Jordan block. The main difference with the semi-simple case lies in the fact that the weighting function also involves the successive derivatives of the inverse tangent function.

Remark 15. *Note that the poles-residues formulation presented in Theorem 12 requires the Jordan canonical form of the matrix A in order to be computed. This decomposition is complex to obtain in practice, hence, from the author's point of view, this formulation must be considered, in this case, mainly as a theoretical tool.*

Example 14. *Let us consider the simple model* $\mathbf{H}$ *considered in Example 1 which realisation is*

$$A = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}, B = C^T = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, D = 0.$$

$\mathbf{H}$ *has a double pole* $\lambda = -1$, i.e. $n_b = 1$ *and* $n_1 = 2$. *Hence, by applying Theorem 12, the frequency-limited* $\mathcal{H}_2$ *norm of* $\mathbf{H}$ *over* $\Omega = [0, \omega]$ *can be written as,*

$$\begin{aligned} \|H\|^2_{\mathcal{H}_{2,\Omega}} &= \frac{j}{2\pi} \Big(\Phi_\lambda^{(1)} \Phi_\lambda^{(1)} I_{11}(\lambda, \lambda, \omega) + \Phi_\lambda^{(1)} \Phi_\lambda^{(2)} I_{12}(\lambda, \lambda, \omega) + \Phi_\lambda^{(2)} \Phi_\lambda^{(1)} I_{21}(\lambda, \lambda, \omega) \ldots \\ &\qquad \ldots + \Phi_\lambda^{(2)} \Phi_\lambda^{(2)} I_{22}(\lambda, \lambda, \omega) \Big), \end{aligned}$$

where $\Phi_\lambda^{(1)} = 2$ *and* $\Phi_\lambda^{(2)} = 1$. *By computing* $I_{11}(\lambda,\lambda,\omega)$, $I_{21}(\lambda,\lambda,\omega)$, $I_{12}(\lambda,\lambda,\omega)$ *and* $I_{22}(\lambda,\lambda,\omega)$, $\|H\|_{\mathcal{H}_{2,\Omega}}$ *can be expressed as an explicit function of* ω,

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \frac{-1}{4\pi(\omega^2+1)}\left(26\left(\omega^2+1\right)\mathbf{atan}\left(-\omega\right) - 10\omega\right). \tag{7.6}$$

The frequency-limited $\mathcal{H}_2$*-norm of* $\mathbf{H}$ *is plotted in Figure 7.3 against the upper bound* ω *of the frequency interval* Ω. *It is computed with equation (7.6) and with the gramian formulation of the norm (see Section 2.1.2). Obviously, both formulations are equivalent and since the model is asymptotically stable, its* $\mathcal{H}_{2,\Omega}$*-norm tends towards its* $\mathcal{H}_2$*-norm as* ω *tends towards infinity.*

7.2.2 Special case 1 : n eigenvalues of multiplicity 1 (semi-simple case)

Let us consider a stable and strictly proper model $\mathbf{H}$ with n simple eigenvalues which corresponding residues are $\Phi_i^{(1)}$ $(i = 1,\ldots,n)$. Since $n_i = 1$ $(i = 1,\ldots,n)$, the general expression of the $\mathcal{H}_{2,\Omega}$-norm (7.5) becomes

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \frac{j}{2\pi}\sum_{i=1}^{n}\sum_{k=1}^{n}\mathbf{tr}\left(\Phi_i^{(1)}{\Phi_k^{(1)}}^T\right) I_{11}(\lambda_i,\lambda_k,\omega),$$

where the weighting functions $I_{11}(\lambda_i,\lambda_k,\omega)$ are only composed of two terms

$$I_{11}(\lambda_i,\lambda_k,\omega) = r_{1,0}(\lambda_i,\lambda_k)\left(W_0\left(j\omega,\lambda_i\right) + W_0\left(j\omega,\lambda_k\right)\right),$$

with

$$W_0(j\omega,\lambda_i) = \mathbf{log}\left(-j\omega-\lambda_i\right) - \mathbf{log}\left(j\omega-\lambda_i\right) := 2j\mathbf{atan}\left(\frac{\omega}{\lambda_i}\right),$$

and

$$r_{1,0}(\lambda_i,\lambda_k) = -\frac{1}{(\lambda_i+\lambda_k)}.$$

Hence

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \frac{j}{2\pi}\sum_{i=1}^{n}\sum_{k=1}^{n} -\frac{2j\mathbf{tr}\left(\Phi_i^{(1)}{\Phi_k^{(1)}}^T\right)}{(\lambda_i+\lambda_k)}\left(\mathbf{atan}\left(\frac{\omega}{\lambda_i}\right) + \mathbf{atan}\left(\frac{\omega}{\lambda_k}\right)\right)$$

By reordering the sums so that the similar terms are grouped together, the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm for model with semi-simple poles (7.1) is retrieved :

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \frac{2}{\pi}\sum_{i=1}^{n}\sum_{k=1}^{n}\frac{\mathbf{tr}\left(\Phi_i^{(1)}{\Phi_k^{(1)}}^T\right)}{\lambda_i+\lambda_k}\mathbf{atan}\left(\frac{\omega}{\lambda_i}\right)$$

7.2.3 Special case 2 : 1 eigenvalue of multiplicity n

Given an asymptotically stable and strictly proper model $\mathbf{H}$ with one single Jordan block of size n associated with the eigenvalue λ which corresponding residues are $\Phi_\lambda^{(i)}$ $(i = 1,\ldots,n)$, the expression presented in [Antoulas, 2005, chap. 5],

$$\|H\|^2_{\mathcal{H}_2} = \mathbf{tr}\left(\sum_{i=1}^{n}\frac{\Phi_\lambda^{(i)}}{(i-1)!}\frac{d^{i-1}}{ds^{i-1}}H(-s)^T\Big|_{s=\lambda}\right), \tag{7.7}$$

is retrieved when ω tends towards infinity from equation (7.5). Indeed, first, note that, for $u > 0$,

$$W_u(j\omega,\lambda) = \frac{1}{u}\left(\frac{1}{(j\omega-\lambda)^u} - \frac{1}{(-j\omega-\lambda)^u}\right),$$

hence for $u > 0$,

$$\lim_{\omega \to \infty} |W_u(j\omega, \lambda)| = 0.$$

Besides, since $\mathbf{H}$ is assumed to be asymptotically stable (see Corollary 1),

$$\lim_{\omega \to \infty} W_0(j\omega, \lambda) = -j\pi.$$

The limits of the weighting functions $I_{lm}(\lambda, \lambda, \omega)$ ($l = 1, \ldots, n$ and $m = 1, \ldots, n$) are thus given by

$$\lim_{\omega \to \infty} I_{lm}(\lambda, \lambda, \omega) = -j2\pi r_{l,k-1}(\lambda, \lambda),$$

where

$$r_{l,k-1}(\lambda, \lambda) = (-1)^{l+k-1} \begin{pmatrix} l+k-2 \\ k-1 \end{pmatrix} \frac{1}{(2\lambda)^{l+k-1}}.$$

By replacing the weighting functions by their limits, one obtains

$$\lim_{\omega \to \infty} \|H\|^2_{\mathcal{H}_{2,\Omega}} = \sum_{k=1}^{n} \sum_{l=1}^{n} \mathbf{tr}\left(\Phi_\lambda^{(k)} {\Phi_\lambda^{(l)}}^T\right) r_{l,k-1}(\lambda, \lambda).$$

Finally, noticing that

$$\sum_{l=1}^{n} {\Phi_\lambda^{(l)}}^T r_{l,k-1}(\lambda, \lambda) = \frac{1}{(k-1)!} \frac{d^{k-1}}{ds^{k-1}} H(-s)^T \Big|_{s=\lambda},$$

leads to expression (7.7).

7.3 Upper bounds of the $\mathcal{H}_\infty$-norm

In this section, the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm is used to derive two upper bounds on the $\mathcal{H}_\infty$-norm of LTI dynamical models which are of great interest in the approximation context. Theses bounds are formulated in Section 7.3.1. Their computation is then presented in Section 7.3.2 and an experimental study of their conservatism is performed in Section 7.3.3. And finally, in Section 7.3.4, a method aimed at building a frequency template of LTI models is developed. Such a frequency template is useful in the context of model approximation to bound the approximation error so that the large-scale model can be represented by an uncertain low-order model and used for control or analysis purposes.

7.3.1 Formulation of the bounds

The link between the $\mathcal{H}_\infty$-norm (Definition 5) and the $\mathcal{H}_{2,\Omega}$-norm is presented in Theorem 13. It is grounded on the Frobenius norm which is firstly recalled in Definition 9.

Definition 9 (Frobenius norm)**.** *The Frobenius norm of a matrix $M \in \mathbb{C}^{m \times n}$ is given by*

$$\|M\|_F := \sqrt{\mathbf{tr}\left(MM^H\right)} = \sqrt{\sum_{i=1}^{min(m,n)} \sigma_i^2},$$

where the σ_i, $i = 1, \ldots, min\,(m, n)$ are the singular values of M.

Theorem 13 ($\mathcal{H}_\infty$ upper bound)**.** *Let us consider an asymptotically stable MIMO LTI dynamical model* $\mathbf{H}$ *whose transfer function is* $H(s)$*. Then its* $\mathcal{H}_\infty$*-norm is upper bounded by the derivative with respect to* ω *of its* $\mathcal{H}_{2,\Omega}$*-norm (with* $\Omega = [0, \omega]$*) as follows*

$$\|H\|_{\mathcal{H}_\infty} \leq \max_{\omega \in \mathbb{R}} \sqrt{\pi \frac{d\|H\|^2_{\mathcal{H}_{2,\Omega}}}{d\omega}}.$$

Moreover, the bound becomes an equality if $\mathbf{H}$ *is SISO.*

Proof. Considering the definition of the $\mathcal{H}_\infty$-norm, one obtains

$$\|H\|_{\mathcal{H}_\infty} = \max_{\omega \in \mathbb{R}} \sigma_{max}(H(j\omega)) \leq \max_{\omega \in \mathbb{R}} \|H(j\omega)\|_F .$$

The $\mathcal{H}_{2,\Omega}$-norm can be written as the integral of the transfer function's Frobenius norm over $[-\omega, \omega]$, *i.e.*

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \frac{1}{2\pi} \int_{-\omega}^{\omega} \|H(j\nu)\|^2_F \, d\nu.$$

Thus by differentiating this expression with respect to ω, it is clear that

$$\|H(j\omega)\|^2_F = \pi \frac{d\|H\|^2_{\mathcal{H}_{2,\Omega}}}{d\omega},$$

which concludes the proof. □

The general bound presented in Theorem 13 can be conveniently expressed based on the spectral expression of the frequency-limited $\mathcal{H}_2$-norm. Indeed, by differentiating (7.2) with respect to ω, one obtains

$$\frac{d\|H\|^2_{\mathcal{H}_{2,[0,\omega]}}}{d\omega} = -\frac{2}{\pi} \sum_{i=1}^{n} \mathbf{tr}\left(\phi_i H(-\lambda_i)^T\right) \frac{\lambda_i}{\lambda_i^2 + \omega^2} + \frac{1}{\pi} \mathbf{tr}\left(DD^T\right) = \sum_{i=1}^{n} f_i(\omega) + \frac{1}{\pi} \mathbf{tr}\left(DD^T\right).$$

Each function f_i is a scalar complex valued function and comes with its complex conjugate which makes the sum real. Therefore, only the real parts of the functions f_i need to be considered, *i.e.*

$$g_i(\omega) = \mathbf{Re}\left(f_i(\omega)\right) = \frac{a_i(x_i^2 - y_i^2 + \omega^2) + 2b_i y_i x_i}{(x_i^2 - y_i^2 + \omega^2)^2 + 4x_i^2 y_i^2}, \tag{7.8}$$

where

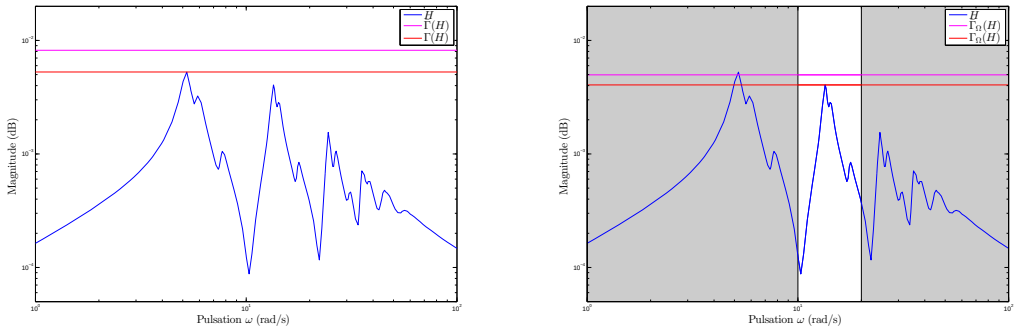
$$\begin{array}{rcl} x_i + jy_i & = & \lambda_i \quad \text{and} \\ a_i + jb_i & = & -\frac{2}{\pi} \mathbf{tr}\left(\phi_i H(-\lambda_i)^T\right) \lambda_i, \end{array} \tag{7.9}$$

The general bound presented in Theorem 13 can then be reformulated and another bound can be derived as presented in Theorem 14.

Theorem 14 (Upper bounds on the $\mathcal{H}_\infty$-norm)**.** *Given a MIMO LTI dynamical system* $\mathbf{H}$ *of order* n *with a diagonalisable realisation matrix* A*, its* $\mathcal{H}_\infty$*-norm is upper bounded by* Γ *and* $\bar{\Gamma}$ *as follows,*

$$\|H\|_{\mathcal{H}_\infty} \leq \underbrace{\sqrt{\mathbf{tr}\left(DD^T\right) + \max_{\omega \in \mathbb{R}_+} \pi \sum_{i=1}^{n} g_i(\omega)}}_{\Gamma} \leq \underbrace{\sqrt{\mathbf{tr}\left(DD^T\right) + \pi \sum_{i=1}^{n} \max_{\omega \in \mathbb{R}_+} g_i(\omega)}}_{\bar{\Gamma}}$$

where the functions g_i *are defined as in* (7.8)*.*



(a) (b)

Figure 7.4: Upper bounds Γ_Ω and $\bar{\Gamma}_\Omega$ over $\Omega_1 = [0, \infty]$ (a) and $\Omega_2 = [10, 20]$ (b) for the LAH model.

Obviously, by restricting the domain of research in Theorem 14 from $\mathbb{R}_+$ to some frequency interval Ω, these bounds can be used to bound the maximal singular value of the transfer function over Ω, *i.e.*

$$\max_{\omega\in\Omega} \sigma_{max}(H(j\omega)) \leq \underbrace{\sqrt{\mathbf{tr}\,(DD^T) + \max_{\omega\in\Omega} \pi \sum_{i=1}^{n} g_i(\omega)}}_{\Gamma_\Omega} \leq \underbrace{\sqrt{\mathbf{tr}\,(DD^T) + \pi \sum_{i=1}^{n} \max_{\omega\in\Omega} g_i(\omega)}}_{\bar{\Gamma}_\Omega}.$$

An illustration of the bounds Γ_Ω and $\bar{\Gamma}_\Omega$ is presented in Example 15. Note that the subscript Ω may be dropped when $\Omega = [0, \infty]$.

About the conservatism of Γ_Ω. From standard norm inequalities, one knows that for a matrix M of rank k,

$$\|M\|_2 \leq \|M\|_F \leq \sqrt{k}\|M\|_2,$$

which gives an indication of how conservative the bound Γ_Ω can be. In particular, by denoting $p = \min\{n_y, n_u\}$, one can say that

$$\frac{\Gamma_\Omega(H)}{\|H\|_{\mathcal{H}_\infty}} \leq \sqrt{p}.$$

Lower bound of the $\mathcal{H}_\infty$-norm for MIMO models. Let us consider a MIMO model $\mathbf{H}$. By computing Γ_Ω for each SISO transfer $H_{i,j}(s)$ of $\mathbf{H}$, one can also find a lower bound for the $\mathcal{H}_\infty$-norm,

$$\max_{\substack{i\in[1,n_y]\\ j\in[1,n_u]}} \Gamma_\Omega(H_{i,j}) \leq \|H\|_{\mathcal{H}_\infty}.$$

Example 15 (Illustration of the upper bounds – code available in appendix C.4)**.** *In this example, the bounds $\Gamma_\Omega(H)$ and $\bar{\Gamma}_\Omega(H)$ are illustrated on the LAH model. The bounds are computed for $\Omega_1 = [0, \infty]$ and for $\Omega_2 = [10, 20]$ and plotted together with the gain of the model frequency response in Figures 7.4a and 7.4b, respectively.*

Since the model is SISO, the bound Γ_{Ω_1} coincides with the $\mathcal{H}_\infty$-norm of the model and Γ_{Ω_2} coincides with the maximal singular value of the transfer function over Ω_2.

7.3.2 Computation of the bounds

Computation of $\bar{\Gamma}_\Omega$

Computing $\bar{\Gamma}_\Omega$ consists in finding the maximum of n simple rational functions over Ω. This can be achieved by evaluating these functions at a finite number of points.

Indeed, the functions g_i have no real pole and tends towards 0 as ω tends towards infinity, hence they are bounded on $\mathbb{R}$. In particular, they are bounded on $\mathbb{R}$ either by 0 when they are strictly negative, or by their value at one of the stationary points. Similarly, on Ω, they are bounded either by their value at a stationary point which belongs to Ω or by their value at one of the bound of the interval.

The stationary points of $g_i(\omega)$ can easily be calculated by differentiating with respect to ω,

$$g_i'(\omega) = \frac{2\omega N_i(\omega)}{((x_i^2 - y_i^2 + \omega^2)^2 + 4x_i^2y_i^2)^2},$$

where,

$$N_i(\omega) = (x_i^2 - y_i^2 + \omega^2)(-a_i(x_i^2 - y_i^2 + \omega^2) - 4b_iy_ix_i)4a_ix_i^2y_i^2.$$

The stationary points are then given by the zeros of the numerator, *i.e.* 0 and the roots $p_k^{(i)}$, $k = 1, \ldots, 4$ of the polynomial N_i which are

$$p_k^{(i)} = \pm\sqrt{\pm 2\frac{\sqrt{x_i^2y_i^2(a_i^2 + b_i^2)}}{a_i} - \frac{2b_ix_iy_i}{a_i} - x_i^2 + y_i^2}. \tag{7.10}$$

The different steps to compute $\bar{\Gamma}_\Omega$ are summarised in Algorithm 7.

Algorithm 7 Computation of $\bar{\Gamma}$

Require: A state-space realisation (A, B, C, D) and a frequency-interval Ω.

1: Compute the eigenvalue decomposition of A, *i.e.* $AX = X\Delta$, where $\Delta = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$.
2: **for** $i = 1, \ldots, n$ **do**
3: Compute a_i, b_i, x_i and y_i from (7.9).
4: Compute the stationary points $p_k^{(i)}$, $k = 1, \ldots, 4$, with (7.10) and keep the real ones that belong to Ω, *i.e.* $\mathcal{S} = \left\{p_k^{(i)} / \mathbf{Im}\left(p_k^{(i)}\right) = 0 \text{ and } \mathbf{Re}\left(p_k^{(i)}\right) \in \Omega\right\}$.
5: Evaluate g_i at 0 (which is also a stationary point), at the points contained in $\mathcal{S}$ and at the limits of Ω.
6: Set $\bar{g}_i$ as the maximum of the values computed at step 5.
7: **end for**
8: Set $\bar{\Gamma} = \sqrt{\pi \sum_{i=1}^n \bar{g}_i + \mathbf{tr}(DD^T)}$.

Computation of Γ_Ω

The computation of the tight bound Γ_Ω requires to find the maximum of a sum of rational functions over Ω. In that case, an optimisation procedure must be used. It can either be a dedicated method [Bugarin et al., 2011] or a generic local optimisation procedure [Fletcher, 2000; Nocedal and Wright, 1999]. Here, only the latter approach has been considered in a mixed method that couples Newton iterations and gradient descent steps with a variable step length depending on the sign of the Hessian[4]. The following remarks can be addressed about the optimisation procedure

[4] Standard solvers such as `fminunc` or `fmincon` of Matlab® have also been tested and work as well.

- when a frequency interval Ω is considered, a constrained optimisation algorithm should logically be used. However, since the maximum is either obtained at a stationary point or at one bound of Ω, it can be omitted in this simple case by initialising the algorithm inside Ω and by stopping it if the point goes outside of Ω. If the frequency interval Ω is composed of disjoint frequency intervals, then several optimisations must necessarily be performed with different starting points.

- The initial point is selected here as the strictly positive imaginary part of the eigenvalue that yields the maximum of the objective function $\mathcal{J}(\omega) = \sum_{i=1}^{n} g_i(\omega)$[5], *i.e.* by denoting λ^+ the set of eigenvalues with a positive imaginary part that belongs to Ω,

$$\omega_0 = \arg\max \left\{ \mathcal{J}(\mathbf{Im}\,(\lambda))/\lambda \in \lambda^+) \right\}.$$

  Depending on the frequency interval Ω, it can happen that the set λ^+ is empty or contains very few elements. Hence it can be completed with arbitrary points that belong to Ω. Since the objective function $\mathcal{J}$ is quite cheap to evaluate, adding points to λ^+ is not prohibitive.

  Since 0 is a stationary point of the objective function, if it belongs to the interval Ω, *i.e.* $0 \in \Omega$, then the maximum between $\mathcal{J}(0)$ and the value of the objective function after the optimisation process is kept to create Γ_Ω.

So far, this optimisation procedure with the aforementioned initialisation strategy has not failed to deliver the correct result. Yet there is no guarantee that the maximum found this way is actually the global maximum.

7.3.3 Experimental study of the bounds quality

In order to evaluate the quality of the bounds Γ and $\bar{\Gamma}$ proposed in Theorem 14, three tests are performed.

- Firstly, to show the good scalability of the method, the first test consists in computing the $\mathcal{H}_\infty$-norm (with Matlab® routine `norm` with default tolerance), Γ and $\bar{\Gamma}$ of randomly generated models which order n varies from 1 to 200 with $n_y = n_u = 5$. For each order n, 50 random models are generated and the CPU time required by each routine is measured. The average of these CPU times over the models having the same order n is reported in Figure 7.5. In Table 7.1 (left), the ratios of the bounds over the real $\mathcal{H}_\infty$-norm (computed with Matlab® routine `norm` with a tolerance equal to the machine precision), *i.e.* $r_\Gamma = \frac{\Gamma(H)}{\|H\|_{\mathcal{H}_\infty}}$ and $r_{\bar{\Gamma}} = \frac{\bar{\Gamma}(H)}{\|H\|_{\mathcal{H}_\infty}}$ are presented.

- Given the nature of the bounds, the number of inputs and outputs is more likely to impact their quality. That is why, in the second test, the order n or the random models is fixed to 20 while their number of inputs and outputs is increased from 1 to 200 (with $n_y = n_u$). Again, for each number of inputs/outputs, 50 models are generated. The average CPU times are plotted in Figure 7.6 and the average ratios of the bounds over the $\mathcal{H}_\infty$-norm are plotted in Figure 7.7. Those ratios are also presented in Table 7.1 (right).

- The bounds Γ and $\bar{\Gamma}$ are compared to the bound provided by the Balanced Truncation (see Section 3.1.2) by reducing the ISS model for several orders r going from 2 to 30. For each reduced-order r, the $\mathcal{H}_\infty$-norm of the approximation error and the various bounds are computed and plotted in Figure 7.8.

Figure 7.5 shows that the computation time of the bounds is less impacted by the order of the model than Matlab® routine. For models of order $n > 50$, computing both bounds takes

[5]The constant term $\frac{1}{\pi}\mathbf{tr}\left(DD^T\right)$ can be discarded during the optimisation process

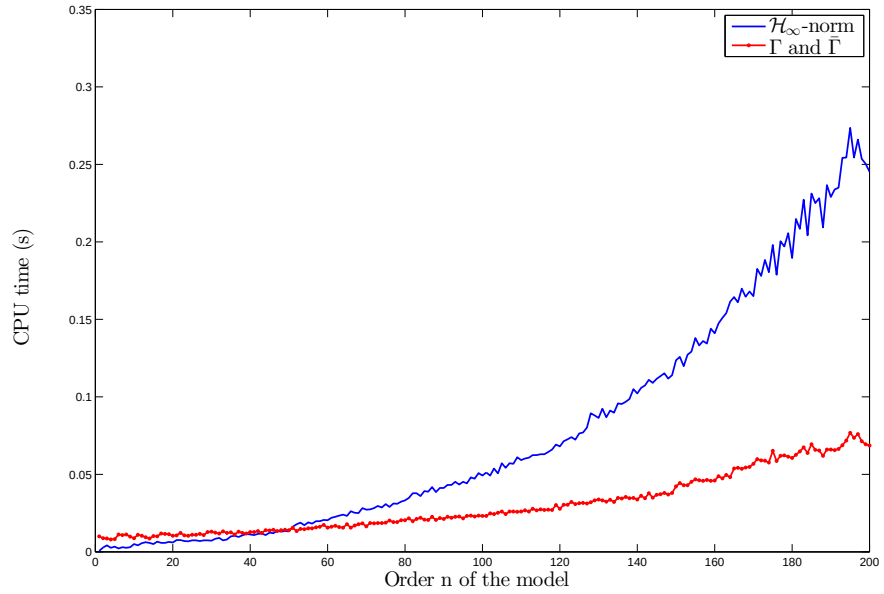


Figure 7.5: Average CPU times required to compute the $\mathcal{H}_\infty$-norm and Γ, $\bar{\Gamma}$ for random models which order varies from 1 to 200.

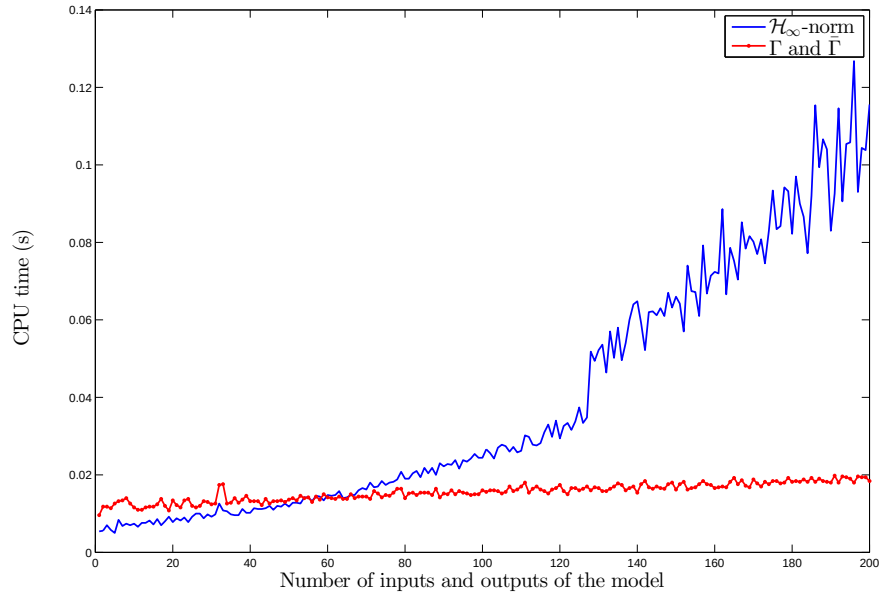


Figure 7.6: Average CPU times required to compute the $\mathcal{H}_\infty$-norm and Γ, $\bar{\Gamma}$ for random models which number of inputs and outputs varies from 1 to 200.

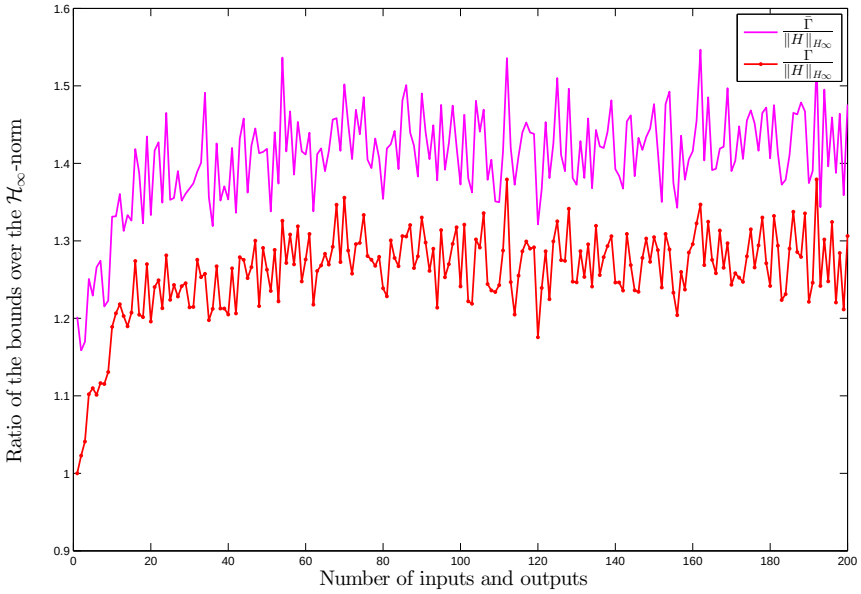


Figure 7.7: Average ratios between the bounds and the $\mathcal{H}_\infty$-norm for random models which number of inputs and outputs varies from 1 to 200.

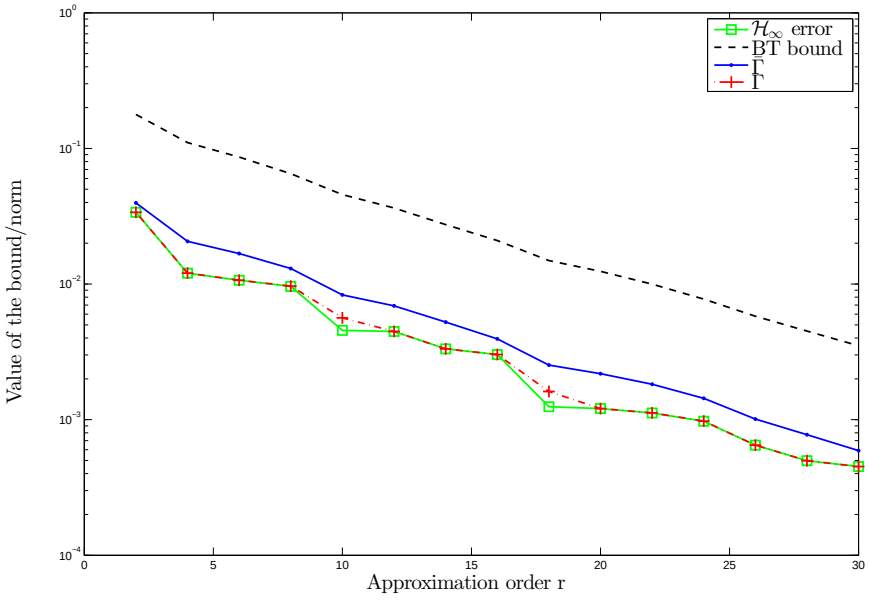


Figure 7.8: Comparison of the bounds Γ and $\bar{\Gamma}$ with the bound of the **BT** on the ISS model.

varying n	r_Γ	$r_{\bar{\Gamma}}$	varying n_y, n_u	r_Γ	$r_{\bar{\Gamma}}$
min	1.00	1.00	min	1.00	1.00
mean	1.12	1.24	mean	1.26	1.41
max	1.61	2.23	max	2.29	2.55
var	0.015	0.11	var	0.067	0.11

Table 7.1: Ratios of the bounds over the real $\mathcal{H}_\infty$-norm for varying order n (left) and varying number of inputs/outputs (right).

in average, less time than computing the $\mathcal{H}_\infty$-norm. The same conclusion can be made with respect to the number of inputs and outputs as illustrated in Figure 7.6. Indeed, increasing the number of inputs and outputs barely increase the CPU time required to compute the bounds. This is not surprising since the number of inputs and outputs has an impact only once in the computation of the trace in the functions $g_i(\omega)$.

From Table 7.1, one can see that on average, the bounds $\Gamma_\Omega(H)$ and $\bar{\Gamma}_\Omega(H)$ are not extremely conservative. As expected, they become more conservative as the number of inputs and outputs increase. Indeed, the mean of the ratios r_Γ and $r_{\bar{\Gamma}}$ increase from 1.12 and 1.24 in the case of models with $n_y = n_u = 5$ (left table) to 1.26 and 1.41 in the case of models with varying number of inputs and outputs (right table). As one can observe in Figure 7.7 this increase happens mainly as the number of inputs and outputs goes from 1 to 20, but then, there is no clear increase anymore. This may be due to the way random models are generated in Matlab® though.

From Figure 7.8, one can see that both Γ and $\bar{\Gamma}$ are tighter than the bound given by the balanced truncation on this example and Γ often coincides with the real $\mathcal{H}_\infty$-norm. Yet, it is important to remember that the proposed bounds cannot be computed a priori since they require the reduced-order model to be built whereas the bound given by the **BT** can be evaluated for several approximation orders r once the gramians have been computed.

7.3.4 Construction of a frequency template

In the context of model approximation, the bound $\Gamma(H - \hat{H})$ developed in the previous section gives information about the worst error induced by the approximation between $\mathbf{H}$ and $\hat{\mathbf{H}}$. With that bound, one could directly build an uncertain reduced-order model which encompasses all the dynamics of the large-scale model and which could then be used to perform robust control or analysis in place of the initial large-scale model. Yet, such a static bound would be terribly conservative and would not provide a sufficiently accurate representation of the approximation error model. The method developed in this section is aimed at providing a more accurate representation of the approximation error between $\mathbf{H}$ and $\hat{\mathbf{H}}$.

In particular the bound Γ_Ω is exploited to develop a method aimed at building a dynamic model $\mathbf{G}$ of *low complexity* which upper bounds another dynamical model $\mathbf{H}$. For simplicity, let us assume that $\mathbf{H}$ is SISO, then the problem consists in finding a minimum phase model $\mathbf{G}$ whose transfer function $G(s)$ satisfies, for all $\omega \in \mathbb{R}_+$

$$|G(j\omega)| \geq |H(j\omega)|. \tag{7.11}$$

Since this general problem is complex, a practical approach is considered here. In particular, the problem is simplified, firstly by imposing a specific structure to the transfer function $G(s)$,

$$G(s) = K \left(\frac{s - z}{s - p} \right)^r, \tag{7.12}$$

with $\mathbf{Re}(p) < 0$ and $\mathbf{Re}(z) < 0$. This structure is extremely restrictive but makes the process way easier. Secondly, the constraint (7.11) is restricted to some relevant pulsations ω_k, $k = 1, \dots, N$,

$$|G(j\omega_k)| \geq |H(j\omega_k)|. \tag{7.13}$$

These relevant pulsations ω_k are selected as the resonant pulsations, smaller than the pulsation ω_∞ where the $\mathcal{H}_\infty$-norm is reached, at which $|H(j\omega)|$ is the largest over the interval $[0, \omega_k]$. These frequencies are determined by computing $\Gamma_\Omega(H)$ where the upper bound ω of Ω decreases progressively from ω_∞ to 0. This process is described in Algorithm 8. At step 1, the $\mathcal{H}_\infty$-norm and the corresponding pulsation ω_∞ are computed. The pulsation ω_∞ is the first resonant pulsation ω_1. Then at each iteration, the next resonant pulsation ω_{k+1} that yields $\tau_{k+1} = \Gamma_\Omega(H)$, $\Omega = [0, \omega_{k+1}]$ is computed by Algorithm 9. The loop stops when the next resonant pulsation is smaller or equal to 0.

Algorithm 8 `getAllMax`

Require: A LTI dynamical model $\mathbf{H}$.

1: Compute $\Gamma(H)$ and the associated pulsation ω_∞.
2: Set $\tau_1 = \Gamma(H)$ and $\omega_1 = \omega_\infty$.
3: $i = 1$
4: **while** $\omega_k > 0$ **do**
5: Compute the next resonant point with Algorithm 9, *i.e.*

$$[\tau_{k+1}, \omega_{k+1}] = \texttt{getNextMax}(H, \omega_k).$$

6: $k \leftarrow k + 1$.
7: **end while**
8: **return** $\boldsymbol{\tau}$, $\boldsymbol{\omega}$

Algorithm 9 `getNextMax`

Require: A LTI dynamical model $\mathbf{H}$, an initial pulsation ω_i.

1: Choose $\rho > 0$.
2: Set $\bar{\omega} = \omega_i$ and $\omega = \omega_i$.
3: **while** $\bar{\omega} = \omega$ **do**
4: Set $\omega \leftarrow \omega - \rho$ and $\Omega = [0, \omega]$.
5: Compute $\Gamma_\Omega(H)$ and the associated pulsation $\bar{\omega}$.
6: **end while**
7: **return** Γ_Ω and $\bar{\omega}$.

Given a dynamical model $\mathbf{H}$ and an initial pulsation ω_i, Algorithm 9 finds the next resonant pulsation $\bar{\omega} < \omega_i$ which yields $\Gamma_\Omega(H)$ where $\Omega = [0, \bar{\omega}]$. This is achieved by computing $\Gamma_\Omega(H)$, $\Omega = [0, \omega]$, with a progressively decreasing upper bound ω. The algorithm stops when the pulsation where the maximum is reached returned by the computation of $\Gamma_\Omega(H)$ is no longer equal to the upper bound ω of the interval Ω. The algorithm is illustrated in Figure 7.9 where the red star is located at the pulsation ω_i, the orange one at $\bar{\omega}$ and the vertical dashed blue lines represent the successive values of ω. For the first values of ω, the maximum $\Gamma_{[0,\omega]}(H)$ is reached at the upper bound ω of the interval, but this is not a relevant frequency because it belongs to the resonant mode located at ω_i, that is why the interval is reduced until the maximum is reached at a point that is not the upper bound of the interval, *i.e.* $\bar{\omega} \neq \omega$. The decrease factor ρ is a tuning parameter that must not be too large in order to avoid to miss some relevant pulsations but not too small either to avoid too many iterations, here it is set to be $0.01\omega_i$.

There are several different approaches that can be considered to determine the gain K, the pole p and the zero z of $G(s)$ (7.12). Here, a simple heuristic approach aimed at constructing

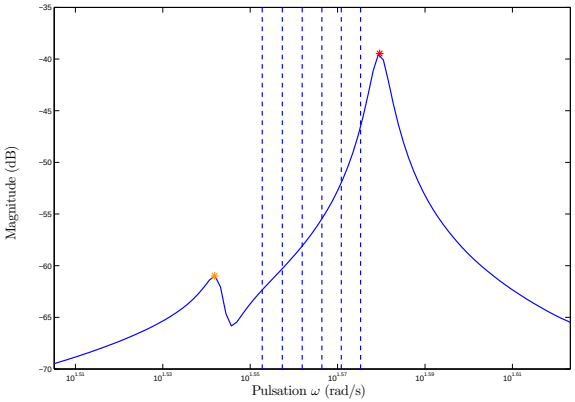


Figure 7.9: Illustration of Algorithm 8.

$\mathbf{G}$ such that $|G(j\omega)|$ is the smallest possible in $[0, \omega_\infty]$ is considered. This choice is motivated by the fact that low frequency are generally important from a physical point of view, hence the approximation is often performed over $[0, \omega]$ which yields a lowest error in that interval.

The approach firstly consists in imposing that the static gain of the frequency template $\mathbf{G}$ is equal to the non-null static gain of the model to be bounded, *i.e.*

$$G(0) = H(0). \tag{7.14}$$

If $H(0) = 0$, one can set an arbitrary value for $G(0)$. This constraint enables to determine one of the three parameters when the other two are fixed, for instance, for each couple $\{K, p\}$, one can determine the zero z as

$$z = \left(\frac{H(0)}{K}\right)^{\frac{1}{r}} p. \tag{7.15}$$

The remaining two parameters K and p are slowly relaxed from an initial value that violates the constraints (7.13) until they are all satisfied. In particular, K is initialised as $K_i = \Gamma(H) = \|H\|_{\mathcal{H}_\infty}$, which means that $G(\infty) = \|H\|_{\mathcal{H}_\infty}$ and p is initialised as $p_i = -\omega_\infty$ which sets the cutoff frequency. Obviously, the initial frequency template $\mathbf{G}_i$ cannot satisfy the constraints (7.13).

The way p and K are iteratively relaxed is a tuning parameter. Indeed, increasing p makes the frequency template slide to the low frequency, while increasing K makes its value at ∞ increases. Since the high frequency are generally less important, here, K is increased faster than p is decreased. The order r of the template has not been mentioned but it can be used to increase the slope of the frequency template.

The complete process to build a frequency template $\mathbf{G}$ with the structure (7.12) that satisfies the constraints (7.13) and (7.14) is summarised in Algorithm 10.

As mentioned before, this method is mainly an heuristic approach that does not offer a strict guarantee that the general constraint (7.11) is really satisfied for all $\omega \in \mathbb{R}_+$. Yet if no relevant pulsation has been missed by Algorithm 8, that should not happen. The overall approach is illustrated in Example 16.

Example 16 (Illustration of the frequency template – code available in Appendix C.5)**.** *To illustrate the proposed approach, the transfer from the first input to the first output of the ISS is reduced to an order* 20 *over* $\Omega = [0, 10]$ *and a frequency template of order* $r = 2$ *is built on the corresponding approximation error. The two main steps of Algorithm 10 are the following :*

Algorithm 10 Algorithm for the construction of frequency template

Require: A LTI dynamical model $\mathbf{H}$, the order r of the frequency template.

1: Compute all the relevant pulsations $\boldsymbol{\omega} \in \mathbb{R}^K$ an associated gains $\boldsymbol{\tau} \in \mathbb{R}^K$ with Algorithm 8.
2: Set $p_1 = -\max(\boldsymbol{\omega})$ and $K_1 = \max(\boldsymbol{\tau})$.
3: Compute z_1 as in equation (7.15).
4: Build G_1 as in equation (7.12).
5: $k = 1$.
6: **while** the constraints (7.13) are not satisfied **do**
7: $\quad p_{i+1} \leftarrow -\alpha|p_i|$ with $0 < \alpha < 1$.
8: $\quad K_{i+1} \leftarrow \beta K_i$ with $\beta > 1$.
9: $\quad$ Update the zero z_{i+1} as in equation (7.15) with p_{i+1} and K_{i+1}.
10: $\quad$ Build $G_{i+1}(s)$ as in equation (7.12).
11: $\quad i \leftarrow i + 1$.
12: **end while**
13: **return** The frequency template $\mathbf{G}$.

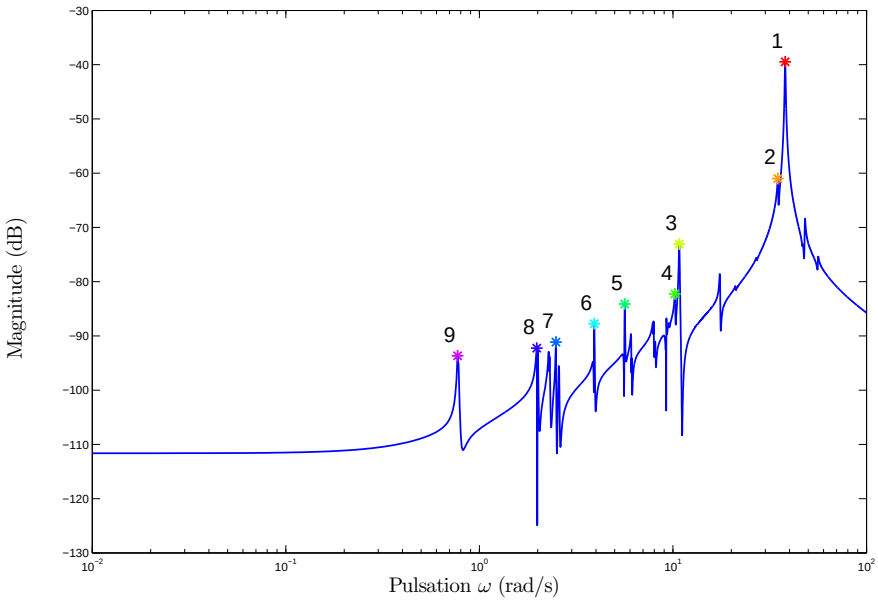


Figure 7.10: Illustration of the Algorithm 10, computation of all the relevant pulsations.

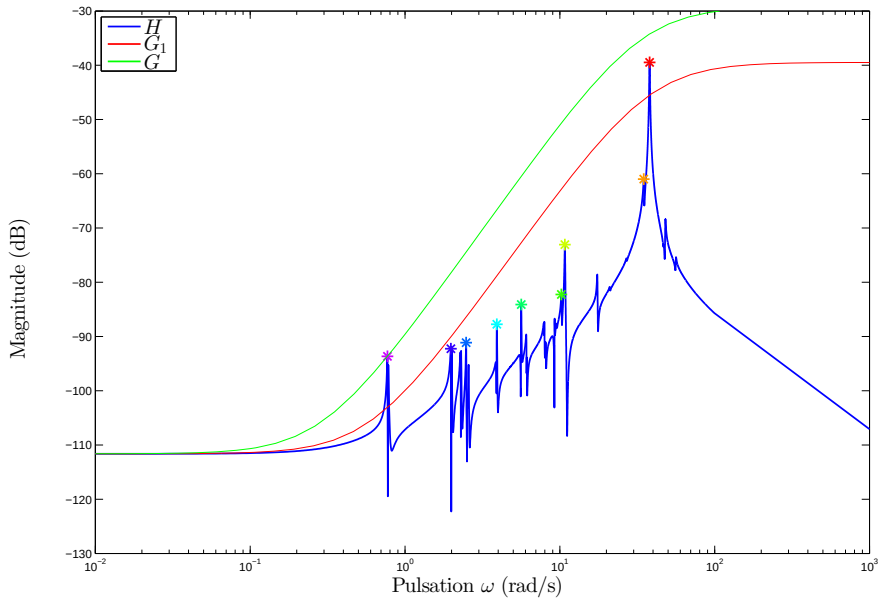


Figure 7.11: Illustration of the Algorithm 10, construction of the template G from G_i.

- *Firstly, the Algorithm 8 is used in order to compute all the relevant pulsations, that is to say the pulsations at which the model is resonant and bounds its response for smaller pulsations. This is illustrated in Figure 7.10 where the numbers indicates the order in which the points have been found. The $\mathcal{H}_\infty$-norm is found first, then the other maximums that are located at lower frequencies are computed. One can see that for each i, $|H(j\omega_i)| > |H(j\omega)|$ for $\omega \in [0, \omega_i]$.*

- *Then, the main loop of Algorithm 10 is executed and transforms the initial frequency template $\mathbf{G}_1$ by successive relaxation of its gain and pole, K and p, into the final one $\mathbf{G}$. Both frequency templates are plotted in Figure 7.11 where $\mathbf{G}_1$ is the red curve and $\mathbf{G}$ is the green one.*

MIMO case

For MIMO models, two uses of Algorithm 10 can be considered,

- firstly, it can be used directly of the whole MIMO model. In that case however, the absolute values are replaced by Frobenius norms and the transfer function of the frequency template $\mathbf{G}$ satisfies

$$\begin{array}{rcl} \|G(j\omega_k)\|_F & \geq & \|H(j\omega_k)\|_F\,,\ k = 1, \ldots, N, \\ \|G(0)\|_F & = & \|H(0)\|_F\,, \end{array} \tag{7.16}$$

  instead of (7.13) and (7.14). Equation (7.16) induces more conservatism than in the SISO case, but it enables to exploit the frequency template for control purposes (see Section 9.2.3).

- Secondly, Algorithm 10 can be used on each SISO transfers separately. This can have an interest for simulation since less conservatism than in the first case is introduced.

Conclusion

In this Chapter, the poles-residues formulation of the $\mathcal{H}_2$-norm has been extended to the $\mathcal{H}_{2,\Omega}$-norm both for models with semi-simple poles only and for models with higher order poles. This has been done by using the partial fraction decomposition of the transfer function in the norm definition. In the case of models with high order poles, the poles-residues formulation is merely a theoretical tool since it requires the Jordan decomposition of the matrix A, but in the semi-simple case, it offers some numerical advantage in terms of computation time in comparison to the gramian formulation.

Besides, it has been shown that the $\mathcal{H}_{2,\Omega}$-norm can be used to bound the $\mathcal{H}_\infty$-norm. In particular, the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm has enabled to derive two bounds which can be computed either analytically or by an optimisation procedure. The quality of these bounds has been highlighted by a numerical study.

Finally, the optimisation procedure used in the computation of the tight bound has also been used in order to develop a method aimed at building a frequency template of LTI dynamical models. These frequency-template can be exploited in the context of control or analysis of large-scale models to dynamically bound the approximation error as illustrated in Chapters 9 and 10.

The poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm is used in the next chapter to express the $\mathcal{H}_{2,\Omega}$ approximation error as a function of the poles and residues of the reduced-order model so that the approximation problem can be reformulated as an optimisation problems in terms of these variables.

Chapter 8

Formulation of the $\mathcal{H}_{2,\Omega}$ error and differentiation of its gradient

In this Chapter, the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm developed in Chapter 7 is exploited to express the $\mathcal{H}_{2,\Omega}$-norm of the approximation error between a large-scale LTI model $\mathbf{H}$ and a reduced-order model $\hat{\mathbf{H}}$ which have both semi-simple poles only. The approach followed in this Chapter is very similar to what have been done to introduce the optimal $\mathcal{H}_2$ model approximation problem in Chapter 4. The similarities and the key differences between the $\mathcal{H}_{2,\Omega}$ and $\mathcal{H}_2$ cases are highlighted.

In Section 8.1, the formulation of the error is presented together with some practical implementation concerns and in Section 8.2, the approximation error is differentiated in order to express the first-order optimality conditions of the optimal $\mathcal{H}_{2,\Omega}$ approximation problem.

The elements presented throughout this Chapter have partly been published in [Vuillemin et al., 2014b].

Contents

8.1 Expression of the approximation error with the poles and residues of the models . . . 111
8.1.1 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$ approximation error . . . 111
8.1.2 On the computation of the approximation error . . . 117
8.2 Gradient of the error . . . 119
8.2.1 Reminder on Wirtinger Calculus . . . 119
8.2.2 Gradient of the $\mathcal{H}_{2,\Omega}$ approximation error . . . 122

8.1 Expression of the approximation error with the poles and residues of the models

This Section is aimed at presenting the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$ approximation error. The theoretical aspects are presented in Section 8.1.1 while the numerical ones are discussed in Section 8.1.2.

8.1.1 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$ approximation error

The poles-residues expression of the squared approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \|H - \hat{H}\|^2_{\mathcal{H}_{2,\Omega}}$ between a large-scale model and a reduced-order one is presented in Theorem 15 for MIMO models and in Remark 16, the SISO case is highlighted. In order to shorten the equations, the following notation is used from now on,

$$\mathbf{a}_{\omega,\lambda} = \frac{2}{\pi}\mathbf{atan}\left(\frac{\omega}{\lambda}\right).$$

Theorem 15 (Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$ approximation error)**.** *Let us consider a frequency interval $\Omega = [0, \omega]$, a n-th order LTI dynamical model $\mathbf{H}$ and a r-th order LTI dynamical model $\hat{\mathbf{H}}$ which have semi-simple poles only and whose transfer functions are respectively*

written as

$$H(s) = \sum_{i=1}^{n} \frac{\Phi_i}{s-\lambda_i} + D \in \mathbb{C}^{n_y \times n_u} \text{ and } \hat{H}(s) = \sum_{i=1}^{r} \frac{\hat{\Phi}_i}{s-\hat{\lambda}_i} + \hat{D} \in \mathbb{C}^{n_y \times n_u}.$$

Let us also assume that $\lambda_i + \lambda_k \neq 0$, $\lambda_i + \hat{\lambda}_l \neq 0$ *and* $\hat{\lambda}_l + \hat{\lambda}_m \neq 0$ *for all* $i, k = 1, \ldots, n$ *and* $l, m = 1, \ldots, r$. *Then, the squared* $\mathcal{H}_{2,\Omega}$*-norm of the error between* $\mathbf{H}$ *and* $\hat{\mathbf{H}}$, i.e. $\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \|H - \hat{H}\|^2_{\mathcal{H}_{2,\Omega}}$, *is given by*

$$\begin{aligned}\mathcal{J}_{\mathcal{H}_{2,\Omega}} \quad = \quad & \sum_{i=1}^{n}\sum_{k=1}^{n} \frac{\mathbf{tr}\left(\Phi_i \Phi_k^T\right)}{\lambda_i + \lambda_k} \mathbf{a}_{\omega,\lambda_i} + \sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\hat{\Phi}_i \hat{\Phi}_k^T\right)}{\hat{\lambda}_i + \hat{\lambda}_k} \mathbf{a}_{\omega,\hat{\lambda}_i} - \sum_{i=1}^{n}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\Phi_i \hat{\Phi}_k^T\right)}{\lambda_i + \hat{\lambda}_k} \left(\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}\right) \\ & \ldots + \sum_{i=1}^{r} \mathbf{tr}\left(\hat{\Phi}_i \tilde{D}^T\right) \mathbf{a}_{\omega,\hat{\lambda}_i} - \sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i \tilde{D}^T\right) \mathbf{a}_{\omega,\lambda_i} + \frac{\omega}{\pi} \mathbf{tr}\left(\tilde{D}\tilde{D}^T\right), \end{aligned} \tag{8.1}$$

where $\mathbf{a}_{\omega,\lambda} = \frac{2}{\pi}\mathbf{atan}\left(\frac{\omega}{\lambda}\right)$ *and* $\tilde{D} = D - \hat{D}$.

Proof. The proof simply consists in writing the $\mathcal{H}_{2,\Omega}$-norm of the approximation error in terms of the residues and poles of the error model and then to separate the various sums. The full proof can be found in Appendix A.3. □

The assumptions $\lambda_i + \lambda_k \neq 0$, $\lambda_i + \hat{\lambda}_l \neq 0$ and $\hat{\lambda}_l + \hat{\lambda}_m \neq 0$ for all $i, k = 1, \ldots, n$ and $l, m = 1, \ldots, r$ in Theorem 15 prevent the denominators in the various terms involved in $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ in equation (8.1) to vanish. It can be easily alleviated by considering the complete poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm presented in Theorem 10 which consists in replacing the elements in the sums where two poles vanish, *i.e.* $\lambda_1 + \lambda_2 = 0$, by $-\frac{\omega}{\pi(\omega^2 - \lambda_1\lambda_2)}$. As an illustration, the general case is considered in Example 17. But to keep the notations simple and since the cases for which the denominators can vanish are quite unusual in practice, the assumptions made in Theorem 15 are maintained in the sequel and is summarised by saying that *the poles are not located symmetrically with respect to the imaginary axis* which includes the case of poles on the imaginary axis.

Remark 16 (SISO case)**.** *In the SISO case, the residues and the direct feedthrough are scalars,* i.e. $\Phi_i = \phi_i \in \mathbb{C}$ *(*$i = 1, \ldots, n$*),* $\hat{\Phi}_k = \hat{\phi}_k \in \mathbb{C}$ *(*$k = 1, \ldots, r$*) and* $\tilde{D} = \tilde{d} \in \mathbb{R}$. *Hence, the expression of the* $\mathcal{H}_{2,\Omega}$ *approximation error* $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ (8.1) *becomes*

$$\begin{aligned}\mathcal{J}_{\mathcal{H}_{2,\Omega}} \quad = \quad & \sum_{i=1}^{n}\sum_{k=1}^{n} \frac{\phi_i \phi_k}{\lambda_i + \lambda_k} \mathbf{a}_{\omega,\lambda_i} + \sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\hat{\phi}_i \hat{\phi}_k}{\hat{\lambda}_i + \hat{\lambda}_k} \mathbf{a}_{\omega,\hat{\lambda}_i} - \sum_{i=1}^{n}\sum_{k=1}^{r} \frac{\phi_i \hat{\phi}_k}{\lambda_i + \hat{\lambda}_k} \left(\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}\right) \ldots \\ & + \sum_{i=1}^{r} \hat{\phi}_i \tilde{d} \mathbf{a}_{\omega,\hat{\lambda}_i} - \sum_{i=1}^{n} \phi_i \tilde{d} \mathbf{a}_{\omega,\lambda_i} + \frac{\omega}{\pi} \tilde{d}^2. \end{aligned} \tag{8.2}$$

For simplicity, let us now assume that $\tilde{d} = 0$. *By reordering the sums in the the reduced-order model* $\mathcal{H}_{2,\Omega}$*-norm (second term),*

$$\|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}} = \sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\hat{\phi}_i \hat{\phi}_k}{\hat{\lambda}_i + \hat{\lambda}_k} \mathbf{a}_{\omega,\hat{\lambda}_i} = \sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\hat{\phi}_i \hat{\phi}_k}{\hat{\lambda}_i + \hat{\lambda}_k} \frac{1}{2} \left(\mathbf{a}_{\omega,\hat{\lambda}_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}\right),$$

equation (8.2) *can be re-written as a quadratic function of the residues like in the* $\mathcal{H}_2$ *case (see Remark 10),*

$$\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \frac{1}{2} \begin{bmatrix} \hat{\phi}_1 & \ldots & \hat{\phi}_r \end{bmatrix} L \begin{bmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_r \end{bmatrix} + h^T \begin{bmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_r \end{bmatrix} + \|H\|^2_{\mathcal{H}_{2,\Omega}},$$

where the symmetric matrix L is defined as

$$L = \begin{bmatrix} \frac{\mathbf{a}_{\omega,\hat{\lambda}_1}}{\hat{\lambda}_1} & \cdots & \frac{\mathbf{a}_{\omega,\hat{\lambda}_1}+\mathbf{a}_{\omega,\hat{\lambda}_r}}{\hat{\lambda}_1+\hat{\lambda}_r} \\ \vdots & \ddots & \vdots \\ \frac{\mathbf{a}_{\omega,\hat{\lambda}_r}+\mathbf{a}_{\omega,\hat{\lambda}_1}}{\hat{\lambda}_r+\hat{\lambda}_1} & \cdots & \frac{\mathbf{a}_{\omega,\hat{\lambda}_r}}{\hat{\lambda}_r} \end{bmatrix} \in \mathbb{C}^{r\times r}, \tag{8.3}$$

and

$$h = -\begin{bmatrix} \sum_{i=1}^{n} \frac{\phi_i}{\lambda_i+\hat{\lambda}_1}\left(\mathbf{a}_{\omega,\lambda_i}+\mathbf{a}_{\omega,\hat{\lambda}_1}\right) & \cdots & \sum_{i=1}^{n} \frac{\phi_i}{\lambda_i+\hat{\lambda}_r}\left(\mathbf{a}_{\omega,\lambda_i}+\mathbf{a}_{\omega,\hat{\lambda}_r}\right) \end{bmatrix}^T \in \mathbb{C}^{r\times 1}. \tag{8.4}$$

For the same reasons as in the $\mathcal{H}_2$ case presented in Remark 8.2, for SISO models, the optimal residues can be determined by solving a linear system. In that case however, L is no longer a Cauchy matrix.

Simplification of the expression of the error

The expression of the error from equation (8.1) can be reformulated in order to make the $\mathcal{H}_{2,\Omega}$-norms of the large-scale and reduced-order models to appear. First, let us expand $\tilde{D} = D - \hat{D}$,

$$\begin{aligned} \sum_{i=1}^{r} \mathbf{tr}\left(\hat{\Phi}_i \tilde{D}^T\right)\mathbf{a}_{\omega,\hat{\lambda}_i} &= \sum_{i=1}^{r} \mathbf{tr}\left(\hat{\Phi}_i D^T\right)\mathbf{a}_{\omega,\hat{\lambda}_i} - \sum_{i=1}^{r} \mathbf{tr}\left(\hat{\Phi}_i \hat{D}^T\right)\mathbf{a}_{\omega,\hat{\lambda}_i}, \\ -\sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i \tilde{D}^T\right)\mathbf{a}_{\omega,\lambda_i} &= -\sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i D^T\right)\mathbf{a}_{\omega,\lambda_i} + \sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i \hat{D}\right)\mathbf{a}_{\omega,\lambda_i}, \\ \frac{\omega}{\pi}\mathbf{tr}\left(\tilde{D}\tilde{D}^T\right) &= \frac{\omega}{\pi}\mathbf{tr}\left(DD^T\right) + \frac{\omega}{\pi}\mathbf{tr}\left(\hat{D}\hat{D}^T\right) - 2\frac{\omega}{\pi}\mathbf{tr}\left(D\hat{D}^T\right). \end{aligned}$$

Then by noticing that

$$\sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\hat{\Phi}_i\hat{\Phi}_k^T\right)}{\hat{\lambda}_i+\hat{\lambda}_k}\mathbf{a}_{\omega,\hat{\lambda}_i} + \frac{\omega}{\pi}\mathbf{tr}\left(\hat{D}\hat{D}^T\right) - \sum_{i=1}^{r}\mathbf{tr}\left(\hat{\Phi}_i\hat{D}^T\right)\mathbf{a}_{\omega,\hat{\lambda}_i} = \|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}},$$

and that

$$\sum_{i=1}^{n}\sum_{k=1}^{n} \frac{\mathbf{tr}\left(\Phi_i\Phi_k^T\right)}{\lambda_i+\lambda_k}\mathbf{a}_{\omega,\lambda_i} + \frac{\omega}{\pi}\mathbf{tr}\left(DD^T\right) - \sum_{i=1}^{n}\mathbf{tr}\left(\Phi_i D^T\right)\mathbf{a}_{\omega,\lambda_i} = \|H\|^2_{\mathcal{H}_{2,\Omega}},$$

one obtains

$$\begin{aligned} \mathcal{J}_{\mathcal{H}_{2,\Omega}} &= \|H\|^2_{\mathcal{H}_{2,\Omega}} + \|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}} - \sum_{i=1}^{n}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\Phi_i\hat{\Phi}_k^T\right)}{\lambda_i+\hat{\lambda}_k}\left(\mathbf{a}_{\omega,\lambda_i}+\mathbf{a}_{\omega,\hat{\lambda}_k}\right) + \sum_{i=1}^{n}\mathbf{tr}\left(\Phi_i\hat{D}^T\right)\mathbf{a}_{\omega,\lambda_i} \\ &\qquad \ldots + \sum_{i=1}^{r}\mathbf{tr}\left(\hat{\Phi}_i D^T\right)\mathbf{a}_{\omega,\hat{\lambda}_i} - 2\frac{\omega}{\pi}\mathbf{tr}\left(D\hat{D}^T\right). \end{aligned} \tag{8.5}$$

If both D and $\hat{D}$ are null, then equation (8.5) simplifies to

$$\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \|H\|^2_{\mathcal{H}_{2,\Omega}} + \|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}} - \sum_{i=1}^{n}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\Phi_i\hat{\Phi}_k^T\right)}{\lambda_i+\hat{\lambda}_k}\left(\mathbf{a}_{\omega,\lambda_i}+\mathbf{a}_{\omega,\hat{\lambda}_k}\right),$$

which is similar to what is obtained in the $\mathcal{H}_2$ case (see equation (4.3)), except that here, the weighting functions $\mathbf{a}_{\omega,\lambda_i}$ prevent from writing the sum in i in the last term as $H(-\hat{\lambda}_k)$. Yet, by defining

$$H_\omega(s) = \sum_{i=1}^{n} \frac{\Phi_i}{s-\lambda_i}\mathbf{a}_{\omega,\lambda_i},$$

the approximation error finally becomes

$$\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \|H\|^2_{\mathcal{H}_{2,\Omega}} + \|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}} + \sum_{i=1}^{r} \mathbf{tr}\left(\left(H_\omega(-\hat{\lambda}_k) + H(-\hat{\lambda}_k)\mathbf{a}_{\omega,\hat{\lambda}_k}\right)\hat{\Phi}_k^T\right). \tag{8.6}$$

If both $\mathbf{H}$ and $\hat{\mathbf{H}}$ are asymptotically stable, then $\lim_{\omega\to\infty} \mathbf{a}_{\omega,\lambda_i} = \lim_{\omega\to\infty} \mathbf{a}_{\omega,\hat{\lambda}_k} = -1$ thus the $\mathcal{H}_2$ approximation error presented in equation (4.3) is retrieved.

Behaviour of the $\mathcal{H}_{2,\Omega}$ approximation error as some poles cross the imaginary axis

The behaviour of the $\mathcal{H}_{2,\Omega}$ approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ as the reduced-order model poles $\hat{\lambda}_k$, $k = 1, \ldots, r$ cross the imaginary axis is not clear in the general case but is highlighted when $\mathbf{H}$ is an asymptotically stable model and $\hat{\mathbf{H}}$ is a second-order model in Example 17. In that example, it is shown that crossing the imaginary axis from the left half-plane to the right half-plane is unlikely to lead to a decrease of the $\mathcal{H}_{2,\Omega}$ approximation error when the large-scale model is asymptotically stable. This is obviously not necessarily true anymore if the large-scale model is unstable. Nevertheless, approximating an unstable model with a stable one does not appear to be desirable, especially for control purpose.

Example 17 (Behaviour of the $\mathcal{H}_{2,\Omega}$ approximation error when approximating to a second-order model - code available in Appendix C.6)**.** *This example is similar to Example 11 that illustrates the non-convexity of the optimal $\mathcal{H}_2$ approximation problem with respect to the poles of the reduced-order model. The LAH model is reduced over $\Omega = [0, \omega] = [0, 20]$ to a second order model whose transfer function is*

$$\hat{H}(s) = \frac{\hat{\phi}}{s - \hat{\lambda}} + \frac{\hat{\phi}^*}{s - \hat{\lambda}^*}$$

where the pole $\hat{\lambda} \in \mathbb{C}$ is varying and the residue $\hat{\phi} \in \mathbb{C}$ is determined in an optimal way (in terms of $\mathcal{H}_{2,\Omega}$-norm) for each value of the pole. Since both models are SISO, one can write the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ as a quadratic function of the residues (see Remark 16).

The optimal residues can then be determined by setting $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\phi}} = 0$ which yields two solutions depending on the location of the reduced-order poles :

- *if the poles are not symmetrically located with respect to the imaginary axis, then the matrix L and the vector h defined by equations (8.3) and (8.4) are given here as,*

$$L = \begin{bmatrix} \frac{\mathbf{a}_{\omega,\hat{\lambda}}}{\hat{\lambda}} & \frac{\mathbf{a}_{\omega,\hat{\lambda}}+\mathbf{a}_{\omega,\hat{\lambda}^*}}{\hat{\lambda}+\hat{\lambda}^*} \\ \frac{\mathbf{a}_{\omega,\hat{\lambda}}+\mathbf{a}_{\omega,\hat{\lambda}^*}}{\hat{\lambda}+\hat{\lambda}^*} & \frac{\mathbf{a}_{\omega,\hat{\lambda}^*}}{\hat{\lambda}^*} \end{bmatrix} \text{ and } h = -\begin{bmatrix} \sum_{i=1}^n \frac{\phi_i}{\lambda_i+\hat{\lambda}}\left(\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}}\right) \\ \sum_{i=1}^n \frac{\phi_i}{\lambda_i+\hat{\lambda}^*}\left(\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}^*}\right) \end{bmatrix}.$$

  The optimal residue and its conjugate are then determined by solving the following linear system

$$L\begin{bmatrix} \hat{\phi} \\ \hat{\phi}^* \end{bmatrix} = h.$$

- *if $\hat{\lambda}$ is on the imaginary axis and $|\hat{\lambda}| > \omega$ (since there are only two complex conjugate poles, this is the only specific case that can occur), then $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is still finite and can be obtained by modifying some entries in the matrix L[1]. By considering the general expression of L*

[1]The process to obtain these modifications is not detailed here, but it is quite straightforward. Indeed, it consists in using the full poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm expressed in Theorem 10, without making additional assumptions on the fact that $\hat{\lambda}_i + \hat{\lambda}_k$ does not vanish, and to express the $\mathcal{H}_{2,\Omega}$-norm of the approximation error. Obviously, since the only term in the approximation error (equation (8.1)) where the denominator $\hat{\lambda}_i + \hat{\lambda}_k$ appears is the $\mathcal{H}_{2,\Omega}$-norm of the reduced-order model, it is the only sum that is affected by this modification.

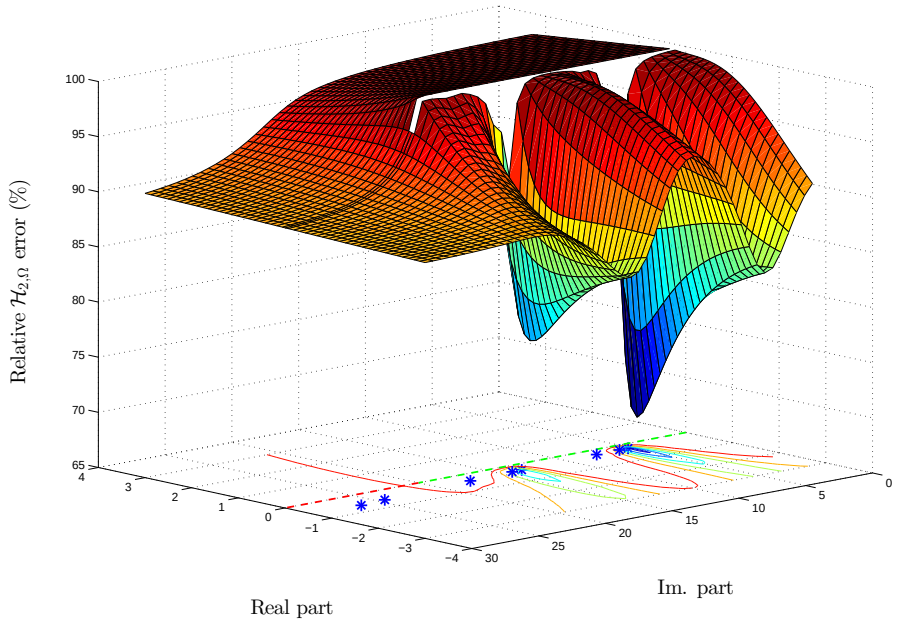


Figure 8.1: $\mathcal{H}_{2,\Omega}$-norm of the approximation error between the LAH model and a second order model with respect to the poles of the reduced-order model for $\Omega = [0, 20]$. The blue stars are the poles of the large-scale model contained in this part of the complex plane, the red/green dashed line represents the frequency bound Ω.

presented in equation (8.3), the indexes i, k for which $\hat{\lambda}_i + \hat{\lambda}_k = 0$ must be modified into

$$[L]_{i,k} = -\frac{\omega}{\pi\left(\omega^2 - \lambda_i \lambda_k\right)}.$$

Here, this means that when $\hat{\lambda}$ is on the imaginary axis over ω, then the expression of h remains the same but L becomes

$$L_{im} = \begin{bmatrix} \frac{\mathbf{a}_{\omega,\hat{\lambda}}}{\hat{\lambda}} & -\frac{2\omega}{\pi\left(\omega^2 - \hat{\lambda}\hat{\lambda}^*\right)} \\ -\frac{2\omega}{\pi\left(\omega^2 - \hat{\lambda}\hat{\lambda}^*\right)} & \frac{\mathbf{a}_{\omega,\hat{\lambda}^*}}{\hat{\lambda}^*} \end{bmatrix},$$

which is simply the limit of L as $\hat{\lambda} + \hat{\lambda}^* \to 0$. In that case, the optimal residues are solution of

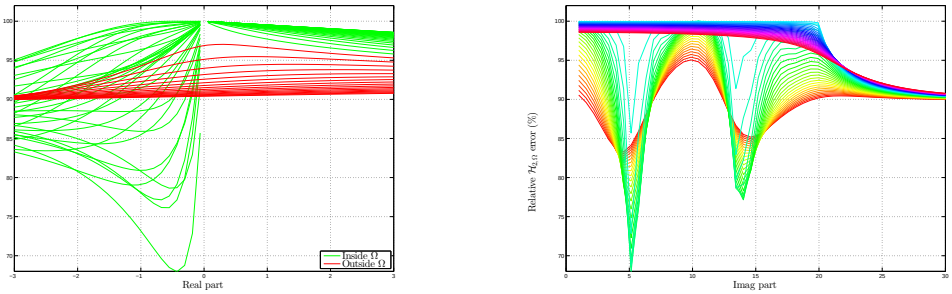
$$L_{im} \begin{bmatrix} \hat{\phi} \\ \hat{\phi}^* \end{bmatrix} = h.$$

- *otherwise, if $\hat{\lambda}$ is on the imaginary axis and $|\hat{\lambda}| \leq \omega$, then the $\mathcal{H}_{2,\Omega}$-norm of the error is, by definition, infinite.*

Again, as in the $\mathcal{H}_2$ case, once the optimal residues have been computed, the $\mathcal{H}_{2,\Omega}$ approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ simplifies (see Section 8.2) and becomes

$$\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \|H\|^2_{\mathcal{H}_{2,\Omega}} - \|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}}.$$

The relative $\mathcal{H}_{2,\Omega}$ error is plotted in Figure 8.1 and projected on the real and imaginary axis in Figure 8.2a and 8.2b, respectively.



(a) (b)

Figure 8.2: Projection of the $\mathcal{H}_{2,\Omega}$-norm presented in Figure 8.1 on the real axis (a) and imaginary axis (b).

In Figure 8.1, one can observe that the behaviour of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ inside the frequency interval Ω is similar to the behaviour of the $\mathcal{H}_2$ approximation error (see Figure 4.1) but outside the interval, the local minima present in the $\mathcal{H}_2$ case have completely vanished.

The behaviour of the objective function in the right-half plane depends on whether $\mathbf{Im}(\hat{\lambda})$ belongs to Ω or not. Indeed, as one can see in Figure 8.2a,

- *when $\mathbf{Im}(\hat{\lambda}) > \omega$ (red line in Figure 8.2a), the objective function increases as $\mathbf{Re}(\hat{\lambda}) < 0$ increases. It means that, with this asymptotically stable model, moving in the positive real direction does not yield a decrease of the objective function and consequently is not likely to represent a descent direction.*

- *If $\mathbf{Im}(\hat{\lambda}) \in \Omega$ (in green), one can see that $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is slowly decreasing as $\mathbf{Re}(\hat{\lambda}) > 0$ increases. Yet, this decrease is generally slower than on the left half-plane[2]. It suggests that when $\mathbf{Im}(\hat{\lambda}) \in \Omega$, switching from the right half-plane to the left half-plane is unlikely to represent a descent direction.*

These two points suggest that reducing an asymptotically stable model with a descent method aimed at minimising $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is very unlikely to produce an unstable reduced-order model. Obviously, this is not necessarily true anymore if the large-scale model is not stable.

In Figure 8.2b, one can see that if $\hat{\lambda}$ is in the right half-plane, then $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ does not evolve along the imaginary axis as long as $\mathbf{Im}(\hat{\lambda}) < \omega$. Then when $\mathbf{Im}(\hat{\lambda}) > \omega$, the objective function decreases slowly and appears to reach a limit. More specifically, let us consider $\hat{\lambda} = x + jy$, $y > 0$, then the limit of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ as y tends towards infinity is equal to the $\mathcal{H}_{2,\Omega}$-norm of the original large-scale model $\mathbf{H}$. Indeed, by noticing that

$$1 + j\frac{\omega}{\hat{\lambda}} = 1 + \frac{\omega y}{x^2 + y^2} + j\frac{\omega x}{x^2 + y^2},$$

it is clear that

$$\lim_{y \to \infty} ln\left(\left|1 + j\frac{\omega}{\hat{\lambda}}\right|\right) = 0,$$

in addition, since $1 + j\frac{\omega}{\hat{\lambda}}$ tends towards 1 as y tends towards infinity,

$$\lim_{y \to \infty} \mathbf{arg}\left(1 + j\frac{\omega}{\hat{\lambda}}\right) = 0$$

[2]This assertion does not hold around the local maxima of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$.

The same limits hold with $1 - j\frac{\omega}{\hat{\lambda}}$*, thus*

$$\lim_{y\to\infty} \mathbf{a}_{\omega,\hat{\lambda}} = 0.$$

This means that $\lim_{y\to\infty} L = 0$ *(and similarly, one can prove that* $\lim_{y\to\infty} L_{im} = 0$*). Obviously* $\lim_{y\to\infty} H(-\hat{\lambda}) = 0$ *thus,* $\lim_{y\to\infty} h = 0$ *which implies that*

$$\lim_{y\to\infty} \mathcal{J}_{\mathcal{H}_{2,\Omega}} = \|H\|^2_{\mathcal{H}_{2,\Omega}}.$$

This result is quite natural since if the reduced-order model poles move away from the frequency-band of interest, its influence in this frequency interval vanishes and only the influence of the large-scale model remains.

The fact that $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ *reaches a plateau as the imaginary parts of the reduced-order model poles increase lead to believe that if a descent optimisation method is used, one should probably not initialise it with a reduced-order model which poles are located outside of* Ω*.*

8.1.2 On the computation of the approximation error

In optimisation, the evaluation of the objective function can be a numerical burden if it is not fast enough. The first step towards a fast computation of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is to express it with matrix and vector operations to exploit the standard and highly optimised linear algebra libraries used in software such as Matlab®.

To this aim, let us define $W_{nn} \in \mathbb{C}^{n\times n}$, $W_{nr} \in \mathbb{C}^{n\times r}$ and $W_{rr} \in \mathbb{C}^{r\times r}$ as

$$[W_{rr}]_{i,k} = \frac{1}{2}\frac{\mathbf{a}_{\omega,\hat{\lambda}_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}}{\hat{\lambda}_i + \hat{\lambda}_k}, \tag{8.7}$$

$$[W_{nr}]_{i,k} = \frac{\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}}{\hat{\lambda}_i + \hat{\lambda}_k}, \tag{8.8}$$

$$[W_{nn}]_{i,k} = \frac{\mathbf{a}_{\omega,\lambda_i}}{\lambda_i + \lambda_k}.$$

Each of these matrix can be easily built as follows

$$W_{nr} = \left(\begin{bmatrix} \mathbf{a}_{\omega,\lambda_1} \\ \vdots \\ \mathbf{a}_{\omega,\lambda_n} \end{bmatrix} \mathbb{1}_n^T + \mathbb{1}_r \begin{bmatrix} \mathbf{a}_{\omega,\hat{\lambda}_1} & \dots, \mathbf{a}_{\omega,\hat{\lambda}_r} \end{bmatrix} \right) \odot \begin{bmatrix} \frac{1}{\lambda_1+\hat{\lambda}_1} & \cdots & \frac{1}{\lambda_1+\hat{\lambda}_r} \\ \vdots & \ddots & \vdots \\ \frac{1}{\lambda_n+\hat{\lambda}_1} & \cdots & \frac{1}{\lambda_n+\hat{\lambda}_r} \end{bmatrix},$$

where $\odot$ is the Hadamard product and $\mathbb{1}_n \in \mathbb{R}^{n\times 1}$ a column vector full of 1. Let $C_\Delta \in \mathbb{C}^{n_y\times n}$ and $\hat{C}_\Delta \in \mathbb{C}^{n_y\times r}$ denote the matrices formed by horizontally concatenating $\mathbf{c_i}^T$ and $\hat{\mathbf{c}}_\mathbf{k}^T$, respectively, *i.e.*

$$C_\Delta = \begin{bmatrix} \mathbf{c_1}^T & \dots \mathbf{c_n}^T \end{bmatrix} \text{ and } \hat{C}_\Delta = \begin{bmatrix} \hat{\mathbf{c}}_\mathbf{1}^T & \dots & \hat{\mathbf{c}}_\mathbf{r}^T \end{bmatrix}, \tag{8.9}$$

and let $B_\Delta \in \mathbb{C}^{n\times n_u}$ and $\hat{B}_\Delta \in \mathbb{C}^{r\times n_u}$ be the matrices built by vertically concatenating $\mathbf{b_i}$ and $\hat{\mathbf{b}}_\mathbf{k}$, *i.e.*

$$B_\Delta = \begin{bmatrix} \mathbf{b_1} \\ \vdots \\ \mathbf{b_n} \end{bmatrix} \text{ and } \hat{B}_\Delta = \begin{bmatrix} \hat{\mathbf{b}}_\mathbf{1} \\ \vdots \\ \hat{\mathbf{b}}_\mathbf{r} \end{bmatrix}. \tag{8.10}$$

Obviously, $C_\Delta = CX$, $\hat{C}_\Delta = \hat{C}\hat{X}$, $B_\Delta = X^{-1}B$ and $\hat{B}_\Delta = \hat{X}^{-1}\hat{B}$, where $X \in \mathbb{C}^{n\times n}$ and $\hat{X} \in \mathbb{C}^{r\times r}$ are the matrices which columns are the right eigenvectors of A and $\hat{A}$, respectively.

Then, as it has been shown in Section 7.1.2, the $\mathcal{H}_{2,\Omega}$-norms of $\mathbf{H}$ and $\hat{\mathbf{H}}$, direct feedthrough excluded, can be computed as

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \mathbb{1}_n^T \left((C_\Delta^T C_\Delta) \odot (B_\Delta B_\Delta^T) \odot W_{nn} \right) \mathbb{1}_n,$$

and

$$\|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}} = \mathbb{1}_r^T \left((\hat{C}_\Delta^T \hat{C}_\Delta) \odot (\hat{B}_\Delta \hat{B}_\Delta^T) \odot W_{rr} \right) \mathbb{1}_r.$$

Note that $[W_{rr}]_{i,k}$ could have been directly defined as $\frac{\mathbf{a}_{\omega,\hat{\lambda}_i}}{\hat{\lambda}_i + \hat{\lambda}_k}$ (like in the poles-residues formulation of the norm in Theorem 10), but here the fact that

$$\|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}} = \sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\hat{\Phi}_i \hat{\Phi}_k\right)}{\hat{\lambda}_i + \hat{\lambda}_k} \mathbf{a}_{\omega,\hat{\lambda}_i} = \sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\hat{\Phi}_i \hat{\Phi}_k\right)}{\hat{\lambda}_i + \hat{\lambda}_k} \frac{1}{2} \left(\mathbf{a}_{\omega,\hat{\lambda}_i} + \mathbf{a}_{\omega,\hat{\lambda}_k} \right),$$

has been exploited. This has no interest for the computation of $\|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}}$ presented above but it is interesting for the the computation of the gradient in Section 8.2.2.

Similarly, the cross term in the error can be computed as

$$-\sum_{i=1}^{n}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\Phi_i \hat{\Phi}_k^T\right)}{\lambda_i + \hat{\lambda}_k} \left(\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_k} \right) = -\mathbb{1}_n^T \left((C_\Delta^T \hat{C}_\Delta) \odot (B_\Delta \hat{B}_\Delta^T) \odot W_{nr} \right) \mathbb{1}_r.$$

and by denoting $\mathbf{diag}(M)$ the column vector formed by the diagonal of the matrix M, the terms in $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ (8.1) associated with the direct feedthrough D and $\hat{D}$ can be expressed as

$$-\sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i \tilde{D}^T\right) \mathbf{a}_{\omega,\lambda_i} = -\mathbf{diag}\left(B_\Delta \tilde{D}^T C_\Delta\right)^T \begin{bmatrix} \mathbf{a}_{\omega,\lambda_1} \\ \vdots \\ \mathbf{a}_{\omega,\lambda_n} \end{bmatrix},$$

and

$$\sum_{i=1}^{r} \mathbf{tr}\left(\hat{\Phi}_i \tilde{D}^T\right) \mathbf{a}_{\omega,\hat{\lambda}_i} = \mathbf{diag}\left(\hat{B}_\Delta \tilde{D}^T \hat{C}_\Delta\right)^T \begin{bmatrix} \mathbf{a}_{\omega,\hat{\lambda}_1} \\ \vdots \\ \mathbf{a}_{\omega,\hat{\lambda}_r} \end{bmatrix}.$$

Note that the number of operations required to compute $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ can be decreased by exploiting the structure of the matrices such as the symmetry of W_{rr} or W_{nn} or the fact that the complex elements come with their complex conjugate, but the benefit would not necessarily be significant in term of computation time, that is why it has not been considered here.

The second step towards a fast computation of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ consists in avoiding to compute the same things multiple times. Here, the matrices C_Δ, B_Δ, $\|H\|^2_{\mathcal{H}_{2,\Omega}}$, the terms associated with D and $\mathbf{a}_{\omega,\lambda_i}$ $(i = 1, \ldots, n)$ can be stored throughout the whole optimisation process. Besides, in the context of minimisation of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$, it is worth noticing that computing $\|H\|_{\mathcal{H}_{2,\Omega}}$ is not mandatory since minimising $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is equivalent to minimise $\mathcal{J}_{\mathcal{H}_{2,\Omega}} - \|H\|^2_{\mathcal{H}_{2,\Omega}}$. Yet, since the computation of the cross terms require the poles and residues of the large-scale model, computing $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ does not implies a lot of additional operations especially as it only needs to be done once.

Note that with respect to the computation time of the approximation error, the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$ approximation error offers some advantage compared to the gramian based one (see Section 5.2.2). Indeed, with the gramian formulation, each evaluation of the approximation error requires to solve a $n \times r$ Sylvester equation, a $r \times r$ Lyapunov equation, due to the cross terms and the $\mathcal{H}_{2,\Omega}$-norm of the reduced-order model respectively, while here, the error is explicit in terms of the poles and residues of the reduced-order model. This advantage comes from the additional hypothesis on the structure of the poles which must be semi-simple and which is not made with the gramian-based formulation.

8.2 Gradient of the error

Based on the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$ approximation error between a large-scale model $\mathbf{H}$ and a reduced-order one $\hat{\mathbf{H}}$, the problem of optimal $\mathcal{H}_{2,\Omega}$ approximation can be re-formulated as follows.

Problem 4 (Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$ optimal model approximation problem)**.** *Given an asymptotically stable LTI dynamical model* $\mathbf{H}$ *of order* n*. The problem of optimal* $\mathcal{H}_{2,\Omega}$ *model approximation consists in finding the poles* $\hat{\lambda}_i \in \mathbb{C}$*, the vectors* $\hat{\mathbf{c}}_\mathbf{i} \in \mathbb{C}^{1\times n_y}$ *and* $\hat{\mathbf{b}}_\mathbf{i} \in \mathbb{C}^{1\times n_u}$ *(*$i = 1, \dots, r$*) and the direct feedthrough* $\hat{D} \in \mathbb{R}^{n_y \times n_u}$ *of the asymptotically stable reduced-order model* $\hat{\mathbf{H}}$ *that minimises the* $\mathcal{H}_{2,\Omega}$*-norm of the approximation error* $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$*.*

Since only stable models have been considered in this study, Problem 4 is restricted to asymptotically stable models, but this assumption is more restrictive than what is required and used here to express the first-order optimality conditions. Indeed, the hypothesis used here is that the poles of both models are not symmetrically located with respect to the imaginary axis. One could diminish these hypothesis even further by using the full expression of the $\mathcal{H}_{2,\Omega}$-norm. However, this has not been considered here.

$\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is a function of $r+rn_y+rn_u+n_yn_u$ variables, thus, Problem 4 is over-parametrised since the minimal number of parameters to represent a LTI model of order r is $rn_y+rn_u+n_yn_u$. This issue could be fixed by imposing some norm constraint on either the $\hat{\mathbf{b}}_\mathbf{i}$ or the $\hat{\mathbf{c}}_\mathbf{i}$ $(i = 1, \dots, r)$, but in order to stick with the simpler unconstrained optimisation framework, no constraint is added.

$\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is a real valued function of complex variables $\hat{\lambda}_i$, $\hat{\mathbf{b}}_\mathbf{i}$, $\hat{\mathbf{c}}_\mathbf{i}$ and their complex conjugates. As it does not satisfy the Cauchy-Riemann equations, such a function is not analytic. The common way of dealing with such kind of functions consists in writing them as functions of the real and imaginary parts of the variables. However here, it would break the structure of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ and that is why keeping it as a function of complex variables is preferred.

In that context, the *Wirtinger calculus* is used to derive the first order optimality conditions for the optimal $\mathcal{H}_{2,\Omega}$ approximation problem. A quick reminder of its main properties is done in Section 8.2.1. It is then applied to differentiate $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ in Section 8.2.2.

8.2.1 Reminder on Wirtinger Calculus

This Section is not meant to be an exhaustive description of the Wirtinger calculus but rather a quick reminder of the properties that are used in the next Section to express the first-order optimality conditions of the approximation problem. It is mainly based on [Brandwood, 1983; Remmert, 1991] where the complex gradient and conjugate gradient operators are defined. Additional useful elements about the use of Wirtinger calculus for optimisation can be found in [Kreutz-Delgado, 2009] and in [Sorber et al., 2012].

Partial complex-variable derivative

Let us consider $g : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ a function of z and its conjugate z^* which is analytic with respect to z and z^* independently, *i.e.* $g(z, a)$ and $g(a, z)$ are analytic in z. Let us also consider $h : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$, the function of real variables x and y such that $z = x + jy$ and $g(z, z^*) = h(x, y)$, then

$$\begin{aligned} \frac{\partial g}{\partial z} &:= \frac{1}{2}\left(\frac{\partial h}{\partial x} - j\frac{\partial h}{\partial y}\right), \\ \frac{\partial g}{\partial z^*} &:= \frac{1}{2}\left(\frac{\partial h}{\partial x} + j\frac{\partial h}{\partial y}\right), \end{aligned} \tag{8.11}$$

where $\frac{\partial g}{\partial z}$ (respectively $\frac{\partial g}{\partial z^*}$) is obtained by treating z^* (respectively z) as a constant[3]. Equation (8.11) can easily be verified with simple functions, for instance,

$$\frac{\partial z}{\partial z} = \frac{\partial z^*}{\partial z^*} = 1 \quad \text{and} \quad \frac{\partial z}{\partial z^*} = \frac{\partial z^*}{\partial z} = 0.$$

Let $g(z, z^*) = u(x, y) + jv(x, y)$ where u and v are real functions, the Cauchy-Riemann conditions can be retrieved, indeed, $\frac{\partial g}{\partial z^*} = 0$ can be rewritten with equation (8.11) as $\frac{\partial u}{\partial x} + j\frac{\partial v}{\partial x} + j\frac{\partial u}{\partial y} - \frac{\partial v}{\partial y} = 0$, which yields, by identifying the real and imaginary parts, to

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.$$

This means that a complex function $f : \mathbb{C} \to \mathbb{C}$ is analytic in z if and only if it only depends exclusively of z and not of z^*.

Stationary points of non-analytic functions

Let us now consider a real-valued function $f : \mathbb{C} \to \mathbb{R}$ of a complex variable z and its conjugate. As mentioned above, such a function is not analytic in z. To characterise its stationary points, one must consider the function $g : \mathbb{C} \times \mathbb{C} \to \mathbb{R}$ of two complex variables and analytic with respect to each variable such that $f(z) = g(z, z^*)$. Then the stationary points of f are the points for which either $\frac{\partial g}{\partial z} = 0$ or $\frac{\partial g}{\partial z^*} = 0$.

This can be proved by setting $g(z, z^*) = u(x, y) + jv(x, y)$ where u and v are real functions of the real variables x and y. Since g is real valued, $v(x, y) = 0$ for all $x, y \in \mathbb{R}$, thus from equation (8.11),

$$\frac{\partial g}{\partial z} = \frac{1}{2}\left(\frac{\partial u}{\partial x} - j\frac{\partial u}{\partial y}\right) \quad \text{and} \quad \frac{\partial g}{\partial z^*} = \frac{1}{2}\left(\frac{\partial u}{\partial x} + j\frac{\partial u}{\partial y}\right).$$

By identifying the real and imaginary parts, it is clear that

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} = 0 \iff \frac{\partial g}{\partial z} = 0 \quad \text{and} \quad \frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} = 0 \iff \frac{\partial g}{\partial z^*} = 0.$$

Note that the differential of g is given as

$$dg = \frac{\partial g}{\partial z} dz + \frac{\partial g}{\partial z^*} dz^*,$$

and since g is real valued, $\frac{\partial g}{\partial z^*} = \left(\frac{\partial g}{\partial z}\right)^*$, thus

$$dg = 2\mathbf{Re}\left(\frac{\partial g}{\partial z} dz\right).$$

The differential dg is maximal when $\frac{\partial g}{\partial z} dz$ is real, thus dz must be proportional to $\left(\frac{\partial g}{\partial z}\right)^* = \frac{\partial g}{\partial z^*}$. It means that $\frac{\partial g}{\partial z^*}$ is the steepest ascent direction. Hence, if one looks for a maximum of f, one could set the following iteration scheme

$$z_{k+1} = z_k + \alpha \frac{\partial g}{\partial z^*}\Big|_{z=z_k},$$

where α is the step length. Example 18 illustrates the various elements of the Wirtinger calculus presented here on a simple function.

[3] In [Kreutz-Delgado, 2009], such derivatives are said to be formal because one cannot vary z while keeping z^* constant.

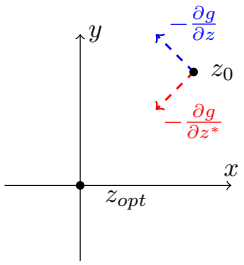


Figure 8.3: Illustration of the descent directions for $g(z, z^*) = zz^*$.

Example 18 (Illustration of the Wirtinger calculus on a simple non-analytic function)**.** *Let us consider the real-valued functions* $f(z) = |z|^2$, $g(z, z^*) = zz^*$ *and* $u(x, y) = x^2 + y^2$ *with* $z = x + jy$. *The function* f *is non-analytic in* z *and clearly*

$$f(z) = g(z, z^*) = u(x, y).$$

The stationary point of f *can be characterised through the gradient of* u,

$$\nabla u = 2 \begin{bmatrix} x \\ y \end{bmatrix} = 0 \iff x = y = 0,$$

or equivalently through the partial complex derivatives

$$\frac{\partial g}{\partial z} = z^* \quad \textit{and} \quad \frac{\partial g}{\partial z^*} = z,$$

which also yield $z = 0$.

Let us now consider an initial point $z_0 = x_0 + jy_0$ *with* $x_0 > 0$ *and* $y_0 > 0$, *then, the steepest descent direction is given by*

$$-\nabla u = -2 \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \quad \textit{or} \quad -\begin{bmatrix} \frac{\partial g}{\partial z^*} \\ \frac{\partial g}{\partial z} \end{bmatrix} = -\begin{bmatrix} z_0 \\ z_0^* \end{bmatrix},$$

but not by $-\begin{bmatrix} \frac{\partial g}{\partial z} \\ \frac{\partial g}{\partial z^*} \end{bmatrix} = -\begin{bmatrix} z_0^* \\ z_0 \end{bmatrix}$ *as illustrated in Figure 8.3 where the directions* $-\frac{\partial g}{\partial z}$ *and* $-\frac{\partial g}{\partial z^*}$ *are represented. Note also the factor* 2 *that is not present in the complex case.*

Complex gradient

The elements mentioned above can be generalised to the multivariate case through the definition of the complex gradient operator with respect to the vector $\mathbf{z} = \begin{bmatrix} z_1 & \dots & z_N \end{bmatrix}^T \in \mathbb{C}^{N \times 1}$,

$$\nabla_{\mathbf{z}} = \begin{bmatrix} \frac{\partial}{\partial z_1} \\ \vdots \\ \frac{\partial}{\partial z_N} \end{bmatrix}.$$

Given a scalar real-valued function $f : \mathbb{C}^N \to \mathbb{R}$ of a complex vector $\mathbf{z}$ (and its conjugate) and $g(\mathbf{z}, \mathbf{z}^*) : \mathbb{C}^N \times \mathbb{C}^N \to \mathbb{R}$, an analytic function with respect to each z_i and z_i^* $(i = 1, \dots, N)$ such that $f(\mathbf{z}) = g(\mathbf{z}, \mathbf{z}^*)$, then, the following results hold,

- either $\nabla_{\mathbf{z}} g = 0$ or $\nabla_{\mathbf{z}^*} g = 0$ enables to characterise the stationary points of f,

- the differential of g is given by

$$\begin{aligned} dg &= (\nabla_{\mathbf{z}} g)^T \, d\mathbf{z} + (\nabla_{\mathbf{z}^*} g)^T \, d\mathbf{z}^* \\ &= 2\mathbf{Re}\left((\nabla_{\mathbf{z}^*} g)^T \, d\mathbf{z} \right) \quad (g \text{ is real-valued}), \end{aligned} \tag{8.12}$$

- by using the Cauchy-Schwarz inequality in equation (8.12), one can see that the differential $\mathbf{Re}((\nabla_{\mathbf{z}^*} g)^T \, d\mathbf{z})$ is maximal when $d\mathbf{z}$ is co-linear to $\nabla_{\mathbf{z}^*} g$. Hence the steepest ascent direction is given by $\nabla_{\mathbf{z}^*} g$.

Note that complex Hessian matrices can similarly be defined (see [Van Den Bos, 1994; Kreutz-Delgado, 2009; Sorber et al., 2012]) but since the Hessian of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is not exploited in the sequel, they are not described here.

8.2.2 Gradient of the $\mathcal{H}_{2,\Omega}$ approximation error

To highlight clearly the link with the previous section, let us focus on the reduced-order poles and consider that the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is a function of the reduced-order model poles $\hat{\lambda}_i$ $(i = 1, \ldots, r)$ only. Let us also assume that these poles are all complex and ordered as

$$\hat{\lambda}_1,\ \hat{\lambda}_2 = \hat{\lambda}_1^*,\ \hat{\lambda}_3,\ \hat{\lambda}_4 = \hat{\lambda}_3^*, \ldots, \hat{\lambda}_{r-1},\ \hat{\lambda}_r = \hat{\lambda}_{r-1}^*.$$

In that case, the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ can either be seen as a function of $\frac{r}{2}$ variables $\hat{\lambda}_1, \hat{\lambda}_3, \ldots, \hat{\lambda}_{r-1}$ or as a function of r variables $\hat{\lambda}_1, \hat{\lambda}_1^*, \ldots, \hat{\lambda}_{r-1}, \hat{\lambda}_{r-1}^*$. In the former case, $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is similar to the function $f(z)$ of Section 8.2.1 while in the latter case, $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is like $g(z, z^*)$. To differentiate $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ with respect to the poles, one must consider it as a function of $\mathbf{w} = \begin{bmatrix} \hat{\lambda}_1 & \hat{\lambda}_3 & \ldots & \hat{\lambda}_{r-1} \end{bmatrix}^T$ and of $\mathbf{w}^*$. The first-order optimality conditions of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ with respect to the reduced-order model poles are then either given by the $\frac{r}{2}$ conditions $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \mathbf{w}} = 0$ or by their conjugate $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \mathbf{w}^*} = 0$.

In order to avoid differentiating between the complex and complex conjugate variables, one can also directly consider $\mathbf{z} = \begin{bmatrix} \mathbf{w}^T & \mathbf{w}^H \end{bmatrix}^T \in \mathbb{C}^{r \times 1}$ and $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \mathbf{z}} = \begin{bmatrix} \frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \mathbf{w}}^T & \frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \mathbf{w}^*}^T \end{bmatrix} = 0$ which leads to the same optimality conditions, excepted that they are redundant.

This is what is done in Theorem 16 where $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is differentiated with respect to $\hat{\lambda}_i$, $\hat{\mathbf{b}}_\mathbf{i}$, $\hat{\mathbf{c}}_\mathbf{i}$ $(i = 1, \ldots, r)$ and $\hat{D}$ without distinction between the variables, conjugated variables and real variables. Refer to Remark 17 for some comments on the notations used in Theorem 16.

The Wirtinger calculus has already been implicitly used in Section 4.1 to express the first-order optimality conditions of the optimal $\mathcal{H}_2$ approximation problem. It has not been mentioned because it does not modify the differentiation rules and it is not mentioned in the literature.

Theorem 16 (Poles-residues formulation of the necessary conditions for the optimal $\mathcal{H}_{2,\Omega}$ approximation problem)**.** *Let us consider two LTI dynamical models* $\mathbf{H}$ *and* $\hat{\mathbf{H}}$ *of order* n *and* $r \ll n$*, respectively, which poles are semi-simple and not symmetrically located with respect to the imaginary axis and which associated transfer functions are given by*

$$H(s) = \sum_{i=1}^{n} \frac{\mathbf{c_i}^T \mathbf{b_i}}{s - \lambda_i} + D \quad \text{and} \quad \hat{H}(s) = \sum_{i=1}^{n} \frac{\hat{\mathbf{c}}_\mathbf{i}^T \hat{\mathbf{b}}_\mathbf{i}}{s - \hat{\lambda}_i} + \hat{D}.$$

Let us also consider a frequency interval $\Omega = [0, \omega]$*. If* $\hat{\mathbf{H}}$ *minimises* $\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \|H - \hat{H}\|_{\mathcal{H}_{2,\Omega}}^2$*, then the derivatives of* $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ *with respect to* $\hat{\lambda}_m$*,* $\hat{\mathbf{b}}_\mathbf{m}$*,* $\hat{\mathbf{c}}_\mathbf{m}$ *and* $\hat{D}$*, given for* $m = 1, \ldots, r$ *by equations*

(8.13), (8.14), (8.15) *and* (8.16)*, respectively, are equal to* 0*.*

$$\begin{aligned}\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\lambda}_m} &= -\sum_{i=1}^{r} \hat{\mathbf{c}}_\mathbf{m}\hat{\mathbf{c}}_\mathbf{i}^T\hat{\mathbf{b}}_\mathbf{i}\hat{\mathbf{b}}_\mathbf{m}^T\left(\frac{\mathbf{a}_{\omega,\hat{\lambda}_i}+\mathbf{a}_{\omega,\hat{\lambda}_m}}{\left(\hat{\lambda}_i+\hat{\lambda}_m\right)^2}+\frac{2}{\pi}\frac{\omega}{\left(\hat{\lambda}_m^2+\omega^2\right)\left(\hat{\lambda}_i+\hat{\lambda}_m\right)}\right)\dots \\ &\quad+\sum_{i=1}^{n} \hat{\mathbf{c}}_\mathbf{m}\mathbf{c}_\mathbf{i}{}^T\mathbf{b}_\mathbf{i}\hat{\mathbf{b}}_\mathbf{m}^T\left(\frac{\mathbf{a}_{\omega,\lambda_i}+\mathbf{a}_{\omega,\hat{\lambda}_m}}{\left(\lambda_i+\hat{\lambda}_m\right)^2}+\frac{2}{\pi}\frac{\omega}{\left(\hat{\lambda}_m^2+\omega^2\right)\left(\lambda_i+\hat{\lambda}_m\right)}\right)\dots \\ &\quad+\frac{2}{\pi}\hat{\mathbf{c}}_\mathbf{m}\tilde{D}\hat{\mathbf{b}}_\mathbf{m}\frac{\omega}{\omega^2+\hat{\lambda}_m^2},\end{aligned} \tag{8.13}$$

$$\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{m}} = \sum_{i=1}^{r}\frac{\hat{\mathbf{b}}_\mathbf{i}^T\hat{\mathbf{c}}_\mathbf{i}\hat{\mathbf{c}}_\mathbf{m}^T}{\hat{\lambda}_i+\hat{\lambda}_m}\left(\mathbf{a}_{\omega,\hat{\lambda}_i}+\mathbf{a}_{\omega,\hat{\lambda}_m}\right)-\sum_{i=1}^{n}\frac{\mathbf{b}_\mathbf{i}{}^T\mathbf{c}_\mathbf{i}\hat{\mathbf{c}}_\mathbf{m}^T}{\lambda_i+\hat{\lambda}_m}\left(\mathbf{a}_{\omega,\lambda_i}+\mathbf{a}_{\omega,\hat{\lambda}_m}\right)+\tilde{D}^T\hat{\mathbf{c}}_\mathbf{m}^T\mathbf{a}_{\omega,\hat{\lambda}_m}, \tag{8.14}$$

$$\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{c}}_\mathbf{m}} = \sum_{i=1}^{r}\frac{\hat{\mathbf{c}}_\mathbf{i}^T\hat{\mathbf{b}}_\mathbf{i}\hat{\mathbf{b}}_\mathbf{m}^T}{\hat{\lambda}_i+\hat{\lambda}_m}\left(\mathbf{a}_{\omega,\hat{\lambda}_i}+\mathbf{a}_{\omega,\hat{\lambda}_m}\right)-\sum_{i=1}^{n}\frac{\mathbf{c}_\mathbf{i}{}^T\mathbf{b}_\mathbf{i}\hat{\mathbf{b}}_\mathbf{m}^T}{\lambda_i+\lambda_m}\left(\mathbf{a}_{\omega,\lambda_i}+\mathbf{a}_{\omega,\hat{\lambda}_m}\right)+\tilde{D}\hat{\mathbf{b}}_\mathbf{m}^T\mathbf{a}_{\omega,\hat{\lambda}_m}, \tag{8.15}$$

$$\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{D}} = -\sum_{i=1}^{r}\hat{\mathbf{c}}_\mathbf{i}^T\hat{\mathbf{b}}_\mathbf{i}\mathbf{a}_{\omega,\hat{\lambda}_i}+\sum_{i=1}^{n}\mathbf{c}_\mathbf{i}{}^T\mathbf{b}_\mathbf{i}\mathbf{a}_{\omega,\lambda_i}-\frac{2}{\pi}\omega\tilde{D}. \tag{8.16}$$

Proof. The derivative are obtained by straight differentiation of the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ given in equation (8.1). □

Remark 17 (About the notations in Theorem 16)**.** *When differentiating with respect to the line vectors* $\hat{\mathbf{b}}_\mathbf{m}$ *or* $\hat{\mathbf{c}}_\mathbf{m}$*, the resulting gradients are column vectors denoted by* $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{c}}_\mathbf{m}} \in \mathbb{C}^{n_y\times 1}$ *and* $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{m}} \in \mathbb{C}^{n_u\times 1}$*. Besides, the derivative with respect to* $\hat{D}$ *is performed element-wise and the resulting derivative is a matrix* $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{D}} \in \mathbb{R}^{n_y\times n_u}$*. This inconsistency has no real impact on the sequel and allows for more clarity, in particular, the derivatives are straightforwardly obtained.*

Reformulation of the first-order optimality conditions as interpolation conditions

The first-order optimality conditions of the $\mathcal{H}_{2,\Omega}$ approximation problem can be reformulated as interpolation conditions similar, in some way, to the interpolation conditions arising in the optimal $\mathcal{H}_2$ model approximation problem (see Section 4.1). Let us define the *filtered* transfer functions $H_\omega(s)$ and $\hat{H}_\omega(s)$ as the transfer functions $H(s)$ and $\hat{H}(s)$ considered without the direct feedthrough D and $\hat{D}$ and which i-th modal contributions are weighted by $\mathbf{a}_{\omega,\lambda_i}$ and $\mathbf{a}_{\omega,\hat{\lambda}_i}$, respectively, *i.e.*

$$H_\omega(s)=\sum_{i=1}^{n}\frac{\mathbf{c}_\mathbf{i}{}^T\mathbf{b}_\mathbf{i}}{s-\lambda_i}\mathbf{a}_{\omega,\lambda_i} \quad \text{and} \quad \hat{H}_\omega(s)=\sum_{i=1}^{r}\frac{\hat{\mathbf{c}}_\mathbf{i}^T\hat{\mathbf{b}}_\mathbf{i}}{s-\hat{\lambda}_i}\mathbf{a}_{\omega,\hat{\lambda}_i}.$$

Derivative with respect to $\hat{\mathbf{b}}_\mathbf{m}$. Let us consider the derivative of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ with respect to $\hat{\mathbf{b}}_\mathbf{m}$ of equation (8.14), the first sum can be written as

$$\begin{aligned}\sum_{i=1}^{r}\frac{\hat{\mathbf{b}}_\mathbf{i}^T\hat{\mathbf{c}}_\mathbf{i}\hat{\mathbf{c}}_\mathbf{m}^T}{\hat{\lambda}_i+\hat{\lambda}_m}\left(\mathbf{a}_{\omega,\hat{\lambda}_i}+\mathbf{a}_{\omega,\hat{\lambda}_m}\right)-\hat{D}^T\hat{\mathbf{c}}_\mathbf{m}^T\mathbf{a}_{\omega,\hat{\lambda}_m} &= \left(\sum_{i=1}^{r}\frac{\hat{\mathbf{b}}_\mathbf{i}^T\hat{\mathbf{c}}_\mathbf{i}}{\hat{\lambda}_i+\hat{\lambda}_m}\mathbf{a}_{\omega,\hat{\lambda}_i}\right. \\ &\quad\left.\dots+\left(\sum_{i=1}^{n}\frac{\hat{\mathbf{b}}_\mathbf{i}^T\hat{\mathbf{c}}_\mathbf{i}}{\hat{\lambda}_i+\hat{\lambda}_m}-\hat{D}^T\right)\mathbf{a}_{\omega,\hat{\lambda}_m}\right)\hat{\mathbf{c}}_\mathbf{m}^T \\ &= \left(-\hat{H}_\omega^T(-\hat{\lambda}_m)-\hat{H}(-\hat{\lambda}_m)^T\mathbf{a}_{\omega,\hat{\lambda}_m}\right)\hat{\mathbf{c}}_\mathbf{m}^T,\end{aligned}$$

and similarly, the second one can be written as

$$-\sum_{i=1}^{n} \frac{\mathbf{b_i}^T \mathbf{c_i} \hat{\mathbf{c}}_\mathbf{m}^T}{\lambda_i + \hat{\lambda}_m} \left(\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_m} \right) + D\hat{\mathbf{c}}_\mathbf{m}^T \mathbf{a}_{\omega,\hat{\lambda}_m} = \left(H_\omega(-\hat{\lambda}_m)^T + H(-\hat{\lambda}_m)^T \mathbf{a}_{\omega,\hat{\lambda}_m} \right) \hat{\mathbf{c}}_\mathbf{m}^T.$$

Hence $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{m}} = 0$ is equivalent to

$$\hat{\mathbf{c}}_\mathbf{m} \left(\hat{H}_\omega(-\hat{\lambda}_m) + \hat{H}(-\hat{\lambda}_m) \mathbf{a}_{\omega,\hat{\lambda}_m} \right) = \hat{\mathbf{c}}_\mathbf{m} \left(H_\omega(-\hat{\lambda}_m) + H(-\hat{\lambda}_m) \mathbf{a}_{\omega,\hat{\lambda}_m} \right). \tag{8.17}$$

Derivative with respect to $\hat{\mathbf{c}}_\mathbf{m}$. Likewise, $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{c}}_\mathbf{m}} = 0$ can also be written as

$$\left(\hat{H}_\omega(-\hat{\lambda}_m) + \hat{H}(-\hat{\lambda}_m) \mathbf{a}_{\omega,\hat{\lambda}_m} \right) \hat{\mathbf{b}}_\mathbf{m}^T = \left(H_\omega(-\hat{\lambda}_m) + H(-\hat{\lambda}_m) \mathbf{a}_{\omega,\hat{\lambda}_m} \right) \hat{\mathbf{b}}_\mathbf{m}^T. \tag{8.18}$$

Derivative with respect to $\hat{\lambda}_m$. This reformulation process is not as obvious concerning the derivative of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ with respect to $\hat{\lambda}_m$. Let us consider the first sum in equation (8.13), note that

$$\sum_{i=1}^{r} \hat{\mathbf{c}}_\mathbf{m} \frac{\hat{\mathbf{c}}_\mathbf{i}^T \hat{\mathbf{b}}_\mathbf{i}}{\left(\hat{\lambda}_i + \hat{\lambda}_m \right)^2} \mathbf{a}_{\omega,\hat{\lambda}_i} \hat{\mathbf{b}}_\mathbf{m}^T = \hat{\mathbf{c}}_\mathbf{m} \hat{H}'_\omega(-\hat{\lambda}_m) \hat{\mathbf{b}}_\mathbf{m}^T,$$

and that

$$\sum_{i=1}^{r} \hat{\mathbf{c}}_\mathbf{m} \frac{\hat{\mathbf{c}}_\mathbf{i}^T \hat{\mathbf{b}}_\mathbf{i}}{\left(\hat{\lambda}_i + \hat{\lambda}_m \right)^2} \mathbf{a}_{\omega,\hat{\lambda}_m} \hat{\mathbf{b}}_\mathbf{m}^T = \hat{\mathbf{c}}_\mathbf{m} \hat{H}'(-\hat{\lambda}_m) \mathbf{a}_{\omega,\hat{\lambda}_m} \hat{\mathbf{b}}_\mathbf{m}^T. \tag{8.19}$$

The last element of the sum is in fact the derivative of $\mathbf{a}_{\omega,-s}$ with respect to s, indeed by noticing that $\frac{d}{ds}\mathbf{atan}\left(-\frac{\omega}{s}\right) = \frac{\omega}{\omega^2+s^2}$, it comes

$$\sum_{i=1}^{r} \hat{\mathbf{c}}_\mathbf{m} \left(\frac{\hat{\mathbf{c}}_\mathbf{i}^T \hat{\mathbf{b}}_\mathbf{i}}{\hat{\lambda}_i + \hat{\lambda}_m} - \hat{D} \right) \frac{2}{\pi} \frac{\omega}{\omega^2 + \hat{\lambda}_m^2} \hat{\mathbf{b}}_\mathbf{m}^T = \hat{\mathbf{c}}_\mathbf{m} \hat{H}(-\hat{\lambda}_m) \frac{2}{\pi} \frac{d}{ds} \mathbf{atan}\left(\frac{\omega}{s}\right) \Big|_{s=-\hat{\lambda}_m} \hat{\mathbf{b}}_\mathbf{m}^T. \tag{8.20}$$

Equations (8.19) and (8.20) can be grouped under the following formulation

$$\hat{\mathbf{c}}_\mathbf{m} \hat{H}'(-\hat{\lambda}_m) \mathbf{a}_{\omega,\hat{\lambda}_m} \hat{\mathbf{b}}_\mathbf{m}^T + \hat{\mathbf{c}}_\mathbf{m} \hat{H}(-\hat{\lambda}_m) \frac{2}{\pi} \frac{d}{ds} \mathbf{atan}\left(\frac{\omega}{s}\right) \Big|_{s=-\hat{\lambda}_m} \hat{\mathbf{b}}_\mathbf{m}^T = \hat{\mathbf{c}}_\mathbf{m} \frac{d}{ds} \left[\hat{H}(s) \mathbf{a}_{\omega,-s} \right]_{s=-\hat{\lambda}_m} \hat{\mathbf{b}}_\mathbf{m}^T.$$

The same result holds for the second sum involved in $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\lambda}_m}$. Thus $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\lambda}_m} = 0$ can finally be re-written as

$$\hat{\mathbf{c}}_\mathbf{m} \left(H'_\omega(-\hat{\lambda}_m) + \frac{d}{ds} \left[H(s) \mathbf{a}_{\omega,-s} \right]_{s=-\hat{\lambda}_m} \right) \hat{\mathbf{b}}_\mathbf{m}^T = \hat{\mathbf{c}}_\mathbf{m} \left(\hat{H}'_\omega(-\hat{\lambda}_m) + \frac{d}{ds} \left[\hat{H}(s) \mathbf{a}_{\omega,-s} \right]_{s=-\hat{\lambda}_m} \right) \hat{\mathbf{b}}_\mathbf{m}^T \tag{8.21}$$

Interpolation conditions. Equations (8.17), (8.18) and (8.21) can be further simplified by defining the two functions[4]

$$T(s) = H_\omega(s) + H(s)\mathbf{a}_{\omega,-s} \quad \text{and} \quad \hat{T}(s) = \hat{H}_\omega(s) + \hat{H}(s)\mathbf{a}_{\omega,-s}.$$

[4]Since $\forall z$, $\mathbf{atan}(-z) = -\mathbf{atan}(z)$, these functions can be written as $T(s) = H_\omega(s) - H(s)\mathbf{a}_{\omega,s}$ and $\hat{T}(s) = \hat{H}_\omega(s) - \hat{H}(s)\mathbf{a}_{\omega,s}$ also.

Indeed, with that notation, the first-order optimality conditions of the $\mathcal{H}_{2,\Omega}$ approximation problem associated with the poles and the residues of the reduced-order model can then be written as

$$\begin{array}{rcl} T(-\hat{\lambda}_m)\hat{\mathbf{b}}_\mathbf{m}^T & = & \hat{T}(-\hat{\lambda}_m)\hat{\mathbf{b}}_\mathbf{m}^T \\ \hat{\mathbf{c}}_\mathbf{m}T(-\hat{\lambda}_m) & = & \hat{\mathbf{c}}_\mathbf{m}\hat{T}(-\hat{\lambda}_m) \\ \hat{\mathbf{c}}_\mathbf{m}T'(-\hat{\lambda}_m)\hat{\mathbf{b}}_\mathbf{m}^T & = & \hat{\mathbf{c}}_\mathbf{m}\hat{T}'(-\hat{\lambda}_m)\hat{\mathbf{b}}_\mathbf{m}^T, \end{array} \tag{8.22}$$

which is similar to the $\mathcal{H}_2$ optimality conditions expressed in Theorem 9 excepted that here the interpolation does not concern directly the large-scale and reduced-order transfer functions. The functions $T(s)$ and $\hat{T}(s)$ are probably related to the weighted functions involved in the frequency weighted $\mathcal{H}_2$ optimality conditions derived in [Anić et al., 2013] (see Section 5.1.2) but a clear link has not been established yet. Note that the function $T(s)$ also appears in the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ in equation (8.6) which becomes,

$$\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \|H\|^2_{\mathcal{H}_{2,\Omega}} + \|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}} + \sum_{i=1}^{n} \mathbf{tr}\left(T(-\hat{\lambda}_k)\hat{\Phi}_k^T\right).$$

Derivative with respect to $\hat{D}$. The derivative of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ with respect to $\hat{D}$ leads to a quite different optimality condition in terms of the transfer functions $H(s)$ and $\hat{H}(s)$, indeed, setting $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{D}} = 0$ in equation (8.16) yields, for $i = 1, \ldots, n_y$ and $k = 1, \ldots, n_u$,

$$\int_{-\omega}^{\omega} H_{i,k}(j\nu)d\nu = \int_{-\omega}^{\omega} \hat{H}_{i,k}(j\nu)d\nu.$$

The direct feedthrough $\hat{D}$ of the reduced-order model is involved in a quadratic and strictly convex way in the $\mathcal{H}_{2,\Omega}$ approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ (8.1), indeed the second-order term is

$$\mathbf{tr}(\hat{D}\hat{D}^T) = \mathbf{vec}(\hat{D})^T\mathbf{vec}(\hat{D}) = \mathbf{vec}(\hat{D})^T I_{n_y n_u}\mathbf{vec}(\hat{D}),$$

where $\mathbf{vec}(\hat{D})$ is the vectorisation of $\hat{D}$, *i.e.* the column vector formed by vertically concatenating the columns of $\hat{D}$. Hence the necessary condition of optimality $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{D}} = 0$ is also sufficient to find the global minimum of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ with respect to $\hat{D}$. The optimal $\hat{D}$ is uniquely determined as a function of the residues and poles of both models $\mathbf{H}$ and $\hat{\mathbf{H}}$, the direct feedthrough D and the upper bound of the frequency interval ω,

$$\hat{D}_{opt} = \frac{\pi}{2\omega}\left(\sum_{i=1}^{r} \hat{\mathbf{c}}_\mathbf{i}^T\hat{\mathbf{b}}_\mathbf{i}\mathbf{a}_{\omega,\hat{\lambda}_i} - \sum_{i=1}^{n} \mathbf{c_i}^T\mathbf{b_i}\mathbf{a}_{\omega,\lambda_i} + \frac{2}{\pi}\omega D\right). \tag{8.23}$$

This implies that $\hat{D}$ can be discarded as an optimisation variable and replaced by its optimal value in each one of the derivatives (8.13), (8.14) and (8.15).

Simplification of the $\mathcal{H}_{2,\Omega}$ approximation error

Let us consider the formulation of the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ of equation (8.5). Let us assume that either $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{c}}_\mathbf{m}} = 0$ or $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{m}} = 0$ for $m = 1, \ldots, r$, then

$$\sum_{k=1}^{r} \hat{\mathbf{b}}_\mathbf{k}\left(\sum_{i=1}^{n} \frac{\mathbf{b_i}^T\mathbf{c_i}}{\lambda_i + \hat{\lambda}_k}\left(\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}\right)\right)\hat{\mathbf{c}}_\mathbf{k}^T = \sum_{k=1}^{r} \hat{\mathbf{b}}_\mathbf{k}\left(\sum_{i=1}^{r} \frac{\hat{\mathbf{b}}_\mathbf{i}^T\hat{\mathbf{c}}_\mathbf{i}}{\hat{\lambda}_i + \hat{\lambda}_k}\left(\mathbf{a}_{\omega,\hat{\lambda}_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}\right) + \tilde{D}^T\mathbf{a}_{\omega,\hat{\lambda}_k}\right)\hat{\mathbf{c}}_\mathbf{k}^T.$$

By taking the trace and reordering the sums so that $\mathbf{a}_{\omega,\hat{\lambda}_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}$ is replaced by $2\mathbf{a}_{\omega,\hat{\lambda}_i}$, the $\mathcal{H}_{2,\Omega}$ approximation error becomes

$$\begin{aligned}\mathcal{J}_{\mathcal{H}_{2,\Omega}} &= \|H\|^2_{\mathcal{H}_{2,\Omega}} + \|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}} - 2\sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\hat{\Phi}_i\hat{\Phi}_k^T\right)}{\hat{\lambda}_i + \hat{\lambda}_k}\mathbf{a}_{\omega,\hat{\lambda}_i} + \sum_{i=1}^{r}\mathbf{tr}\left(\hat{\Phi}_i\hat{D}^T\right)\mathbf{a}_{\omega,\hat{\lambda}_i}\cdots \\ &\quad + \sum_{i=1}^{n}\mathbf{tr}\left(\Phi_i\hat{D}^T\right)\mathbf{a}_{\omega,\lambda_i} - \frac{2}{\pi}\omega\mathbf{tr}\left(D\hat{D}^T\right).\end{aligned}$$

If in addition $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{D}} = 0$, then

$$\mathbf{tr}\left(\left(\sum_{i=1}^{n}\Phi_i\mathbf{a}_{\omega,\lambda_i} - \frac{2}{\pi}\omega D\right)\hat{D}^T\right) = \mathbf{tr}\left(\left(\sum_{i=1}^{r}\hat{\Phi}_i\mathbf{a}_{\omega,\hat{\lambda}_i} - \frac{2}{\pi}\omega\hat{D}\right)\hat{D}^T\right),$$

thus

$$\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \|H\|^2_{\mathcal{H}_{2,\Omega}} + \|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}} - 2\left(\sum_{i=1}^{r}\sum_{k=1}^{r}\frac{\mathbf{tr}\left(\hat{\Phi}_i\hat{\Phi}_k^T\right)}{\hat{\lambda}_i + \hat{\lambda}_k}\mathbf{a}_{\omega,\hat{\lambda}_i} - \sum_{i=1}^{r}\mathbf{tr}\left(\hat{\Phi}_i\hat{D}^T\right)\mathbf{a}_{\omega,\hat{\lambda}_i} + \frac{\omega}{\pi}\mathbf{tr}\left(\hat{D}\hat{D}^T\right)\right).$$

The last term is equal to $\|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}}$, hence if $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{D}} = 0$ and either $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{c}}_\mathbf{m}} = 0$ or $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{m}} = 0$, the squared $\mathcal{H}_{2,\Omega}$ approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is equal to the difference of the squared $\mathcal{H}_{2,\Omega}$-norms of $\mathbf{H}$ and $\hat{\mathbf{H}}$,

$$\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \|H - \hat{H}\|^2_{\mathcal{H}_{2,\Omega}} = \|H\|^2_{\mathcal{H}_{2,\Omega}} - \|\hat{H}\|^2_{\mathcal{H}_{2,\Omega}}.$$

On the computation of the gradient of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$

The computation of the gradient is also an important part in an optimisation algorithm which must be efficient in order not to become a numerical burden for the process. The process here is similar to what has been done for the computation of the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ in the previous section and the matrices built to compute $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ are also involved here. Indeed, let us consider the matrices $W_{rr} \in \mathbb{C}^{n\times n}$, $W_{nr} \in \mathbb{C}^{r\times r}$, $C_\Delta \in \mathbb{C}^{n_y\times n}$, $\hat{C}_\Delta \in \mathbb{C}^{n_y\times r}$, $B_\Delta \in \mathbb{C}^{n\times n_u}$ and $\hat{B}_\Delta \in \mathbb{C}^{r\times n_u}$ defined in equations (8.7), (8.8), (8.9) and (8.10), respectively.

Let us consider $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{m}}$ which expression is

$$\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{m}} = \sum_{i=1}^{r}\frac{\hat{\mathbf{b}}_\mathbf{i}^T\hat{\mathbf{c}}_\mathbf{i}\hat{\mathbf{c}}_\mathbf{m}^T}{\hat{\lambda}_i + \hat{\lambda}_m}\left(\mathbf{a}_{\omega,\hat{\lambda}_i} + \mathbf{a}_{\omega,\hat{\lambda}_m}\right) - \sum_{i=1}^{n}\frac{\mathbf{b}_\mathbf{i}{}^T\mathbf{c}_\mathbf{i}\hat{\mathbf{c}}_\mathbf{m}^T}{\lambda_i + \hat{\lambda}_m}\left(\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_m}\right) + \tilde{D}^T\hat{\mathbf{c}}_\mathbf{m}^T\mathbf{a}_{\omega,\hat{\lambda}_m}.$$

First, notice that

$$\hat{C}_\Delta^T\hat{C}_\Delta = \begin{bmatrix}\hat{\mathbf{c}}_\mathbf{1}\\ \vdots \\ \hat{\mathbf{c}}_\mathbf{r}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{c}}_\mathbf{1}^T & \dots & \hat{\mathbf{c}}_\mathbf{r}^T\end{bmatrix} = \begin{bmatrix}\hat{\mathbf{c}}_\mathbf{1}\hat{\mathbf{c}}_\mathbf{1}^T & \dots & \hat{\mathbf{c}}_\mathbf{1}\hat{\mathbf{c}}_\mathbf{r}^T \\ \vdots & \ddots & \vdots \\ \hat{\mathbf{c}}_\mathbf{r}\hat{\mathbf{c}}_\mathbf{1}^T & \dots & \hat{\mathbf{c}}_\mathbf{r}\hat{\mathbf{c}}_\mathbf{r}^T\end{bmatrix} \in \mathbb{C}^{r\times r},$$

By multiplying $\hat{C}_\Delta^T\hat{C}_\Delta$, element-wise, with $2W_{rr}$, one obtains, in the m-th column,

$$\mathbf{t}_\mathbf{m} = \begin{bmatrix}\hat{\mathbf{c}}_\mathbf{1}\hat{\mathbf{c}}_\mathbf{m}^T\left(\mathbf{a}_{\omega,\hat{\lambda}_1} + \mathbf{a}_{\omega,\hat{\lambda}_m}\right) & \dots & \hat{\mathbf{c}}_\mathbf{r}\hat{\mathbf{c}}_\mathbf{m}^T\left(\mathbf{a}_{\omega,\hat{\lambda}_r} + \mathbf{a}_{\omega,\hat{\lambda}_m}\right)\end{bmatrix}^T \in \mathbb{C}^{r\times 1}, \tag{8.24}$$

which is a part of the elements involved in the first sum of $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{m}}$. By left multiplying the vector $\mathbf{t}_\mathbf{m}$ from equation (8.24) by $\hat{B}_\Delta^T$, it comes

$$\hat{B}_\Delta^T\mathbf{t}\mathbf{m} = \begin{bmatrix}\hat{\mathbf{b}}_\mathbf{1}^T & \dots & \hat{\mathbf{b}}_\mathbf{r}^T\end{bmatrix}\mathbf{t}_\mathbf{m} = \sum_{i=1}^{r}\frac{\hat{\mathbf{b}}_\mathbf{i}^T\hat{\mathbf{c}}_\mathbf{i}\hat{\mathbf{c}}_\mathbf{m}^T}{\hat{\lambda}_i + \hat{\lambda}_m}\left(\mathbf{a}_{\omega,\hat{\lambda}_i} + \mathbf{a}_{\omega,\hat{\lambda}_m}\right).$$

Thus the first sum of $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{m}}$ is the m-th column of $\hat{B}_\Delta^T \left(\left(\hat{C}_\Delta^T \hat{C} \right) \odot 2W_{rr} \right)$, *i.e.*

$$\left[\hat{B}_\Delta^T \left(\left(\hat{C}_\Delta^T \hat{C} \right) \odot 2W_{rr} \right) \right] \mathbf{e_m} = \sum_{i=1}^{r} \frac{\hat{\mathbf{b}}_\mathbf{i}^T \hat{\mathbf{c}}_\mathbf{i} \hat{\mathbf{c}}_\mathbf{m}^T}{\hat{\lambda}_i + \hat{\lambda}_m} \left(\mathbf{a}_{\omega,\hat{\lambda}_i} + \mathbf{a}_{\omega,\hat{\lambda}_m} \right). \tag{8.25}$$

Similarly for the second sum,

$$\left[B_\Delta^T \left(\left(C_\Delta^T \hat{C}_\Delta \right) \odot 2W_{nr} \right) \right] \mathbf{e_m} = \sum_{i=1}^{n} \frac{\mathbf{b_i}^T \mathbf{c_i} \hat{\mathbf{c}}_\mathbf{m}^T}{\lambda_i + \hat{\lambda}_m} \left(\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_m} \right). \tag{8.26}$$

By following the same idea, it is easy to see that,

$$\tilde{D}^T \hat{\mathbf{c}}_\mathbf{m}^T \mathbf{a}_{\omega,\hat{\lambda}_m} = \left[\left(\tilde{D}^T \hat{C}_\Delta \right) \odot \left(\mathbb{1}_{n_u} \left[\begin{array}{ccc} \mathbf{a}_{\omega,\hat{\lambda}_1} & \dots & \mathbf{a}_{\omega,\hat{\lambda}_r} \end{array} \right] \right) \right] \mathbf{e_m}. \tag{8.27}$$

By coupling equations (8.25), (8.26) and (8.27), on finally obtains

$$\begin{aligned} \left[\begin{array}{ccc} \frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{1}} & \dots & \frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{b}}_\mathbf{r}} \end{array} \right] &= \hat{B}_\Delta^T \left(\left(\hat{C}_\Delta^T \hat{C} \right) \odot 2W_{rr} \right) - B_\Delta^T \left(\left(C_\Delta^T \hat{C}_\Delta \right) \odot 2W_{nr} \right) \dots \\ &\quad + \left(\tilde{D}^T \hat{C}_\Delta \right) \odot \left(\mathbb{1}_{n_u} \left[\begin{array}{ccc} \mathbf{a}_{\omega,\hat{\lambda}_1} & \dots & \mathbf{a}_{\omega,\hat{\lambda}_r} \end{array} \right] \right). \end{aligned}$$

The same idea can be used for $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{c}}_\mathbf{m}}$ and leads to

$$\begin{aligned} \left[\begin{array}{ccc} \frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{c}}_\mathbf{1}} & \dots & \frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\mathbf{c}}_\mathbf{r}} \end{array} \right] &= \hat{C}_\Delta \left(\left(\hat{B}_\Delta \hat{B}_\Delta^T \right) \odot 2W_{rr} \right) - C_\Delta \left(\left(B_\Delta \hat{B}_\Delta^T \right) \odot 2W_{nr} \right) \dots \\ &\quad + \left(\tilde{D} \hat{B}_\Delta^T \right) \odot \left(\mathbb{1}_{n_y} \left[\begin{array}{ccc} \mathbf{a}_{\omega,\hat{\lambda}_1} & \dots & \mathbf{a}_{\omega,\hat{\lambda}_r} \end{array} \right] \right). \end{aligned}$$

The derivative with respect to the reduced-order model poles, $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\lambda}_m}$, is obtained in a very similar way to the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$. First, let us consider the matrices $L_{rr} \in \mathbb{C}^{r \times r}$ and $L_{nr} \in \mathbb{C}^{n \times r}$ which elements are defined as,

$$[L_{rr}]_{i,k} = \frac{\mathbf{a}_{\omega,\hat{\lambda}_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}}{\left(\hat{\lambda}_i + \hat{\lambda}_k \right)^2} + \frac{2\omega}{\pi \left(\hat{\lambda}_k^2 + \omega^2 \right) \left(\hat{\lambda}_i + \hat{\lambda}_k \right)},$$

and

$$[L_{nr}]_{i,k} = \frac{\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}}{\left(\lambda_i + \hat{\lambda}_k \right)^2} + \frac{2\omega}{\pi \left(\hat{\lambda}_k^2 + \omega^2 \right) \left(\lambda_i + \hat{\lambda}_k \right)}.$$

By multiplying L_{rr} (respectively L_{nr}), element-wise, by $\hat{B}_\Delta \hat{B}_\Delta^T$ (resp. $B_\Delta \hat{B}_\Delta^T$) and by $\hat{C}_\Delta^T \hat{C}_\Delta$ (resp. $\hat{C}_\Delta^T C_\Delta$), one obtains a matrix which $m-th$ column contains the terms in the first (resp. second) sum of $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \hat{\lambda}_m}$, *i.e.*

$$\left[\left(\hat{B}_\Delta \hat{B}_\Delta^T \right) \odot \left(\hat{C}_\Delta^T \hat{C}_\Delta \right) \odot L_{rr} \right]_{i,m} = \hat{\mathbf{c}}_\mathbf{m} \hat{\mathbf{c}}_\mathbf{i}^T \hat{\mathbf{b}}_\mathbf{i} \hat{\mathbf{b}}_\mathbf{m}^T \left(\frac{\mathbf{a}_{\omega,\hat{\lambda}_i} + \mathbf{a}_{\omega,\hat{\lambda}_m}}{\left(\hat{\lambda}_i + \hat{\lambda}_m \right)^2} + \frac{2}{\pi} \frac{\omega}{\left(\hat{\lambda}_m^2 + \omega^2 \right) \left(\hat{\lambda}_i + \hat{\lambda}_m \right)} \right),$$

and

$$\left[\left(B_\Delta \hat{B}_\Delta^T \right) \odot \left(\hat{C}_\Delta^T C_\Delta \right) \odot L_{nr} \right]_{i,m} = \hat{\mathbf{c}}_\mathbf{m} \mathbf{c_i}^T \mathbf{b_i} \hat{\mathbf{b}}_\mathbf{m}^T \left(\frac{\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_m}}{\left(\lambda_i + \hat{\lambda}_m \right)^2} + \frac{2}{\pi} \frac{\omega}{\left(\hat{\lambda}_m^2 + \omega^2 \right) \left(\lambda_i + \hat{\lambda}_m \right)} \right).$$

Besides, the term relative to the direct feedthrough can be obtained as

$$\frac{2}{\pi}\hat{\mathbf{c}}_\mathbf{m}\tilde{D}\hat{\mathbf{b}}_\mathbf{m}\frac{\omega}{\omega^2+\hat{\lambda}_m^2} = \mathbf{e_m}^T\left[\mathbf{diag}\left(\hat{B}_\Delta\tilde{D}^T\hat{C}_\Delta\right)\odot\begin{bmatrix}\frac{2\omega}{\pi\left(\hat{\lambda}_1^2+\omega^2\right)}\\ \vdots \\ \frac{2\omega}{\pi\left(\hat{\lambda}_r^2+\omega^2\right)}\end{bmatrix}\right].$$

Hence, it finally comes that

$$\begin{aligned}\begin{bmatrix}\frac{\partial\mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial\hat{\lambda}_1} & \dots & \frac{\partial\mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial\hat{\lambda}_r}\end{bmatrix} = & -\mathbb{1}_r^T\left(\left(\hat{B}_\Delta\hat{B}_\Delta^T\right)\odot\left(\hat{C}_\Delta^T\hat{C}_\Delta\right)\odot L_{rr}\right)\dots\\ & +\mathbb{1}_n^T\left(\left(B_\Delta\hat{B}_\Delta^T\right)\odot\left(\hat{C}_\Delta^T C_\Delta\right)\odot L_{nr}\right)\\ & +\mathbf{diag}\left(\hat{B}_\Delta\tilde{D}^T\hat{C}_\Delta\right)^T\odot\begin{bmatrix}\frac{2\omega}{\pi\left(\hat{\lambda}_1^2+\omega^2\right)} & \cdots & \frac{2\omega}{\pi\left(\hat{\lambda}_r^2+\omega^2\right)}\end{bmatrix}.\end{aligned}$$

Just like the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$, its gradient with respect to the reduced-order parameters can be easily computed with simple matrix and vector operations. Again, in comparison, with the gramian-based approach, in addition of what is required to compute the error, one $n\times r$ Sylvester equation and one $r\times r$ Lyapunov equation must be solved and the derivative of a complex matrix logarithm must be computed (see Section 5.2.2).

Conclusion

In this Chapter, the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm has been used to express the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ between a large-scale model and a reduced-order one as a function of its poles and residues. There exist similarities between the structures of the $\mathcal{H}_2$ approximation error (4.3) and $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ (8.5) but in the frequency-limited case, additional weightings prevent from formulating the cross terms as the transfer function of the large-scale model evaluated at the opposite of the reduced-order model poles.

The approximation error $\mathcal{H}_{2,\Omega}$ has then been differentiated with respect to the reduced-order model poles and residues in order to express the necessary conditions for optimality. Similarly to the $\mathcal{H}_2$ case, these optimality conditions can be expressed as interpolation ones. Yet they concern some modified transfer functions (8.22) and not directly the models transfer functions. Hence it is not clear if an iterative procedure as the **IRKA** (see Section 4.2.1) could be developed for the $\mathcal{H}_{2,\Omega}$ approximation. That is why an optimisation procedure is developed in the next Chapter.

Chapter 9

Development of a descent algorithm for the optimal $\mathcal{H}_{2,\Omega}$ approximation problem

Since most of the readily available optimisation solvers do not handle real functions of complex variables directly in the complex domain[1], the poles-residues formulation of the $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ error and its gradient derived in the previous chapter are used here to develop a complex-domain descent algorithm. It is called *Descent Algorithm for Residues and Poles Optimisation* (**DARPO**) and relies on the complex-domain optimisation framework developed in [Sorber et al., 2012]. It is aimed at finding a reduced-order model $\hat{\mathbf{H}}$ that satisfies the first-order optimality conditions for the optimal $\mathcal{H}_{2,\Omega}$ approximation problem. More specifically, by writing the transfer function $\hat{H}(s)$ associated with $\hat{\mathbf{H}}$ as

$$\hat{H}(s) = \sum_{i=1}^{r} \frac{\hat{\mathbf{c}}_\mathbf{i}^T \hat{\mathbf{b}}_\mathbf{i}}{s - \hat{\lambda}_i} + \hat{D},$$

DARPO is aimed at finding the poles $\hat{\lambda}_i$, the residues $\{\hat{\mathbf{c}}_\mathbf{i}, \hat{\mathbf{b}}_\mathbf{i}\}$ $(i = 1, \ldots, r)$ and the direct feedthrough $\hat{D}$ that fulfil the necessary conditions for optimality presented in Theorem 16. Using this framework implies that both models $\mathbf{H}$ and $\hat{\mathbf{H}}$ have semi-simple poles only and that their eigenvalues are not located symmetrically with respect to the imaginary axis. These hypothesis on the structure of the poles restrict the models on which the method can be applied but enables to efficiently compute the approximation error and its gradient.

The algorithm is described in Section 9.1 and its behaviour is illustrated on academic examples in Section 9.2.

The results presented in this Chapter have partly been published in [Vuillemin et al., 2014b].

Contents

9.1 A descent algorithm in the complex domain 129
9.1.1 Reminder of unconstrained optimisation 130
9.1.2 Descent Algorithm for Residues and Poles Optimisation 133
9.1.3 Initialisation of **DARPO** 137
9.2 Numerical illustration 139
9.2.1 Illustration of the behaviour of **DARPO** 139
9.2.2 Improvement provided by **DARPO** 146
9.2.3 Application for the control of a large-scale model 151

9.1 A descent algorithm in the complex domain

In Section 9.1.1, a quick reminder of some aspects of unconstrained optimisation of real functions of real variables are recalled. Then, in Section 9.1.2, the optimisation algorithm for optimal $\mathcal{H}_{2,\Omega}$ model approximation is presented. Finally, the initialisation process used in the algorithm is detailed in Section 9.1.3.

[1] L. Sorber and his colleagues have now proposed a complex-domain optimisation toolbox in [Sorber et al., 2013].

9.1.1 Reminder of unconstrained optimisation

In this Section, some elements about unconstrained optimisation are recalled. For sake of simplicity, only the case of real functions of real variables is presented, but the ideas are the same for the unconstrained optimisation of real functions of complex variables and the corresponding algorithms are similar.

Several algorithms are available to perform unconstrained optimisation (see for instance [Fletcher, 2000]). Here, a quasi-Newton method based on the *Broyden–Fletcher–Goldfarb–Shanno* (BFGS) update of the inverse Hessian approximation is considered. This method has been chosen because it does not require the Hessian of the objective function and is known for its good performances in practice [Li and Fukushima, 2001].

In the sequel, the ideas behind the BFGS update and the algorithm used to address the line-search sub-problem are presented.

The BFGS update

Let us consider a twice differentiable real-valued function, $f : \mathbb{R}^n \to \mathbb{R}$ of a real vector $\mathbf{x}$ that has to be minimised. The idea behind the *Newton iteration* is to replace f by its quadratic approximation q_k around the iterate $\mathbf{x_k}$,

$$f(\mathbf{x_k} + \mathbf{p}) \approx q_k(\mathbf{p}) = f_k + \mathbf{p}^T \mathbf{g_k} + \frac{1}{2}\mathbf{p}^T H_k \mathbf{p},$$

where f_k, $\mathbf{g_k}$ and H_k are the function, its gradient and its Hessian at $\mathbf{x_k}$, respectively, and to minimise q_k with respect to $\mathbf{p}$ instead of f. Since q_k is quadratic in $\mathbf{p}$ the necessary conditions for optimality leads to

$$\mathbf{p_k} = -{H_k}^{-1}\mathbf{g_k},$$

which is a minimum if and only if H_k is positive definite. Then, the Newton iteration is given by $\mathbf{x_{k+1}} = \mathbf{x_k} + \alpha_k \mathbf{p_k}$ with $\alpha_k = 1$.

The Hessian H can be large, hence computing and inverting it can be tedious. Alternatively, it can simply not be available. The *quasi-Newton* methods are aimed at alleviating these issues by estimating the Hessian H_k or directly its inverse ${H_k}^{-1}$ from the successive values f_k and $\mathbf{g_k}$. To this aim, let us define the position and gradient increments $\mathbf{s_k}$ and $\mathbf{y_k}$, respectively, as,

$$\mathbf{s_k} = \mathbf{x_{k+1}} - \mathbf{x_k} \quad \text{and} \quad \mathbf{y_k} = \mathbf{g_{k+1}} - \mathbf{g_k}.$$

From a Taylor expansion, one obtains

$$\mathbf{y_k} = H_k \mathbf{s_k} + o(\mathbf{s_k}). \tag{9.1}$$

One looks for a matrix B_{k+1} that mimics the property of the true Hessian (9.1), *i.e.* that satisfies *the secant equation*

$$\mathbf{y_k} = B_{k+1}\mathbf{s_k}. \tag{9.2}$$

For $n > 1$, there is no unique solution to equation (9.2), thus there are several ways to compute B_{k+1} or its inverse. Two common methods are the *Davidson-Fletcher-Powell* (DFP) method and the BFGS update (and its limited memory extension). The latter consists in selecting the next inverse approximation of the Hessian, B_{k+1}^{-1}, as the symmetric positive definite matrix satisfying (9.2) which is the closest of B_k^{-1} with respect to the Frobenius norm. The inverse BFGS update is given by

$$B_{k+1}^{-1} = \left(I_n - \frac{\mathbf{s_k}{\mathbf{y_k}}^T}{{\mathbf{y_k}}^T\mathbf{s_k}}\right) B_k^{-1} \left(I_n - \frac{\mathbf{y_k}{\mathbf{s_k}}^T}{{\mathbf{y_k}}^T\mathbf{s_k}}\right) + \frac{\mathbf{s_k}{\mathbf{s_k}}^T}{{\mathbf{y_k}}^T\mathbf{s_k}}.$$

With this quasi-Newton method, the search direction is given by $\mathbf{p_k} = -B_k^{-1}\mathbf{g_k}$. If B_k is positive definite, then

$${\mathbf{p_k}}^T\mathbf{g_k} = -{\mathbf{g_k}}^T B_k^{-1}\mathbf{g_k} < 0,$$

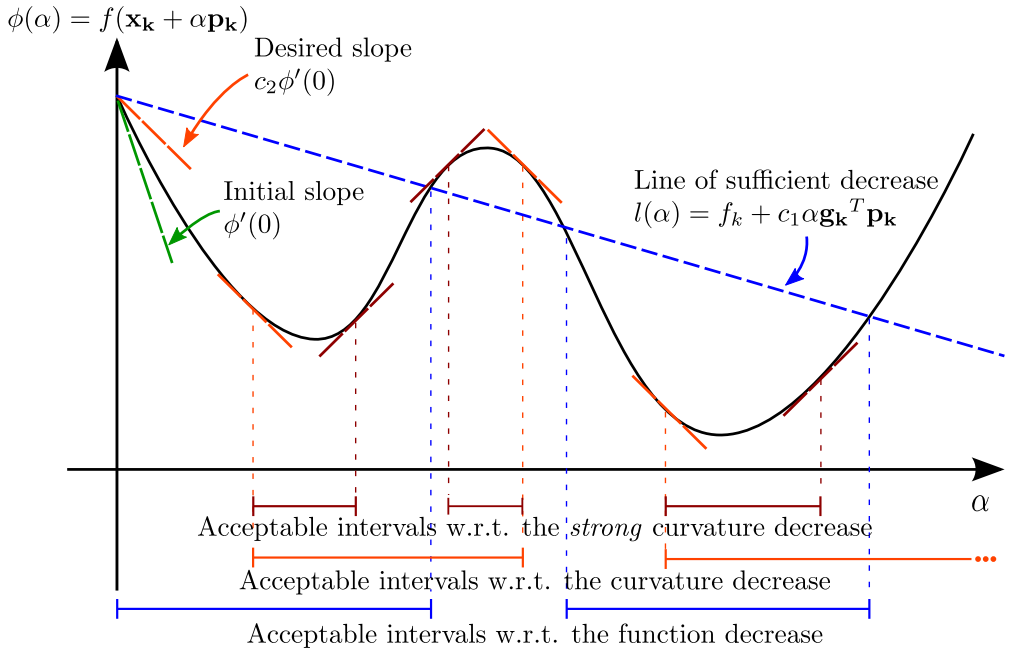


Figure 9.1: Illustration of the Armijo, Wolfe and strong Wolfe conditions.

which means that $\mathbf{p_k}$ is a descent direction.

The position update is given by $\mathbf{x_{k+1}} = \mathbf{x_k} + \alpha_k \mathbf{p_k}$, where the step length α_k is no longer necessarily equal to one as in the Newton iteration. It has to be determined with a line-search algorithm which basic principle is described thereafter.

The line-search algorithm

The step length α_k should ideally be chosen so that it is the global minimiser of the function

$$\phi(\alpha) = f(\mathbf{x_k} + \alpha \mathbf{p_k}), \quad (\alpha > 0)$$

which represents the evolution of the objective function f in the direction $\mathbf{p_k}$ from the point $\mathbf{x_k}$. In practice, finding the optimal value of α is too expensive and an *inexact line-search* is used instead. This approach consists in finding a step-length α_k that ensures a *sufficient decrease* of f in the direction $\mathbf{p_k}$.

The sufficient decrease of the objective function is generally represented by the *Armijo* condition which states that α_k should satisfies

$$\phi(\alpha_k) = f(\mathbf{x_k} + \alpha_k \mathbf{p_k}) \leq \underbrace{f_k + c_1 \alpha_k \mathbf{g_k}^T \mathbf{p_k}}_{l(\alpha_k)}, \tag{9.3}$$

with $0 < c_1 < 1$. Equation (9.3) means that α_k must yield a value $\phi(\alpha_k)$ below the line $l(\alpha)$ which has a negative slope and begins at $f_k = \phi(0)$. This is illustrated by the blue intervals in Figure 9.1.

Since any sufficiently small α satisfies equation (9.3), the Armijo condition is not enough to ensure a sufficient progress along the direction $\mathbf{p_k}$. To alleviate this issue, it is completed by the *curvature condition*

$$\underbrace{\nabla f(\mathbf{x_k} + \alpha_k \mathbf{p_k})^T \mathbf{p_k}}_{\phi'(\alpha_k)} \geq c_2 \underbrace{\mathbf{g_k}^T \mathbf{p_k}}_{\phi'(0)}, \tag{9.4}$$

where $c_1 < c_2 < 1$. Equation (9.4) states that the slope of ϕ at the point α_k must be greater than the slope at the initial point $\phi'(0) < 0$. In other words, α_k must lead to a point where moving along $\mathbf{p_k}$ does not result anymore in a decrease of the objective function as significant as at the initial point.

Equation (9.3) together with (9.4) represent the *Wolfe conditions*. These conditions are given by the intersection of the blue and orange intervals in Figure 9.1. Still, these conditions do not ensure that α_k leads to a point close to a local minimum, for instance, the rightmost point of the first blue interval in Figure 9.1 satisfies the Wolfe conditions since at that point $\phi'(\alpha) > 0 > c_2\phi'(0)$.

In order to be close to a local minimum of ϕ, the Wolfe conditions can be modified as

$$\begin{array}{rcl} f(\mathbf{x_k} + \alpha_k \mathbf{p_k}) & \leq & f_k + c_1 \alpha_k {\mathbf{g_k}}^T \mathbf{p_k} \\ |\nabla f(\mathbf{x_k} + \alpha_k \mathbf{p_k})^T \mathbf{p_k}| & \leq & c_2 |{\mathbf{g_k}}^T \mathbf{p_k}|, \end{array} \tag{9.5}$$

which are known as the *strong Wolfe conditions*. Adding the absolute value prevents the points α where ϕ has a too large positive slope to be kept as acceptable. The points that satisfy the strong curvature condition are represented by the dark red interval in Figure 9.1 and those that satisfy the strong Wolfe conditions are given by the intersection of the blue and dark red intervals.

Several algorithms have been developed to determine a step-length α_k that satisfies the (strong) Wolfe conditions. Here, the method proposed in [Fletcher, 2000, Section 2.6] has been chosen. It can be decomposed in two phases:

- **Bracketing phase.** This first step consists in finding an interval $[a, b]$ which contains an acceptable step length. This is achieved by creating a sequence of increasing step-lengths α_i. This phase stops either when one α_i is an acceptable step-length (*i.e.* it satisfies the strong Wolfe conditions) or when an interval that necessarily contains an acceptable step-length is found. In the former case, the line-search algorithm ends. The latter case happens when $\phi(\alpha_i) \geq l(\alpha_i)$, when $\phi(\alpha_i) \geq \phi(\alpha_{i-1})$ or when $\phi'(\alpha_i) \geq 0$ and leads to the sectioning phase.

- **Sectioning phase.** The second step generates a sequence of brackets $[a_i, b_i]$ which length decreases until an acceptable step-length α_k has been found. At each iteration i of the sectioning phase, a trial step-length α_i is chosen to lie inside $[a_i, b_i]$. This selection is usually achieved such that α_i minimises a quadratic or cubic interpolation of ϕ on the interval. Then either a_i or b_i is replaced by α_i to generate the next interval $[a_{i+1}, b_{i+1}]$. An iteration of this phase is illustrated in Figure 9.2 for the various cases that can arise. The sectioning step stops when an acceptable step-length is found.

In practice, additional precautions must be taken with the sectioning phase. In particular

(i) when selecting α_i in $[a_i, b_i]$ by interpolation, one must prevent α_i from being too close to the bounds of the interval. In [Fletcher, 2000], this is achieved by selecting α_i inside $[a_i + \tau_2(b_i - a_i), b_i - \tau_3(b_i - a_i)]$ where $0 < \tau_2 < \tau_3 \leq \frac{1}{2}$.

(ii) Near a stationary point, two consecutive values of the objective function $\phi(0) = f(\mathbf{x_k})$ and $\phi(\alpha_k) = f(\mathbf{x_k} + \alpha_k \mathbf{p_k})$ might become too close, *i.e.* of the order of magnitude of the machine precision. Hence a stopping criterion should be added to avoid any further irrelevant iteration.

(iii) Similarly, the optimisation process should stop whenever the change in position from $\mathbf{x_k}$ to $\mathbf{x_{k+1}}$ is no longer sufficient.

The tools for unconstrained optimisation that have been described here are used in the next session to perform optimal approximation in the $\mathcal{H}_{2,\Omega}$-norm.

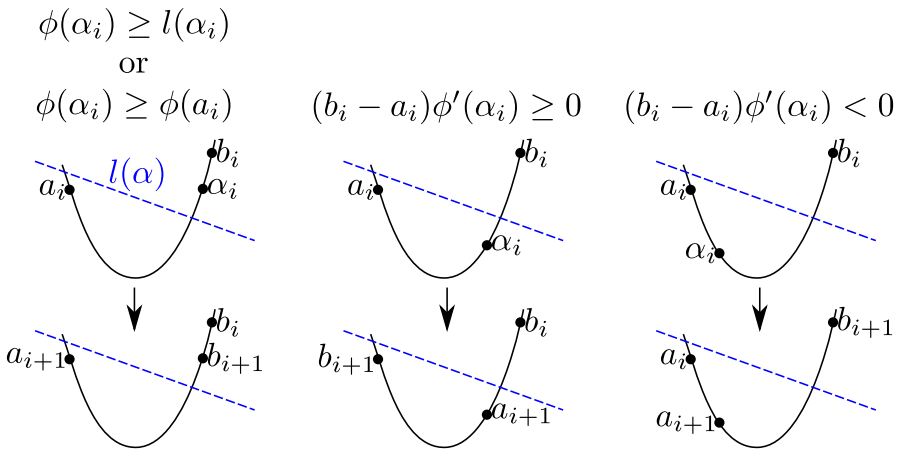


Figure 9.2: Illustration of the sectioning phase of the line-search algorithm.

9.1.2 Descent Algorithm for Residues and Poles Optimisation

The complex-domain optimisation algorithm developed here is mainly based on [Sorber et al., 2012] where the BFGS algorithm is formally extended for the optimisation of real functions of complex variables. In practice, the resulting optimisation scheme is the same as its real counterpart described in the previous Section excepted that the gradient becomes the conjugate complex gradient, and that the transpose operations are replaced by conjugate transpose operations.

The method is called **DARPO** and is presented in Algorithm 11. Given the realisation of a large-scale model $\mathbf{H} := (A, B, C, D)$ and a frequency interval Ω, the first step consists in computing the eigenvalue decomposition of the large-scale matrix A in order to build the matrices C_Δ and B_Δ involved in the computation of the approximation error and its gradient (see Remark 18 for some comments on this step). Then an initial reduced-order model is chosen as a vector $\mathbf{z_0}$ containing the poles $\hat{\lambda}_i$ and residues $\{\hat{\mathbf{c}}_\mathbf{i}, \hat{\mathbf{b}}_\mathbf{i}\}$ $(i = 1, \ldots, r)$. The process of initialisation is detailed in Section 9.1.3. Note that as explained in Remark 19, $\hat{D}$ does not need to be taken into account among the optimisation variables.

Then, while a stationary point has not been reached, the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ is computed together with its gradient (step 6) at the current point $\mathbf{z_k}$. Then the conjugate of the gradient of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ with respect to $\mathbf{z}$ is used, together with the current inverse Hessian approximation ${B_k}^{-1}$ to set the descent direction $\mathbf{p_k}$. At step 8, the step-length α_k that yields a sufficient decrease of the objective function in the direction $\mathbf{p_k}$ is determined with the inexact line-search algorithm described in Section 9.1.1. Then, at steps 9 and 10, the next point $\mathbf{z_{k+1}}$ is created and the inverse Hessian approximation is updated. Finally, the reduced-order model is built from the optimal poles, residues and direct feedthrough at step 13.

Additionally, the tools developed in Section 7.3 can be used to bound the $\mathcal{H}_\infty$-norm of the error model and to build a frequency template that bounds the Frobenius norm of the error model transfer function (and thus its maximum singular value). Since the eigen-decomposition of the large-scale and reduced-order models are already available, these optional steps can be achieved with a limited computational cost.

As the method **FL-ISTIA** presented in Chapter 6, **DARPO** has been integrated to the MORE Toolbox [Poussot-Vassal and Vuillemin, 2012] and an example of its use can be found in Appendix C.7.

Remark 18 (Computation of the eigen-decomposition of the large-scale model)**.** *The first step of **DARPO** requires to compute the full eigen-decomposition of the initial large-scale model* $\mathbf{H}$*. This is computationally expensive and cannot be performed for too large models. And, unlike in*

the $\mathcal{H}_2$-case, one cannot avoid this decomposition because here the eigenvalue of the large-scale model explicitly appear in the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$.

Hence, and as the other results based on the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm, this algorithm is restricted to the cases for which the eigen-decomposition is actually available. This issue is not specific to this formulation though. Indeed, with the gramian formulation, the logarithm of a complex $n \times n$ matrix has to be computed and at each iteration, two $n \times r$ Sylvester equations have to be solved.

Some hints to alleviate the issue related to the computation of the full eigen-decomposition are presented as a perspective in Chapter 11.

Remark 19 (About the direct feedthrough $\hat{D}$ in **DARPO**)**.** *As mentioned in Section 8.2.2, the optimal direct feedthrough $\hat{D}$ can be determined uniquely from equation (8.23). Hence in Algorithm 11, it is assumed that each time the objective function $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ or its gradient is evaluated, $\hat{D}$ is replaced with its optimal value.*

The following remarks can be addressed about Algorithm 11

- Step 6 : one part of the gradient of the approximation error is the complex conjugate of the other. To decrease the number of operations required to evaluate the gradient, one can compute it for the real variables and those which imaginary part is positive (or negative) only and then re-construct the complete gradient by conjugating its complex elements.

- Step 13 : the realisation of the reduced-order model $\hat{\mathbf{H}} = (\hat{A}, \hat{B}, \hat{C})$ is obtained by creating arbitrary complex vectors $\hat{\mathbf{x}}_\mathbf{i} \in \mathbb{C}^{r\times 1}$ associated with the eigenvalues $\hat{\lambda}_i$ which respect the complex conjugation of the eigenvalues, *i.e.* if $\hat{\lambda}_j = \hat{\lambda}_i^*$ then $\hat{\mathbf{x}}_\mathbf{j} = \hat{\mathbf{x}}_\mathbf{i}^*$, and which are linearly independent. Then, by denoting $\hat{X} \in \mathbb{C}^{r\times r}$ the matrix which columns are the vectors $\hat{\mathbf{x}}_\mathbf{i}$, $\Delta \in \mathbb{C}^{r\times r}$ the diagonal matrix containing the eigenvalues $\hat{\lambda}_i$, $\hat{C}_\Delta \in \mathbb{C}^{n_y\times r}$ (respectively $\hat{B}_\Delta \in \mathbb{C}^{r\times n_u}$) the matrix which columns (lines) are the $\hat{\mathbf{c}}_\mathbf{i}^T$ ($\hat{\mathbf{b}}_\mathbf{i}$),

$$\begin{aligned} \hat{A} &= \hat{X}\Delta\hat{X}^{-1} \\ \hat{B} &= \hat{X}\hat{B}_\Delta \\ \hat{C} &= \hat{C}_\Delta\hat{X}^{-1}. \end{aligned}$$

  This is a basis change between the complex diagonal realisation $(\Delta, \hat{B}_\Delta, \hat{C}_\Delta)$ of $\hat{\mathbf{H}}$ and a real realisation. The vectors $\hat{\mathbf{x}}_\mathbf{i}$ can for instance be chosen so that the real realisation $(\hat{A}, \hat{B}, \hat{C})$ is block diagonal.

- Step 14 : since the poles and residues of both the large-scale and reduced-order models are already known, the computation of the upper bounds $\bar{\Gamma}_\Omega(H - \hat{H})$ and $\Gamma_\Omega(H - \hat{H})$ does not require to solve any additional large-scale eigenvalue problem. Indeed the poles of the error model $\mathbf{E} = \mathbf{H} - \hat{\mathbf{H}}$ are the poles of $\mathbf{H}$ and $\hat{\mathbf{H}}$ while its residues are composed of the residues of $\mathbf{H}$ and the opposite of the residues of $\hat{\mathbf{H}}$. The same remark holds for Step 15.

- Note that the stopping criterion $\|\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial z}|_{\mathbf{z}=\mathbf{z_k}}\| \leq \epsilon$ is the ideal one[2] and indicates that the algorithm has reached a stationary point. Yet, as suggested in Section 9.1.1, one also checks the progress made at each iteration in the direction $\mathbf{p_k}$ through the difference between two consecutive iterates $\mathbf{z_k}$ and $\mathbf{z_{k+1}}$ and between the corresponding values of the objective function. One can also directly check the norm of the derivative of $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ in the direction $\mathbf{p_k}$ before performing a line-search, *i.e.*

$$\left\| \mathbf{p_k^H} \frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial z^*}\Big|_{\mathbf{z}=\mathbf{z_k}} \right\|. \tag{9.6}$$

[2] By ideal stopping criterion, one means that it is the most relevant criterion available here to determine if the current point is close to a stationary point.

Algorithm 11 Optimisation scheme of **DARPO**

Require: An initial model realisation $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n_u}$, $C \in \mathbb{R}^{n_y\times n}$, $D \in \mathbb{R}^{n_y\times n_u}$, a frequency interval Ω and a reduction order $r \in \mathbb{N}^*$.

1: Solve the large-scale eigenvalue problem

$$AX = X\Delta,$$

where $\Delta = \mathbf{diag}\,(\lambda_1, \ldots, \lambda_n)$, and compute

$$C_\Delta = \left[\begin{array}{ccc} \mathbf{c_1}^T & \ldots & \mathbf{c_n}^T \end{array}\right] = CX \quad \text{and} \quad B_\Delta = \left[\begin{array}{ccc} \mathbf{b_1}^T & \ldots & \mathbf{b_n}^T \end{array}\right]^T = X^{-1}B.$$

2: Choose an initial point z_0,

$$\mathbf{z_0} = \left[\begin{array}{cccccccc} \hat\lambda_1^{(0)} & \ldots & \hat\lambda_r^{(0)} & \hat{\mathbf{c}}_\mathbf{1}^{(0)} & \ldots & \hat{\mathbf{c}}_\mathbf{r}^{(0)} & \hat{\mathbf{b}}_\mathbf{1}^{(0)} & \ldots & \hat{\mathbf{b}}_\mathbf{r}^{(0)} \end{array}\right]^T \in \mathbb{C}^{N\times 1},$$

which contains the poles $\hat\lambda_i$ and residues $\{\hat{\mathbf{c}}_\mathbf{i}, \hat{\mathbf{b}}_\mathbf{i}\}$ $(i = 1, \ldots, r)$ that represent the initial reduced-order model.

3: Set the initial inverse Hessian approximation $B_0^{-1} = I_N$.

4: k $\leftarrow$ 0.

5: **while** $\left\| \frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial z}\Big|_{\mathbf{z=z_k}} \right\| > \epsilon$ **do**

6: Compute the $\mathcal{H}_{2,\Omega}$ approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}(\mathbf{z_k})$ from equation (8.1) and its gradient at the point $\mathbf{z_k}$, $\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \mathbf{z}}\Big|_{\mathbf{z=z_k}}$ by using the derivatives with respect to $\hat\lambda_m$ and $\hat{\mathbf{c}}_\mathbf{m}$ and $\hat{\mathbf{b}}_\mathbf{m}$ presented in equations (8.13), (8.15) and (8.14).

7: Set the descent direction as $\mathbf{p_k} = -B_k^{-1}\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \mathbf{z}^*}\Big|_{\mathbf{z=z_k}} = -B_k^{-1}\left(\frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \mathbf{z}}\Big|_{\mathbf{z=z_k}}\right)^*$.

8: Determine the step length α_k that minimises $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ in the direction $\mathbf{p_k}$ and such that $\mathcal{J}_{\mathcal{H}_{2,\Omega}}(\mathbf{z_k} + \alpha_k\mathbf{p_k})$ satisfies the complex strong Wolfe conditions (9.5).

9: Set the new point as $\mathbf{z_{k+1}} = \mathbf{z_k} + \alpha_k\mathbf{p_k}$.

10: Build the next inverse Hessian approximation with the BFGS update

$$B_{k+1}^{-1} = \left(I_N - \frac{\mathbf{s_k y_k^H}}{\mathbf{y_k^H s_k}}\right) B_k^{-1} \left(I_N - \frac{\mathbf{y_k s_k^H}}{\mathbf{y_k^H s_k}}\right) + \frac{\mathbf{s_k s_k^H}}{\mathbf{y_k^H y_k}},$$

where $\mathbf{z_k} = \mathbf{z_{k+1}} - \mathbf{z_k}$ and $\mathbf{y_k} = \frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \mathbf{z}}\Big|_{\mathbf{z=z^*_{k+1}}} - \frac{\partial \mathcal{J}_{\mathcal{H}_{2,\Omega}}}{\partial \mathbf{z}^*}\Big|_{\mathbf{z=z_k}}$.

11: k $\leftarrow$ k+1.

12: **end while**

13: Use $\hat\lambda_i^{(k)}$, $\hat{\mathbf{c}}_\mathbf{i}^{(k)}$, $\hat{\mathbf{b}}_\mathbf{i}^{(k)}$ $(i = 1, \ldots, r)$ to construct $\hat A$, $\hat B$ and $\hat C$ and use equation (8.23) to build the optimal direct feedthrough $\hat D$.

14: **[Optional]** Compute the upper bounds $\bar\Gamma_\Omega(H - \hat H)$ and $\Gamma_\Omega(H - \hat H)$ with the methods from Section 7.3.

15: **[Optional]** Build a frequency template $G(s)$ that satisfies $\|G(j\omega)\|_F \geq \left\|H(j\omega) - \hat H(j\omega)\right\|_F$ for all $\omega \in \mathbb{R}$ with the method from Section 7.3.

If (9.6) is low (of the order of magnitude of the machine precision), then it means that the objective function is almost *flat* in the direction $\mathbf{p_k}$, *i.e.* that current point is a stationary point. Hence, the line-search algorithm is unlikely to find a significant step-length α_k.

About the stability of the reduced-order model. In its current state, there is no formal guarantee that **DARPO** preserves the stability of the initial large-scale model. Even if an unstable reduced-order model is not expected to be obtained from an initial stable reduced-order model (see Example 17), two problematic cases can arise

(i) there is actually an unstable local minimum,

(ii) the local minimum is stable but the line-search algorithm produces a step-length α_k that makes the poles switch from the left half-plane to the right half plane.

There is not much that can be done in case *(i)* without using constrained optimisation but case *(ii)* can be handled fairly easily in the unconstrained optimisation framework.

Indeed, let us consider a pole $\hat{\lambda}$ located in the left half-plane at the left (along the real axis) of its optimal value $\hat{\lambda}_{opt}$ which is also in the left half-plane, *i.e.*

$$\mathbf{Re}(\hat{\lambda}) < \mathbf{Re}(\hat{\lambda}_{opt}) < 0.$$

In that case, the real part of the descent direction $p \in \mathbb{C}$ associated with the pole $\hat{\lambda}$ is positive. If a too large step is performed in the direction p, then the next iterate of the pole, $\hat{\lambda} + \alpha p$, might ends up in the right half-plane. Whether this will happen depends on the line-search algorithm used to determine the step-length α (see Section 9.1.1).

First, note that the parameter $0 < c_1 < 1$ used in the definition of the line decrease $l(\alpha)$ in equation (9.3) controls the slope of l. More specifically, the smaller c_1 is, the flatter the decrease line is and the larger the interval of acceptable points with respect to the decrease condition is (this can be seen in Figure 9.1). The line $l(\alpha)$ is involved in both the bracketing and sectioning phase of the line-search algorithm and determines how far in the descent direction the algorithm can look. The further one looks in the direction p, the more likely a large acceptable step-length α will be found.

The initial step-length trial used in the bracketing phase also contributes to determine the size of the initial interval containing an acceptable step length. Choosing 1 as initial value is standard since the Newton iteration is mimicked when the approximate Hessian is representative enough of the real Hessian. Yet this might directly leads to consider an interval containing step-length sufficiently large to make $\hat{\lambda}$ moves in the right half-plane.

Hence, to avoid selecting a step-length α that satisfies the strong Wolfe conditions but leads $\hat{\lambda}$ in the right half-plane, two solutions can be considered,

- the first one consists in increasing c_1 to prevent the line-search algorithm from looking too far from the current point,
- the second one consists in verifying if the initial step-length trial $\alpha = 1$ leads $\hat{\lambda}$ to move in the right half-plane. If it does, then it can be replaced by

$$\alpha_{max} = -\frac{\mathbf{Re}(\hat{\lambda})}{\mathbf{Re}(p)},$$

which represents the maximum allowed step-length for the next pole $\hat{\lambda} + \alpha p$ to lies in the left half-plane. One must then ensure that the line-search algorithm chose a point in $[0, \alpha_{\max}[$.

Note that if $\mathbf{Im}(\hat{\lambda} + \alpha_{max} p) \in \Omega$, then α_{max} is necessarily an acceptable upper bound for the bracketing phase since the objective function is, by definition, infinite when some poles of the reduced-order model are on the imaginary axis inside Ω.

About the convergence. The optimisation method used in **DARPO** produces a sequences of reduced-order models that yield monotonically decreasing $\mathcal{H}_{2,\Omega}$ approximation errors. Whether the final reduced-order model is a local minimum or not cannot be ensured without precise knowledge of the curvature of the objective function. Indeed, the global convergence of the BFGS update used here is not fully known for non-convex functions. However in practice, the BFGS update is known to be efficient and in [Li and Fukushima, 2001], the authors show that the BFGS update is a specific case of their *cautious BFGS update* which is proven to make the optimisation method globally convergent. Hence, **DARPO** is expected to lead to a local minimum of the optimal $\mathcal{H}_{2,\Omega}$ approximation problem.

9.1.3 Initialisation of DARPO

Unlike the **FL-BT** or the **FL-ISTIA**, a relevant reduced-order model should be used as initial point with **DARPO**. Hence this algorithm should be considered as a refinement process which cannot be considered separately from how it is initialised. Indeed, **DARPO** is a local optimisation algorithm applied to a non-linear and non-convex optimisation problem, thus the selection of the initial reduced-order model is extremely important and determines the quality of the solution.

Various initialisation strategies can be considered :

- **Random initialisation.** The simplest strategy consists in using a randomly generated model as initial point.

  This initialisation method is straightforward but *should not be used* because *(i)* it is not deterministic and might produce different reduced-order models through successive launches on an identical use case, *(ii)* it does not take into account the frequency interval Ω over which the initial large-scale model has to be reduced and *(iii)* it does not exploit any information about the initial large-scale model while its eigenvalues and eigenvectors are known from step 1. In practice, this initialisation strategy performs poorly on non-trivial applications and leads to a local minimum far from being satisfactory.

- **Initialisation with another method.** The second initialisation strategy consists in using another model approximation method to produce an initial reduced-order model. Since the aim here is to perform optimal $\mathcal{H}_{2,\Omega}$ approximation, the model approximation method should take into account the frequency interval Ω. For instance, one could use one of the methods presented in Chapter 5 (such as the **FW-BT** or the **FL-BT**) or one could also use **FL-ISTIA**, the empirical method developed during this thesis and presented in Chapter 6.

  This strategy generally leads to reduced-order model of good quality (with respect to the $\mathcal{H}_{2,\Omega}$-norm of the error) and, depending on the initial model approximation that is used, it can be deterministic and exploit the eigen-decomposition of the matrix A (for computing the gramians for instance).

  Yet, since it implies to use two different model approximation methods, this strategy is not completely satisfactory but it may be considered when computation time is not an issue.

- **Initialisation with a modal truncation.** This initialisation strategy consists in setting the initial reduced-order model as a modal truncation of the initial large-scale model. This initial reduced-order model is the most natural since the eigen-decomposition of the initial model is computed at step 1 of Algorithm 11. As the objective here is to minimise $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$, a relevant criterion must be built to determine the poles that must be retained. To that aim, various criteria have been considered and the most relevant are,

– to keep the poles λ_i and associated residues Φ_i that maximise

$$\mathcal{J}_{init}^{(1)}(\lambda_i, \Phi_i) = \left\| \Phi_i \mathbf{atan}\left(\frac{\omega}{\lambda_i}\right) \right\|_2.$$

where $\Omega = [0, \omega]$ (see Remark 20 for more complex intervals). This criterion is similar to what can appear in standard modal approximation with in addition the weighting term which decreases the influence of the residues associated with poles which imaginary part is outside of Ω. $\mathcal{J}_{init}^{(1)}$ can easily be evaluated and does not add much computation.

– The second criterion consists in keeping the poles λ_i and associated residues Φ_i that have the maximum influence in the $\mathcal{H}_{2,\Omega}$-norm of $\mathbf{H}$, *i.e.* those that yield the largest

$$\mathcal{J}_{init}^{(2)}(\lambda_i, \Phi_i) = -\mathbf{Re}\left(\mathbf{tr}\left(\Phi_i H(-\lambda_i)^T\right)\mathbf{atan}\left(\frac{\omega}{\lambda_i}\right)\right), \tag{9.7}$$

where $\Omega = [0, \omega]$. Again, the function $\mathcal{J}_{init}^{(1)}$ can easily be evaluated[3].

– The third criterion consists in keeping the poles λ_i and associated residues Φ_i that would make the best second-order (or first-order if the pole is real) approximation of $\mathbf{H}$ over Ω, *i.e.* by denoting $H_i(s)$ the transfer functions defined as

$$\hat{H}_i(s) = \frac{\Phi_i}{s - \lambda_i} + \frac{\Phi_i^*}{s - \lambda_i^*},$$

the poles to retain are those that maximise

$$\mathcal{J}_{init}^{(3)}(\lambda_i, \Phi_i) = -\|H - \hat{H}_i\|_{\mathcal{H}_{2,\Omega}}^2.$$

This selection method is computationally more demanding than the two others since it requires to evaluate several approximation errors.

None of these poles selection method really stands out of the other in term of global performance but the two first ones are cheaper. The criterion $\mathcal{J}_{init}^{(1)}$ tends to be more relevant than $\mathcal{J}_{init}^{(2)}$ when the considered frequency interval contains only poles associated with small residues in comparison to other areas but might fail on simple unbounded intervals like $[0, \infty)$. Hence, $\mathcal{J}_{init}^{(2)}$ is considered as the default initialisation procedure in the sequel if nothing else is specified.

Independently of the poles selection method, this initialisation strategy based on the modal truncation fully exploits the fact that the eigen-decomposition of the matrix A is computed in **DARPO** and in addition it does not add much extra computation if $\mathcal{J}_{init}^{(1)}$ or $\mathcal{J}_{init}^{(2)}$ is used. Besides, the poles of the reduced-order model are generally not too far from the poles of the initial large-scale model which can be attractive from a practical point of view.

The main flaw of this initialisation strategy is that if the frequency interval Ω does not contain enough poles, then irrelevant poles outside of Ω are selected and they might not be able to move inside the frequency interval of interest as suggested by Example 17.

From a practical point of view, this initialisation strategy must also be completed with a safeguard procedure dealing with the incoherence that can appear between the reduction order r and the poles selected. For instance, when r is odd while the candidate poles are all complex, etc.

[3]Indeed, since the eigen-decomposition of the initial model is known, $H(-\lambda_i)$ can be expressed with the poles and residues of the transfer function and the criterion $\mathcal{J}_{init}^{(2)}$ can thus be computed in a similar way to the approximation error $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$ in Section 8.1.2.

Remark 20. *For a more complex interval* $\Omega = \bigcup_{k=1}^{n_\Omega} \left[\omega_1^{(k)}, \omega_2^{(k)}\right]$, *where* $\omega_1^{(k)} < \omega_2^{(k)}$, *the weighting function* $\mathbf{atan}(\frac{\omega}{\lambda_i})$ *becomes* $\sum_{k=1}^{n_\Omega} \mathbf{atan}(\frac{\omega_2^{(k)}}{\lambda_i}) - \mathbf{atan}(\frac{\omega_1^{(k)}}{\lambda_i})$.

Example 19 (Illustration of the initialisation strategy of **DARPO**)**.** *In this example, the poles selection criterion* $\mathcal{J}_{init}^{(2)}$ *given in equation* (9.7) *is computed as a function of the upper bound* ω *of the frequency interval* Ω *in order to determine how the influence of a pole evolves as the frequency interval cross its imaginary part.*

For each pole and associated residue $\{\lambda_i, \Phi_i\}$, $\mathcal{J}_{init}^{(2)}(\lambda_i, \Phi_i)$ *is computed for various values of* ω. *This is done with the LAH model and with the first input to first output transfer of the ISS model. The frequency-responses of the models, as well as the values of the criterion are plotted in Figure 9.3. The vertical dashed lines correspond to the imaginary part of the eigenvalues and share the same colour as their corresponding criterion* $\mathcal{J}_{init}^{(2)}$.

For both models, one can see that for poorly damped modes, the criterion $\mathcal{J}_{init}^{(2)}$ *increases drastically when* ω *crosses the imaginary part of the corresponding pole. This is particularly clear with the ISS model which has three modes that strongly dominate the others.*

Note that $\mathcal{J}_{init}^{(2)}$ *does not produce perfectly intuitive results. Indeed, if one would like to reduce the ISS model over* $\Omega = [0, \infty)$ *to an order* 4, *then one needs to select two couples of complex conjugate poles. The initialisation process will select firstly the pole associated with resonance (1), then the (3) and surprisingly not the (2) while the latter yields a higher resonance around* 20 *rad/s than the former does around* 40 *rad/s.*

It turns out that in that case, the intuitive poles selection strategy, when used as initial reduced-order model in ***DARPO****, yields a larger* $\mathcal{H}_{2,\Omega}$ *approximation error than if the initialisation is done with the criterion* $\mathcal{J}_{init}^{(2)}$. *This is illustrated in Figure 9.4 where the frequency responses of the initial model is plotted together with the frequency responses of the two* 4*-th order models obtained with* ***DARPO*** *by using the two different initialisation strategies.*

The error between the initial model and the reduced-order model is lower when the non-intuitive pole selection approach given by $\mathcal{J}_{init}^{(2)}$ *is used.*

9.2 Numerical illustration

In Section 9.2.1, the behaviour of **DARPO** such as its convergence and the decrease of the error is illustrated, then, in Section 9.2.2, the gain that it can provide with respect to its initial point is evaluated. Finally in Section 9.2.3, it is used to demonstrate the process that can be followed for the control of large-scale model.

9.2.1 Illustration of the behaviour of DARPO

Reduction of the ISS model over $\Omega = [0, \infty)$

The examples presented in [Beattie and Gugercin, 2009] (see Section 4.2.3 for more information on the method developed in this article) are reproduced to illustrate the behaviour of **DARPO** and to show that its performances are similar to the optimal $\mathcal{H}_2$ approximation model presented in the paper when Ω is set to $[0, \infty)$. The examples are based on the first input to first output transfer of the ISS model.

The $r = 2$ **case.** In that case, the initial model is reduced to an order $r = 2$ in order to illustrate the convergence of the poles and residues of the reduced-order model. The poles and residues are initialised as

$$\hat{\lambda} = -4 \times 10^{-3} \pm j7.8 \times 10^{-1} \quad \text{and} \quad \hat{\phi} = 2.5 \times 10^{-4} \pm j2.2 \times 10^{-4},$$

which is similar to what is done in [Beattie and Gugercin, 2009].

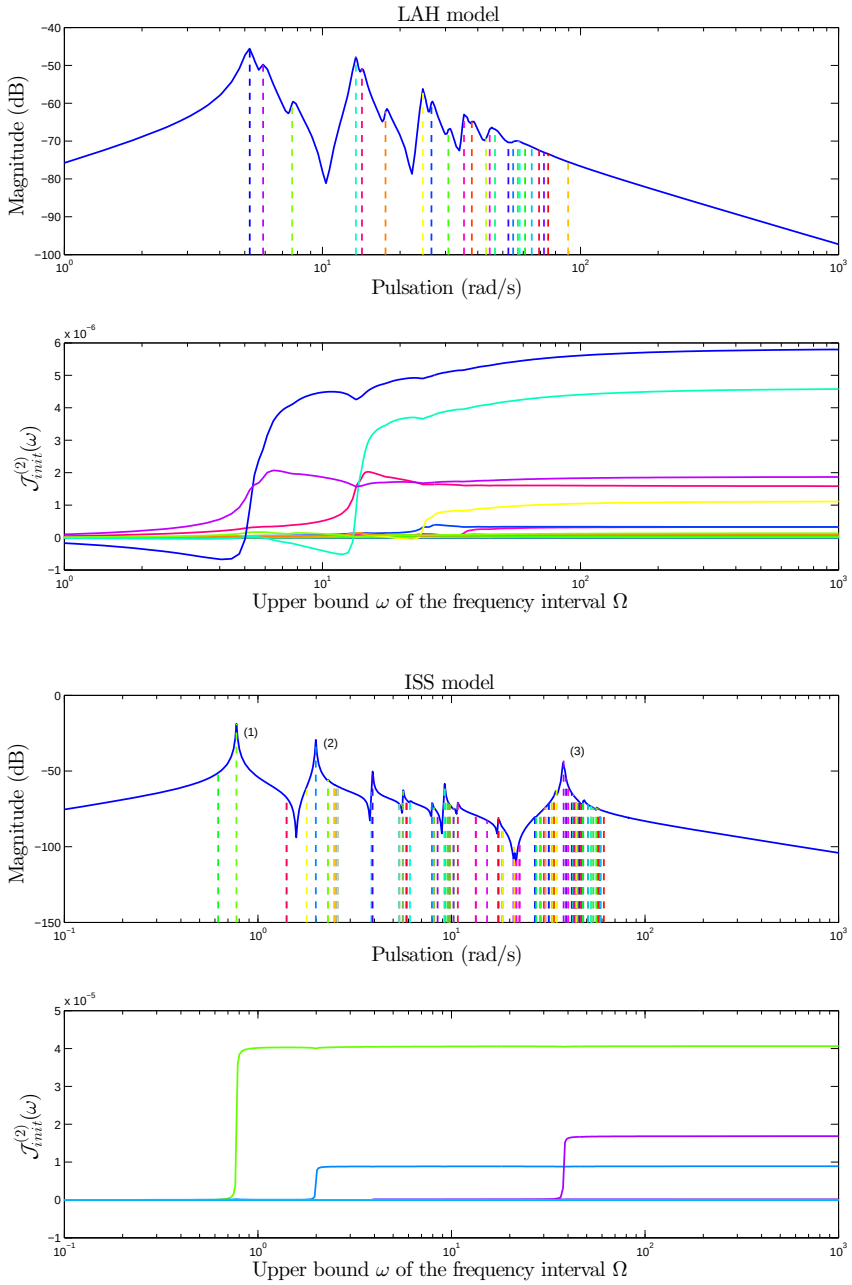


Figure 9.3: Value of the criterion $\mathcal{J}_{init}^{(2)}$ with respect to ω on the LAH model (top) and the ISS model (bottom).

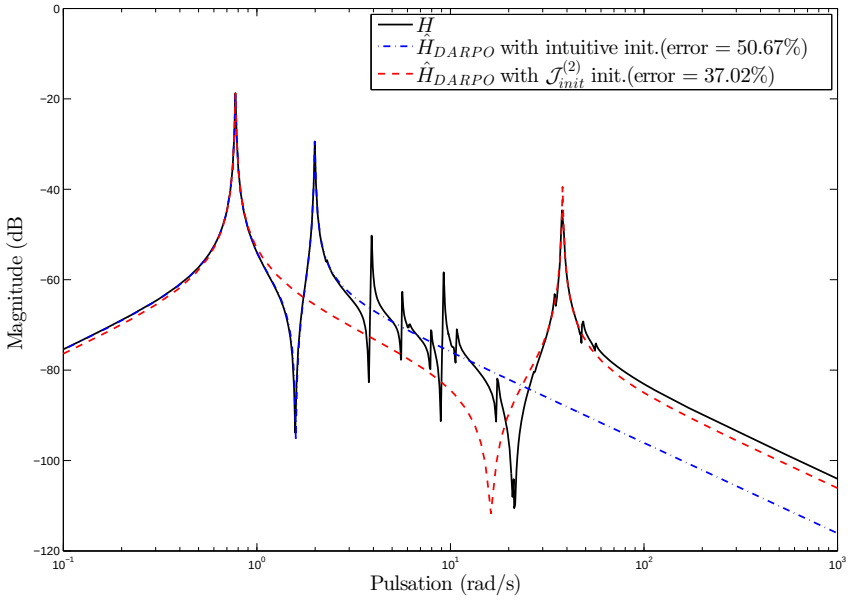


Figure 9.4: Comparison of the reduced-order models obtained with **DARPO** initialised intuitively or with the criterion $\mathcal{J}_{init}^{(2)}$ on the ISS model for $\Omega = [0, \infty)$ and $r = 4$.

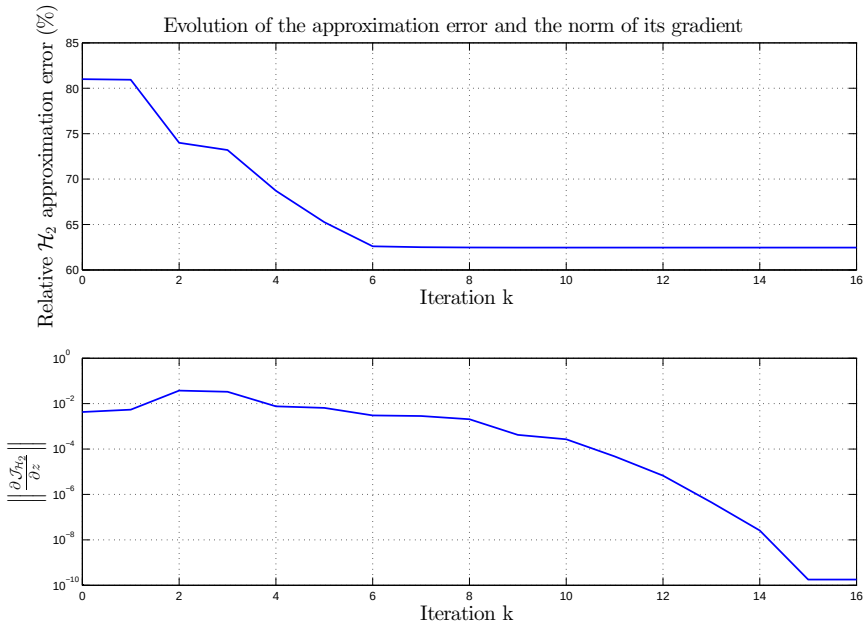
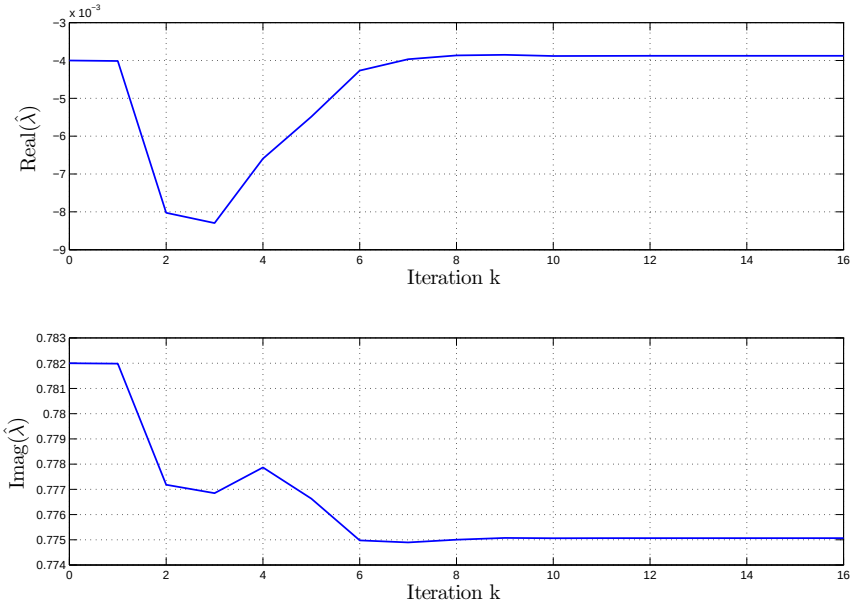
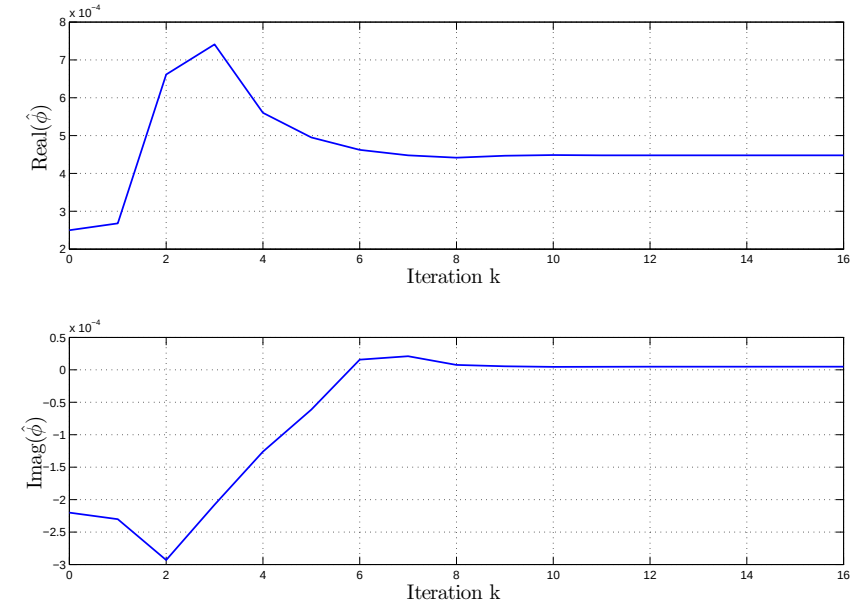


Figure 9.5: Evolution of the relative $\mathcal{H}_2$ approximation error (top) and of the norm of the gradient (bottom) for $r = 2$.



(a) Evolution of the pole $\hat{\lambda}$.



(b) Evolution of the residue $\hat{\phi}$.

Figure 9.6: Evolution of the pole (a) and residue (b) of the second-order approximation of the ISS.

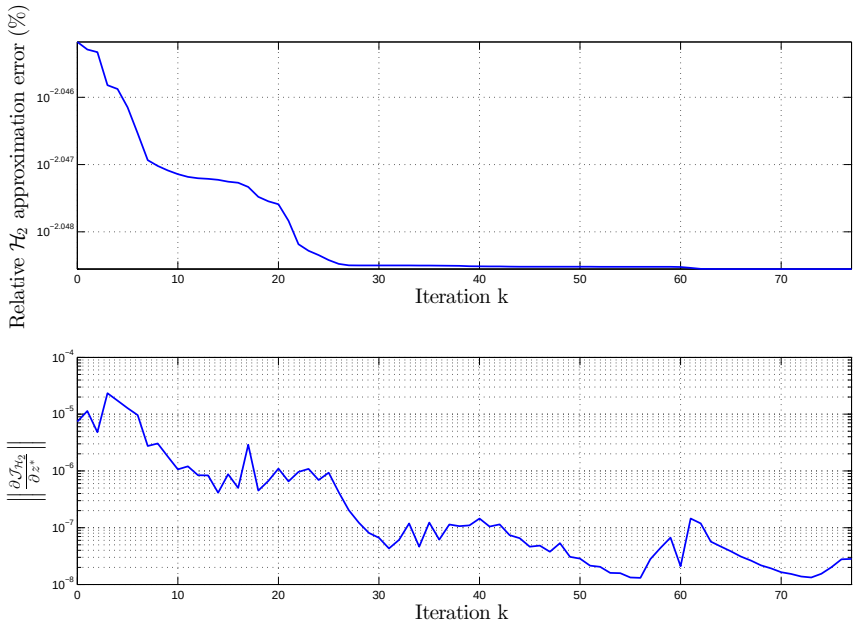


Figure 9.7: Evolution of the relative $\mathcal{H}_2$ approximation error (top) and of the norm of the gradient (bottom) for $r = 20$.

In Figure 9.5, the evolution of the relative $\mathcal{H}_2$ approximation error $\frac{\|H-\hat{H}\|_{\mathcal{H}_2}}{\|H\|_{\mathcal{H}_2}}$ is plotted together with the norm of the gradient. The evolution of the real and imaginary parts of the pole $\hat{\lambda}$ and the residue $\hat{\phi}$ are plotted in Figure 9.6a and Figure 9.6b, respectively.

The resulting reduced-order model is the same as in [Beattie and Gugercin, 2009]. One can observe that **DARPO** stops after 17 steps, but the approximation error, the poles and the residues do not evolve much after 10 iterations. In the article, the convergence is reached after ~ 5 steps, this difference comes from the method used to determine a new trial point. Indeed, in [Beattie and Gugercin, 2009], the authors use a trust-region method while **DARPO** is based on a line-search method. The former is more robust and has stronger convergence properties but obtaining a new trial point is more costly than with the latter approach [Nocedal and Yuan, 1998].

Moreover, the final value reached by the norm of the gradient is here $\sim 10^{-9}$ which can seem high. Nevertheless, the norm of the derivative of $\mathcal{J}_{\mathcal{H}_2}$ at the last iteration k in the descent direction $\mathbf{p_k}$ given by equation (9.6) is around 10^{-20}, lower than the machine precision. It means that at convergence, the objective function is almost flat in the direction given by $\mathbf{p_k}$. Hence, the line-search algorithm cannot find a better trial point in that direction.

The $r = 20$ case. In that case, the ISS model is reduced to an order $r = 20$. In [Beattie and Gugercin, 2009], a randomly generated initial point is used thus one cannot reproduce the same test. Here, **DARPO** is initialised by the modal truncation with the criterion $\mathcal{J}_{init}^{(2)}$ presented in Section 9.1.3. The relative $\mathcal{H}_2$ approximation error is reported in Figure 9.7 together with the norm of the gradient and, in Figure 9.8, the frequency responses of both models are plotted.

In Figure 9.7, one can see that the initial approximation error is already extremely low (0.9 %), hence **DARPO** barely decreases the error which finally falls to 0.89 %. The initial point was already almost an optimal choice. This is not surprising since the ISS model has only poorly damped modes, the modal approximation used as initialisation is already a good approximation. On this kind of models and with this initialisation strategy, only the residues are modified by

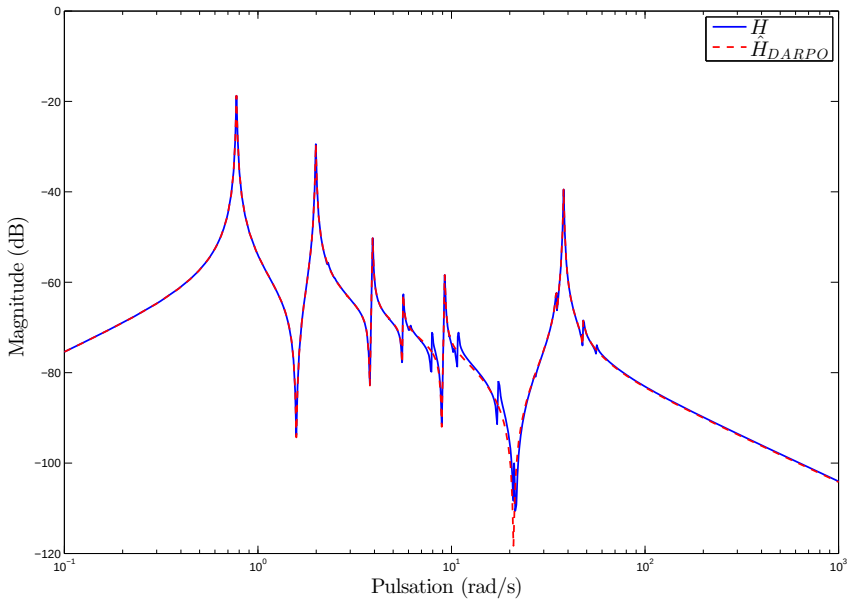


Figure 9.8: Frequency responses the reduced-order model obtained with **DARPO** ($r = 20$) and of the initial model.

DARPO and the poles barely moves.

One can observe in Figure 9.8 that the reduced-order model matches perfectly the initial model. Only some small dynamics around 10 rad/s are missed.

Evolution of the $\mathcal{H}_2$ and $\mathcal{H}_\infty$-norms of the error for varying approximation orders

In this example, the ISS model is reduced to several orders r going from 2 to 30 with **DARPO** over

- $\Omega = [0, \infty)$: in that case, the bounds $\Gamma(H-\hat{H})$ and $\bar{\Gamma}(H-\hat{H})$ are computed and compared to the real $\mathcal{H}_\infty$-norm of the error $\|H - \hat{H}\|_{\mathcal{H}_\infty}$ through the relative quantities

$$\frac{\Gamma(H-\hat{H})}{\|H\|_{\mathcal{H}_\infty}}, \quad \frac{\bar{\Gamma}(H-\hat{H})}{\|H\|_{\mathcal{H}_\infty}} \quad \text{and} \quad \frac{\|H-\hat{H}\|_{\mathcal{H}_\infty}}{\|H\|_{\mathcal{H}_\infty}}.$$

  The resulting values are plotted in Figure 9.9 (a). The relative $\mathcal{H}_2$ approximation error $\frac{\|H-\hat{H}\|_{\mathcal{H}_2}}{\|H\|_{\mathcal{H}_2}}$ is also plotted in Figure 9.9 (b).

- $\Omega = [0, 5]$: here the bounds $\Gamma_\Omega(H-\hat{H})$ and $\bar{\Gamma}_\Omega(H-\hat{H})$ are computed and the maximal singular value $\bar{\sigma}_\Omega$ of $H(j\omega)$ over Ω is measured. In Figure 9.9 (c), the relative quantities

$$\frac{\Gamma_\Omega(H-\hat{H})}{\bar{\sigma}_\Omega} \quad \text{and} \quad \frac{\bar{\Gamma}_\Omega(H-\hat{H})}{\bar{\sigma}_\Omega},$$

  are plotted. In Figure 9.9 (d), the relative $\mathcal{H}_{2,\Omega}$-norm of the error is also plotted.

One can observe in Figure 9.9 (b) and 9.9 (d) that in both cases, the $\mathcal{H}_2$ and $\mathcal{H}_{2,\Omega}$ norm of the error globally decrease as the approximation order r increases. Yet, the decrease is more regular

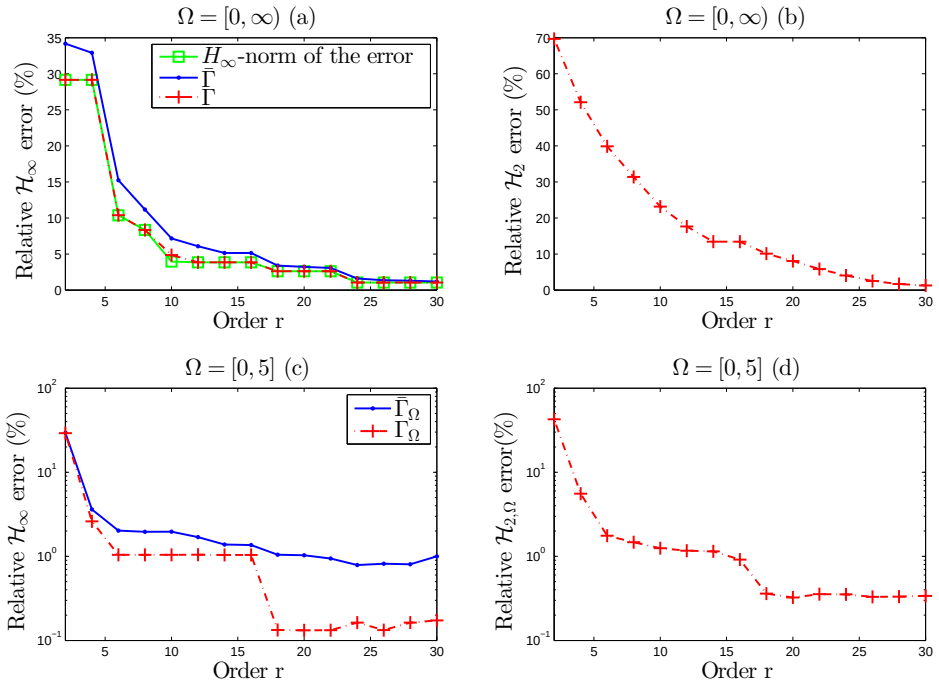


Figure 9.9: Evolution of the the $\mathcal{H}_{2,\Omega}$ approximation error and of Γ_Ω over $\Omega = [0, \infty)$ - (a) and (b) - and over $\Omega = [0, 5]$ - (c) and (d).

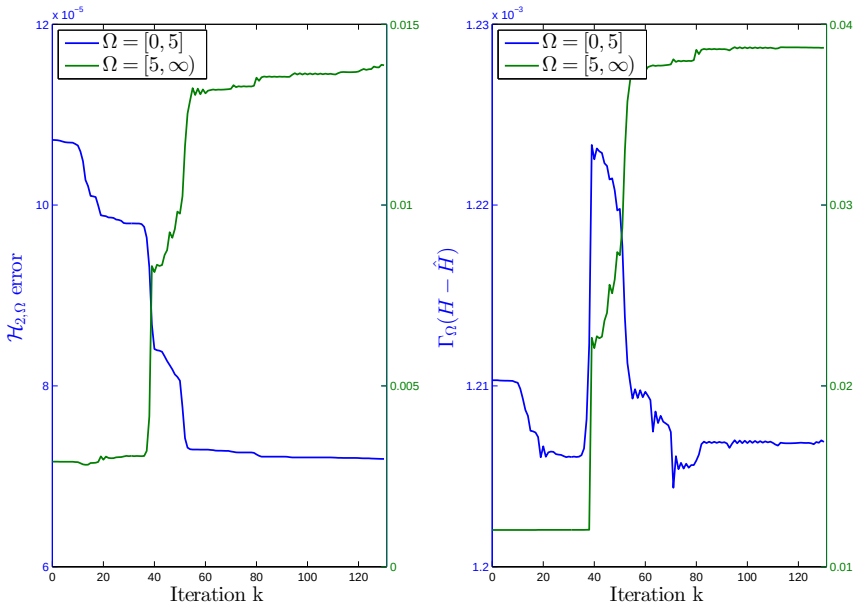


Figure 9.10: Evolution of the the $\mathcal{H}_{2,\Omega}$-norm of the approximation error and of $\Gamma_\Omega(H - \hat{H})$ over $[0, 5]$ and $[5, \infty)$.

LAH	g_{modal}	$g_{FLISTIA}$	g_{FLBT}
mean	21.88	7.11	7.72
max	99.82	96.97	100
min	2.17	0.12	0.084
std	27.91	16.76	16.90
ISS			
mean	4.99	0.95	0.67
max	58.69	99.99	100
min	7.61×10^{-4}	3.45×10^{-5}	1.08×10^{-5}
std	8.07	7.91	5.51
CBM			
mean	46.45	15.17	24.08
max	100	99.99	100
min	0.26	1.30×10^{-5}	2.55×10^{-5}
std	38.51	19.75	28.67

Table 9.1: Gain provided by the use **DARPO** with respect to the modal truncation, the **FL-ISTIA** and the **FL-BT** (in %).

in the $\mathcal{H}_2$ case than in the $\mathcal{H}_{2,\Omega}$ case. Indeed, one can for instance see in Figure 9.9 (d) that the approximation error barely decreases between $r = 20$ and $r = 25$ and even increases between $r = 20$ and $r = 22$. This comes from the initialisation process of **DARPO**. Indeed, for $r \geq 20$, the selection criterion $\mathcal{J}_{init}^{(2)}$ presented in Section 9.1.3 starts to select irrelevant poles outside the frequency interval which does not necessarily improve the approximation error inside the frequency interval Ω. In that case, a lower approximation error can for instance be obtained if **FL-ISTIA** is used to initialise **DARPO**.

In Figure 9.9 (a), one can see that in spite of the fact that the model is MIMO, the bound Γ is very close to the true $\mathcal{H}_\infty$-norm and that similarly to the $\mathcal{H}_2$-norm, the $\mathcal{H}_\infty$-norm of the approximation error decreases as the approximation order increases. In Figure 9.9 (c), one can observe that globally, $\bar{\Gamma}_\Omega$ is decreasing with the increase of the order and that Γ_Ω has an evolution similar to the $\mathcal{H}_{2,\Omega}$-norm of the approximation error.

Note that if the approximation error globally decreases both in mean value (through $\mathcal{J}_{\mathcal{H}_{2,\Omega}}$) and in maximum value (through Γ_Ω), this is not true outside the interval Ω. To highlight this point, let us consider $\Omega_2 = [5, \infty)$ and compute, for $r = 16$, $\|H - \hat{H}\|_{\mathcal{H}_{2,\Omega_2}}$ and $\Gamma_{\Omega_2}(H - \hat{H})$ at each iteration of **DARPO**. The $\mathcal{H}_{2,\Omega}$-norms of the approximation error over Ω and Ω_2 as well as the bounds Γ_Ω and Γ_{Ω_2} are plotted in Figure 9.10.

As expected the error $\|H - \hat{H}\|_{\mathcal{H}_{2,\Omega}}$ decreases in the interval Ω but the norm of the approximation over Ω_2 has an inverse evolution and increases during the iterations. The same increase can be observed with $\Gamma_{\Omega_2}(H - \hat{H})$. Over Ω, the upper bound $\Gamma_\Omega(H - \hat{H})$ is lower at the last iteration that at the first one but this could not have been the case given its evolution.

9.2.2 Improvement provided by DARPO

As mentioned earlier, **DARPO** is first of all a refinement process that improves, locally, the quality of the given initial reduced-order model. Hence, one must be cautious when comparing it. For instance, comparing **DARPO** initialised with a modal truncation based on the criterion $\mathcal{J}_{init}^{(2)}$ directly with other methods such as the **FL-ISTIA** only enables to say whether the modal truncation is close to a local minimum better than the model obtained with **FL-ISTIA**. Instead, here, one rather evaluates the improvement that **DARPO** can achieve given an initial reduced-order model obtained with another approximation method.

In particular, the LAH, the ISS and the CBM models are reduced for various approximation

orders and various frequency intervals (the same as in Section 6.2) with **DARPO** initialised with

- a modal truncation obtained with the criterion $\mathcal{J}_{init}^{(2)}$,
- the **modified FL-ISTIA** obtained with the same parameters as in Section 6.2,
- the **FL-BT**.

The improvement led by the use of **DARPO** to each stable initial point is measured through the gain[4]

$$g = \frac{e_0 - e_f}{e_0},$$

where e_0 is the initial $\mathcal{H}_{2,\Omega}$ approximation error, *i.e.* the one obtained with the initial approximation method used, and e_f is the final approximation error. Since **DARPO** is a descent method, necessarily $e_f \leq e_0$. The statistical indicators of the gain provided by **DARPO** are reported in Table 9.1 and its mean values for each approximation order r and for each upper bound ω of the frequency interval are plotted in Figures 9.11 and 9.12, respectively. For each couple of approximation order and frequency interval, the best initialisation point, *i.e.* the one that led **DARPO** to the lowest approximation error, is represented in Figure 9.13.

One can observe that the gain provided by the use of **DARPO** greatly varies between the models. Indeed, while it provides, in average, a great improvement for the approximation of the CBM model, the gain is not extremely significant with the ISS model. Similarly, depending on the initial method used for initialisation, the gain is more or less important. **DARPO** improves, in average, more the modal truncation than the **FL-ISTIA** or the **FL-BT**.

One can also see that with the **FL-ISTIA** and the **FL-BT**, the minimum value of g is equal to $\sim 10^{-5}\%$. This means that **DARPO** does not modify much the initial point and does not provide any significant improvement. It also means that the initial reduced-order model given by **FL-ISTIA** or the **FL-BT** is already located extremely close to a stationary point. From Figure 9.11, one can observe that, in average, the smallest values of gain are reached for small approximation orders r. This is not surprising since the search space is smaller for low values of r, hence there are less local minima. In general, the discrepancies between various model approximation methods appear for higher approximation orders.

From Figure 9.12, one can also see that the gain tends to decrease as the frequency interval widens. This is especially visible when the **FL-ISTIA** is used as initial method. This was expected since as ω increases, the $\mathcal{H}_{2,\Omega}$-norm tends towards the $\mathcal{H}_2$-norm and **FL-ISTIA** reduces to **ISRKA**, a sub-optimal $\mathcal{H}_2$-model approximation method (see Section 4.2.2), when $\omega \to \infty$. Hence, unless **DARPO** moves the poles of the reduced-order model, it is already optimal.

This suggests that **DARPO** is particularly interesting when the frequency interval is small and the approximation order large since in those cases, the other approximation methods are generally further from a local minimum.

Finally, from Figure 9.13, one can see that none of the initialisation method considered here stands really out of the two others. Indeed, while the **FL-BT** leads in general to better reduced-order model with the LAH model, it is not true anymore with the ISS and CBM models. Yet, this representation does not show that often, the errors obtained from the three initialisation strategies are actually extremely close.

[4]The cases for which the initial reduced-order model is unstable are discarded.

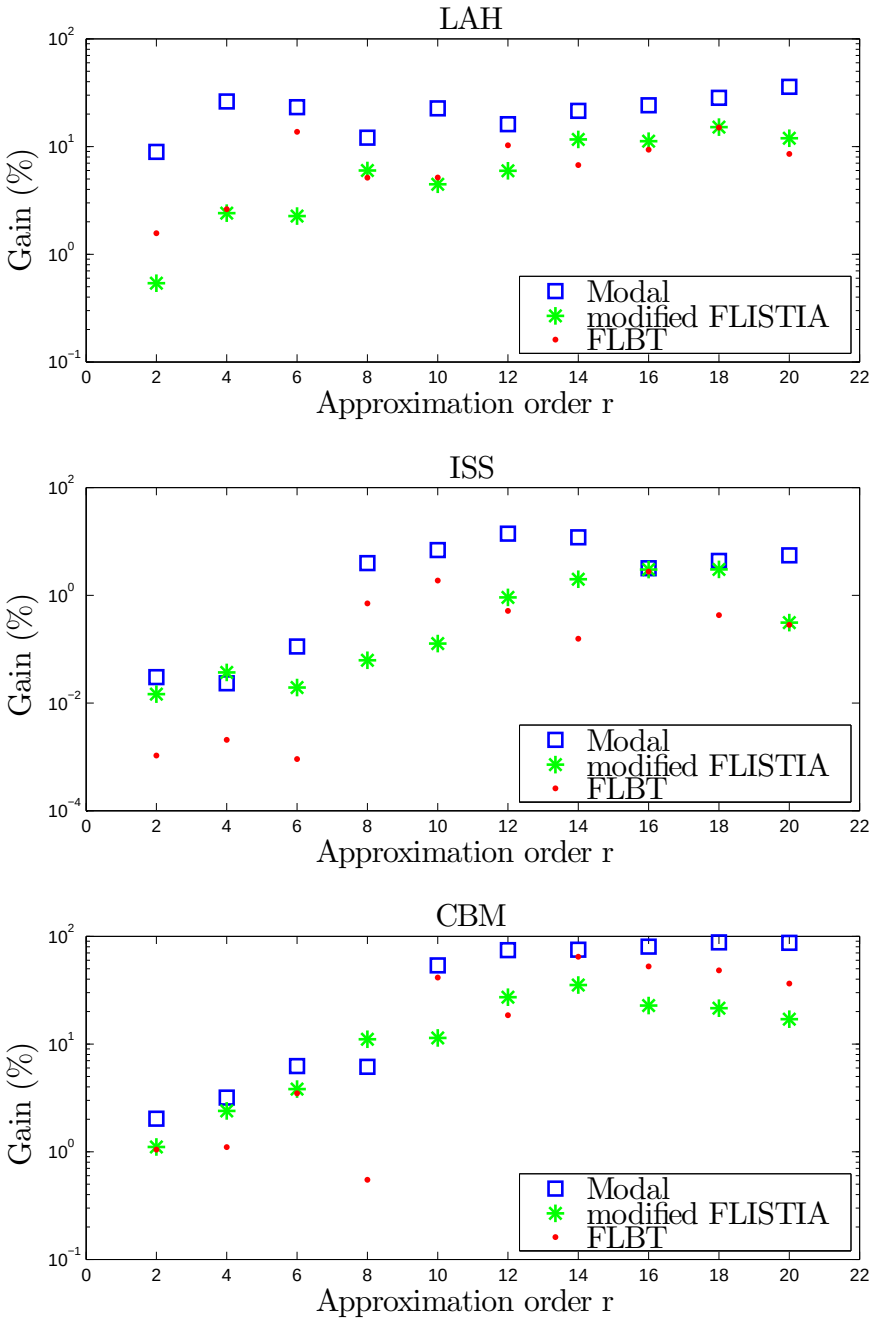


Figure 9.11: Mean gain provided by **DARPO** on the LAH (top), ISS (middle) and CBM (bottom) models for each approximation order r.

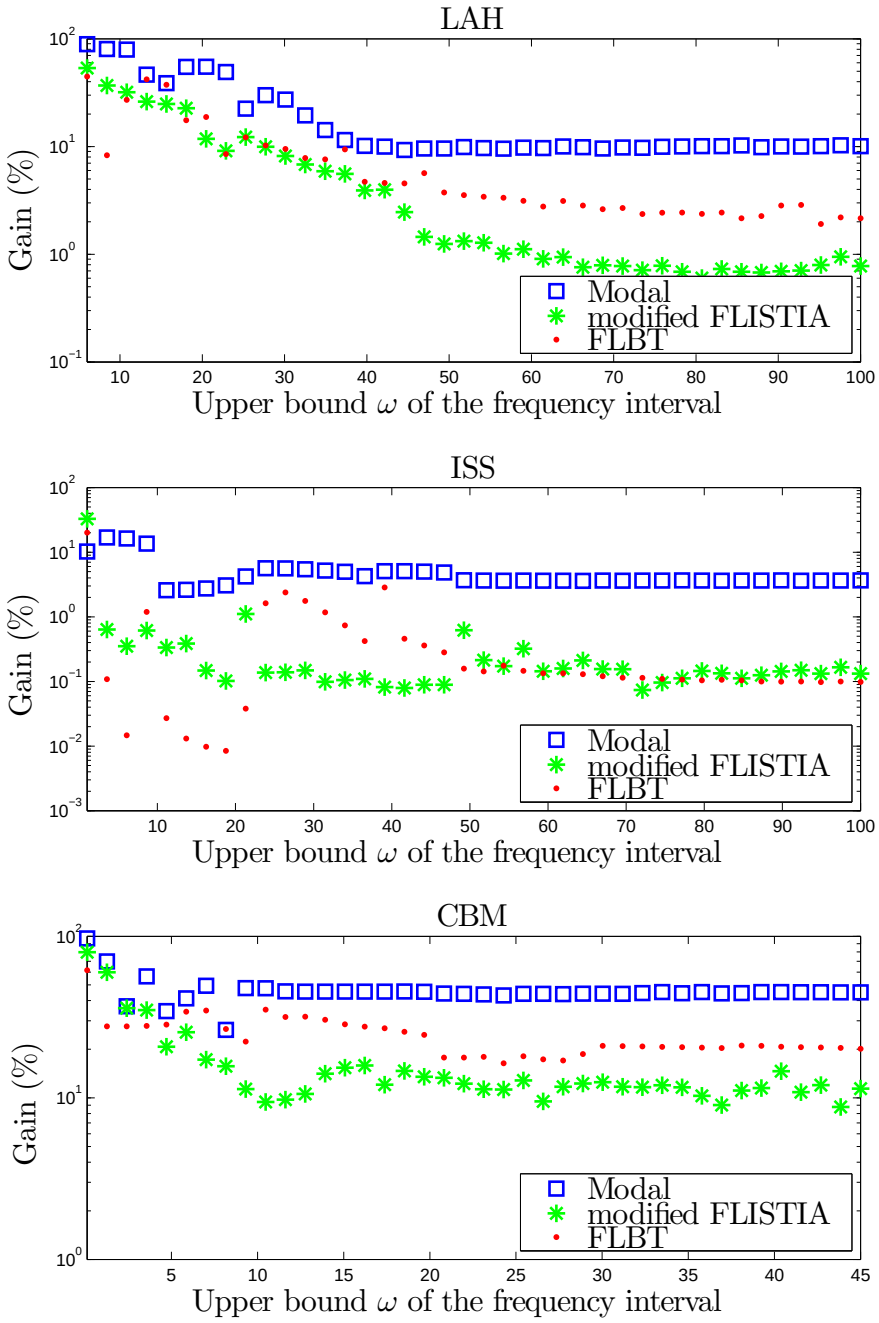


Figure 9.12: Mean gain provided by **DARPO** on the LAH (top), ISS (middle) and CBM (bottom) models for each upper bound ω of the frequency interval.

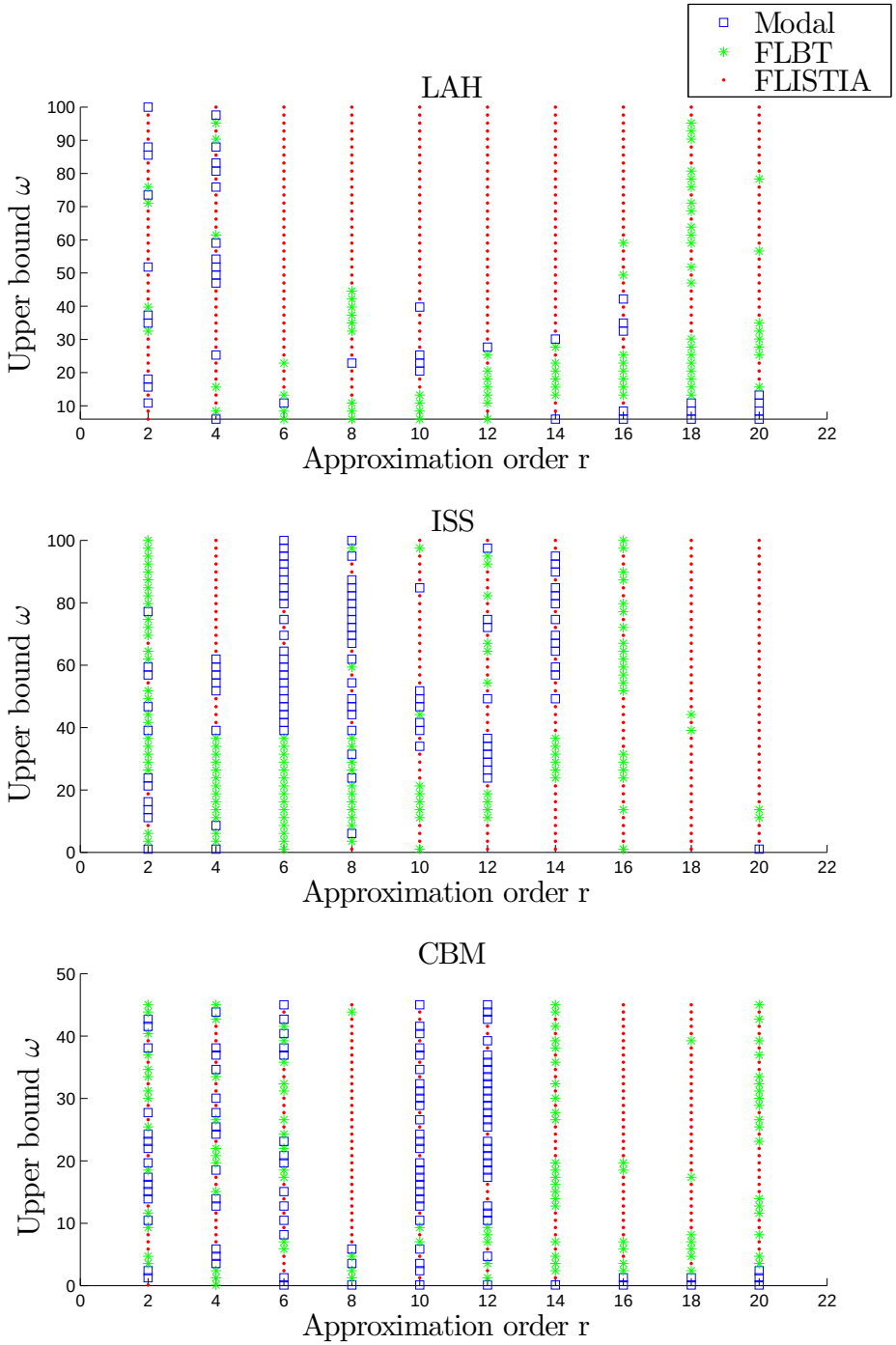


Figure 9.13: Best initialisation for **DARPO** among the modal truncation, the **FL-BT** and the **FL-ISTIA** on the LAH, ISS and CBM models for varying approximation orders and frequency intervals.

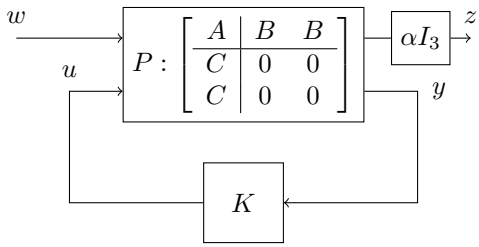


Figure 9.14: $\mathcal{F}_l(P, K)$

9.2.3 Application for the control of a large-scale model

In this Section, **DARPO** is used to demonstrate the process that can be followed for the design of controller on large-scale models. Note that this example is not a real industrial application and has been built up. Yet, the specifications have been chosen to be as representative as possible of what could be found in a real application. The code of this example is available in appendix C.8.

Description of the specifications

Initial control problem. Let us consider the ISS model $\mathbf{H} := (A, B, C)$, ($n = 270$, $n_y = 3$ and $n_u = 3$) which represents the transfer of a large flexible space structure with collocated actuators and rate sensors. The initial control objective is to design a controller K of order $n_k \ll n$ that performs an input disturbance rejection of $x_{dB} = 30$ dB. In other words, one would like to find the controller K such that the difference between the $\mathcal{H}_\infty$-norms of the open and closed loops is, at least, of x_{dB}, *i.e.*

$$20log_{10}\left(\|H\|_{\mathcal{H}_\infty}\right) - 20log_{10}\left(\|H_{BF}\|_{\mathcal{H}_\infty}\right) \geq x_{dB},$$

which can be rewritten as

$$\|H_{BF}\|_{\mathcal{H}_\infty}^{-1} \geq \|H\|_{\mathcal{H}_\infty}^{-1} 10^{x_{dB}/20} = \alpha, \tag{9.8}$$

where $\alpha = 272.88$ here. This disturbance rejection problem can be formulated as a standard $\mathcal{H}_\infty$ control problem aimed at finding K which solves the non-smooth optimisation problem,

$$K = \arg \min_{\substack{\tilde{K} \text{ stabilises } P \\ \tilde{K} \in \mathcal{K}}} \|\mathcal{F}_l(P, \tilde{K})\|_{\mathcal{H}_\infty},$$

where $\mathcal{F}_l(P, K)$ is the lower Linear Fractional Representation (LFR)[5] formed by interconnecting the plant $P(s)$ with the gain $K(s)$ as represented in Figure 9.14 and $\mathcal{K}$ represents the set of n_k-th order rational transfer matrices. Note that the $\mathcal{H}_\infty$-norm of the LFR $\mathcal{F}_l(P, K)$ is simply the $\mathcal{H}_\infty$-norm of the closed-loop H_{BF} from (9.8) scaled by α,

$$\|H_{BF}\|_{\mathcal{H}_\infty} = \frac{1}{\alpha}\|\mathcal{F}_l(P, K)\|_{\mathcal{H}_\infty}$$

hence the disturbance rejection specification (9.8) can be rewritten as

$$\|\mathcal{F}_l(P, K)\|_{\mathcal{H}_\infty} \leq 1. \tag{9.9}$$

Since the model $\mathbf{H}$ is positive, finding a controller K satisfying the rejection specification (9.9) is simple (see [Balakrishnan, 1996]). Indeed, any sufficiently large negative static feedback

[5] A very comprehensive description of the LFR and associated manipulation can be found in [Magni, 2006].

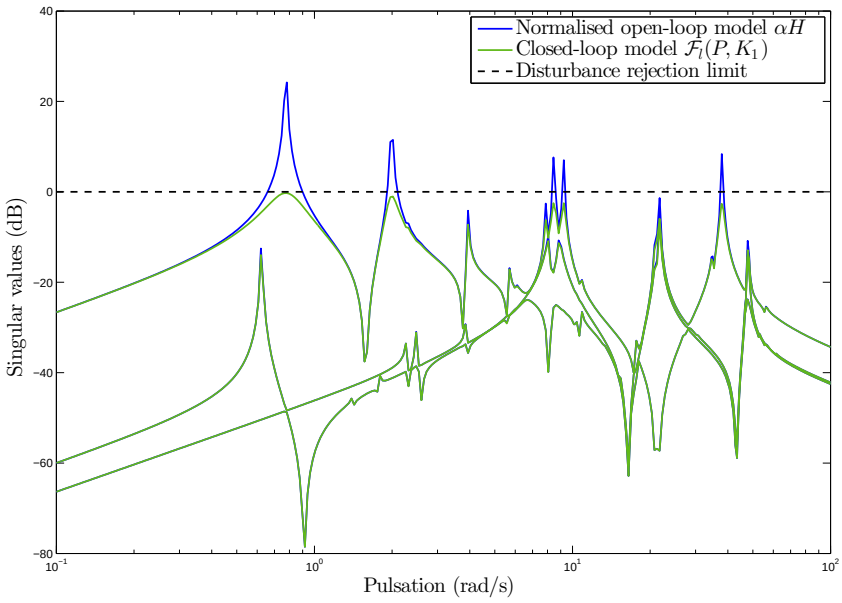


Figure 9.15: Singular values of the open-loop model **H** and of the closed-loop model $\mathcal{F}_l(P, K_1)$ obtained with the static output feedback $K_1(s)$.

is sufficient to damp all frequencies. For instance, let us consider the static feedback gain $K_1 = -\alpha I_3$, then the $\mathcal{H}_\infty$-norm of the lower LFR $\mathcal{F}_l(P, K_1)$ satisfies inequality (9.9), indeed

$$\|\mathcal{F}_l(P, K_1)\|_{\mathcal{H}_\infty} = 0.9693 < 1.$$

This is illustrated in Figure 9.15, where the singular values of the open-loop and closed-loop models **H** and $\mathcal{F}_l(P, K_1)$ are plotted.

Adding actuators. To make the control problem more interesting and more realistic, a dynamic is added to represent the actuators. Let us consider the actuator $A_c(s) = a_c(s)I_3$ where $a_c(s)$ is a damped second-order low-pass filter of bandwidth 20 rad/s, *i.e.*

$$a_c(s) = \frac{\omega_a^2}{s^2 + \xi_a \omega_a s + \omega_a^2},$$

with $\xi_a = 1.4$ and $\omega_a = 20$rad/s. Note that the actuator is built to have a static gain of 1, which means that the assumption is made that it does not disturb the low frequency and does not modify the $\mathcal{H}_\infty$-norm of the LFR.

The $\mathcal{H}_\infty$ control problem is consequently modified and consists now in finding the controller K which solves

$$K = \arg \min_{\substack{\tilde{K} \text{ stabilises } P \\ \tilde{K} \in \mathcal{K}}} \|\mathcal{F}_l^{(2)}(P, \tilde{K})\|_{\mathcal{H}_\infty}, \tag{9.10}$$

where $\mathcal{F}_l^{(2)}(P, K)$ is the LFR obtained by adding the actuator dynamics to $\mathcal{F}_l(P, K)$ as represented in Figure 9.16. One can easily checks that the former static output feedback K_1 is no longer sufficient when the actuators dynamic is taken into account, indeed

$$\|\mathcal{F}_l^{(2)}(P, K_1)\|_{\mathcal{H}_\infty} = 6.5079 > 1.$$

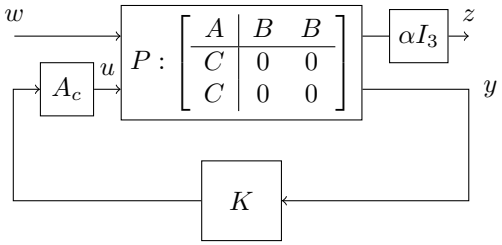


Figure 9.16: $\mathcal{F}_l^{(2)}(P, K)$

This can also be seen in Figure 9.17 (left) where the singular values of the open-loop and closed-loop models are plotted. More importantly, the closed-loop becomes unstable when the feedback gain increases, for instance one can observe in Figure 9.17 (right) that the closed-loop $\mathcal{F}_l^{(2)}(P, K_2)$ where $K_2 = 3K_1$ has unstable poles. Hence, a dynamic controller has to be found to fulfil the disturbance rejection when the actuator dynamic is taken into account.

Roll-off specification for the controller. In practice, it is common to restrain the dynamics of a controller in high frequency in order to *(i)* attenuate the transmission of measurement noise and *(ii)* to avoid to obtain the inverse of the actuator with a $\mathcal{H}_\infty$ synthesis[6]. This constraint can be translated in the $\mathcal{H}_\infty$ control framework by adding the optimisation objective $\|WK\|_{\mathcal{H}_\infty}$, where $W(s)$ is a roll-off filter, to the previous optimisation problem (9.10) which therefore becomes multiobjective,

$$K = \arg \min_{\substack{\tilde{K} \text{ stabilises } P \\ \tilde{K} \in \mathcal{K}}} \max\{\|\mathcal{F}_l^{(2)}(P, \tilde{K})\|_{\mathcal{H}_\infty}, \|W\tilde{K}\|_{\mathcal{H}_\infty}\}. \tag{9.11}$$

If the controller $K(s)$ is such that $\|WK\|_{\mathcal{H}_\infty} \leq 1$, then it means that its dynamics are upper bounded by the inverse of the roll-off filter $W(s)$, *i.e.*

$$\sigma_i(K(j\omega)) \leq \sigma_i(W^{-1}(j\omega)),\ \forall \omega \in \mathbb{R},\ i = 1, \ldots, 3.$$

In addition, $\|WK\|_{\mathcal{H}_\infty} \leq 1$ ensures the *strong stabilisation*, that is to say that the controller $K(s)$ is stable. This property is quite recommended in space applications from a practical implementation point of view.

Here the roll-off filter $W(s)$ is designed in the following way

- one would like a roll-off of 40db/decade on each axis, hence the roll-off filter $W(s)$ is chosen as a diagonal transfer matrix where each diagonal entry is a damped, invertible, second-order model.

- In addition, since one knows that α is the required gain to achieve the disturbance rejection specification without actuator, the controller must, at least, be allowed to reach this value for frequencies lower than 40rad/s where the last resonance of the model **H** is located. Hence, one wants that

  $$|W^{-1}(j\omega_{ro})| = \rho\alpha,\ \rho > 1,$$

  Here, ρ is set to 3 in order not to be too restrictive.

[6] Indeed, $K(s) = -\alpha A_c^{-1}(s)$ cancels the dynamics of the actuator and achieves the disturbance rejection. Yet in practice, it is not acceptable because the dynamic of the actuator is not perfectly known and noise is likely to be present on the output y.

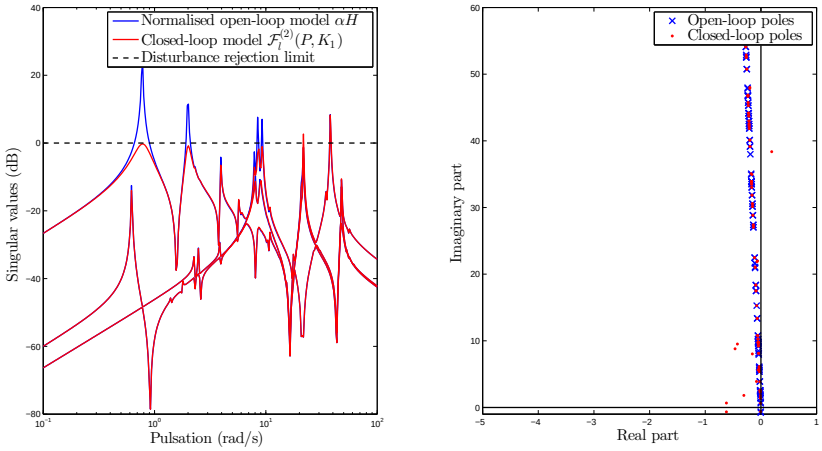


Figure 9.17: Singular values of the open-loop and closed-loop models $\mathbf{H}$ and $\mathcal{F}_l^{(2)}(P, K_1)$ (left) and poles of the open-loop and closed-loop models $\mathbf{H}$ and $\mathcal{F}_l^{(2)}(P, K_2)$ where $K_2 = 3K_1$ (right).

By taking into account these two points, the filter can be chosen for instance as

$$W(s) = \frac{1}{\rho\alpha} \frac{\frac{s^2}{\omega_{ro}^2}}{\frac{1}{(p\omega_{ro})^2}s^2 + \frac{\xi}{p\omega_{ro}}s + 1} I_3,$$

where $\rho = 3$, $\omega_{ro} = 40$ rad/s, $p = 300$, $\xi = 1.4$. Note that decreasing ρ or ω_{ro} makes the control problem harder since the inverse of $W(s)$ is then constraining the dynamics of the controller directly below 40rad/s where the last resonance of the model is located.

Note that since the filter $W(s)$ imposes a second-order roll-off on each one of the three axis, the order of the controller must at least be equal to 6.

Solving the $\mathcal{H}_\infty$ control problem

The multiobjective $\mathcal{H}_\infty$ control problem (9.11) can be solved using dedicated non-smooth optimisation tools [Apkarian and Noll, 2006; Gumussoy et al., 2009]. However, the size of the model $\mathbf{H}$ slows down the optimisation process and can prevent it from satisfying the constraints due to the presence of more local minima. Model approximation can be used to alleviate this issue.

Indeed, let us consider $\hat{\mathbf{H}}$, the r-th order approximation of $\mathbf{H}$ and $\mathbf{G}$, the frequency template built with the method proposed in Section 7.3.4 and whose transfer function $G(s)$ satisfies, for all $\omega \in \mathbb{R}$,

$$\|G(j\omega)\|_F \geq \left\|H(j\omega) - \hat{H}(j\omega)\right\|_F \geq \sigma_{max}(H(j\omega) - \hat{H}(j\omega)). \tag{9.12}$$

The frequency template $\mathbf{G}$ represents the error induced by the approximation and can be seen as an additive uncertainty on the reduced-order model $\hat{\mathbf{H}}$. This reduced-order uncertain model encompasses all the dynamics of the large-scale model.

By synthesising a controller $K(s)$ that ensures the *robust stability* of $\hat{\mathbf{H}}$ with respect to the uncertainty $\mathbf{G}$, then it is guaranteed to stabilise the large-scale model $\mathbf{H}$. Hence, a stabilising controller $K(s)$ for $\mathbf{H}$ can be found by solving the following $\mathcal{H}_\infty$ control problem

$$K = \arg \min_{\substack{\tilde{K} \text{ stabilises } \hat{P} \\ \tilde{K} \in \mathcal{K}}} \|\mathcal{F}_l^{(3)}(\hat{P}, \tilde{K})\|_{\mathcal{H}_\infty},$$

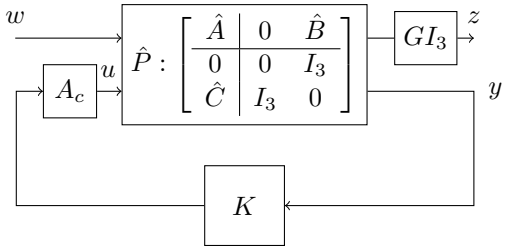


Figure 9.18: $\mathcal{F}_l^{(3)}(\hat{P}, K)$

where $\mathcal{F}_l^3(\hat{P}, K)$ is the low-order LFR formed by interconnecting the reduced-order plant $\hat{P}$ with $K(s)$ and $A_c(s)$ and by adding the uncertainty $G(s)$ in output of the performance channel as represented in Figure 9.18. Note that the method to build the frequency template $\mathbf{G}$ proposed in Section 7.3.4 produces a SISO model while the model here is MIMO. That is why one has to create the MIMO transfer matrix $G(s)I_3$.

Hence, instead of (9.11), one rather wants to find $K(s)$ that solves

$$K = \min_{\substack{\tilde{K} \text{ stabilises } \hat{P} \\ \tilde{K} \in \mathcal{K}}} \max\{\|\mathcal{F}_l^{(2)}(\hat{P}, \tilde{K})\|_{\mathcal{H}_\infty}, \|W\tilde{K}\|_{\mathcal{H}_\infty}, \|\mathcal{F}_l^{(3)}(\hat{P}, \tilde{K})\|_{\mathcal{H}_\infty}\}. \tag{9.13}$$

The differences between the $\mathcal{H}_\infty$ control problem (9.11) and (9.13) are

(i) the new $\mathcal{H}_\infty$ problem involves only low-order models while the problem (9.11) involves the initial large-scale plant P,

(ii) the optimisation channel $\|\mathcal{F}_l^{(3)}(\hat{P}, \tilde{K})\|_{\mathcal{H}_\infty}$ in (9.13) ensures that the controller $K(s)$ stabilises the *large-scale* plant P by synthesising a robust controller on the low-order uncertain model,

(iii) the performance channel $\|\mathcal{F}_l^{(2)}(\hat{P}, \tilde{K})\|_{\mathcal{H}_\infty}$ guarantees that the controller $K(s)$ ensures the disturbance rejection *on the low-order LFR* $\mathcal{F}_l^{(2)}(\hat{P}, \tilde{K})$ but it is not necessarily true on the large-scale one $\mathcal{F}_l^{(2)}(P, \tilde{K})$.

The last point implies that the $\mathcal{H}_\infty$-norm of $\mathcal{F}_l^{(2)}(P, \tilde{K})$ must be recomputed afterwards to check if it is inferior to 1 (see Remark 21). With a sufficiently accurate reduced-order model $\hat{\mathbf{H}}$, one expects the constraint to be satisfied though.

Remark 21 (Robust performance)**.** *In order to guarantee that the disturbance rejection is satisfied on the large-scale model, the controller $K(s)$ should be synthesised using μ-synthesis to ensure the* robust performance *of the controller with respect to the approximation error.*

The available methods for μ-synthesis lead to controllers of the same dimension as the plant which is not necessarily interesting. One way to address the issue here is to perform a $\mathcal{H}_\infty$ synthesis based on the $\mathcal{H}_\infty$ control problem represented in Figure 9.20. However, with this approach, only sufficient conditions can be fulfilled, hence the controller might be conservative. Here, with the specifications considered in the control problem, no satisfactory solution is found with this approach.

In this example, the $\mathcal{H}_\infty$ control problem (9.11) on the large-scale plant P has not led to a solution which satisfies all the constraints ($\gamma = 2.8$). In particular, the template imposed by the roll-off filter has not been satisfied. By relaxing the latter, *i.e.* by increasing ρ, a satisfying controller can be found.

For the problem (9.13), the large-scale model $\mathbf{H}$ has been approximated to an order 30 with **DARPO** over $\Omega = [0, 40]$ and the frequency template has an order 2. The singular values of the

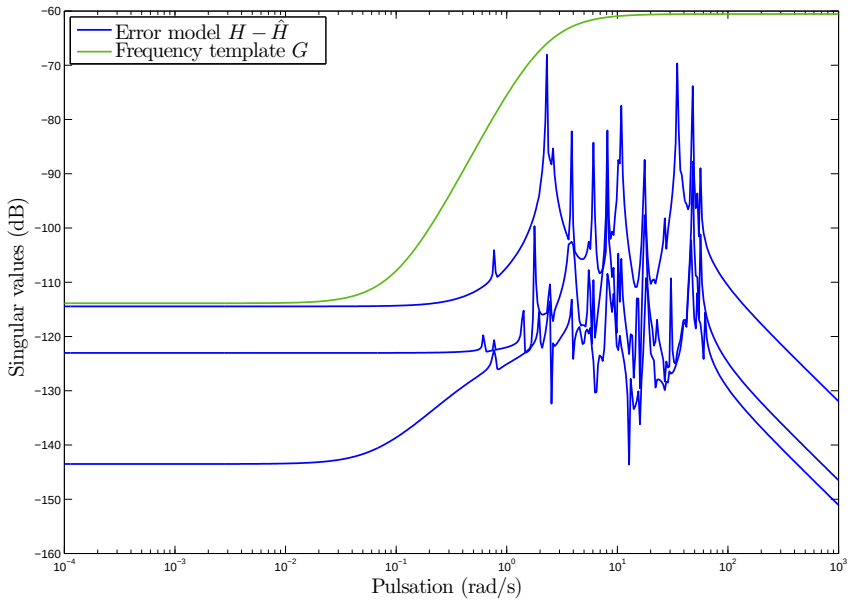


Figure 9.19: Singular values of the approximation error $\mathbf{H}-\hat{\mathbf{H}}$ and of the corresponding frequency template $\mathbf{G}$.

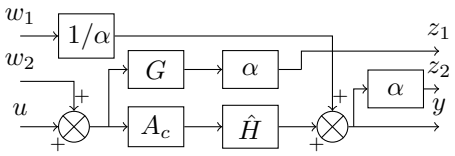


Figure 9.20: $\mathcal{H}_\infty$ control problem for robust performance.

approximation error $\mathbf{H} - \hat{\mathbf{H}}$ and of the frequency template $\mathbf{G}$ are plotted in Figure 9.19 where one can observe that the constraint (9.12) is indeed satisfied.

On this low-order $\mathcal{H}_\infty$ problem, a controller $K(s)$ is found such that

$$\max\{\|\mathcal{F}_l^{(2)}(\hat{P}, K)\|_{\mathcal{H}_\infty}, \|WK\|_{\mathcal{H}_\infty}, \|\mathcal{F}_l^{(3)}(\hat{P}, K)\|_{\mathcal{H}_\infty}\} = 0.9086 < 1,$$

which means that the robust stability, the roll-off specification and the disturbance rejection on the low-order plant are all satisfied. To verify that the disturbance rejection is also satisfied on the large plant P, one must compute the $\mathcal{H}_\infty$-norm of the large LFR, *i.e.*

$$\|\mathcal{F}_l^{(2)}(P, K)\|_{\mathcal{H}_\infty} = 0.9097 < 1.$$

The disturbance rejection of 30dB can be seen in Figure 9.21 where the singular values of the open-loop, low-order closed-loop and large-scale closed-loop are plotted. The difference between the low-order (green) and the large-scale (red) closed-loops is barely noticeable here. One can also see that the synthesised controller $K(s)$ satisfies the constraint imposed by the roll-off filter $W(s)$ in Figure 9.22.

Note also that by approximating the large-scale model $\mathbf{H}$ over $\Omega = [0, \infty)$, no satisfying controller is found. This is certainly due to the fact that the 30-th order approximation over $[0, \infty)$ misses a low frequency resonance on the second input to second output transfer while the reduced-order model obtained by approximation over $[0, 40]$ catches it.

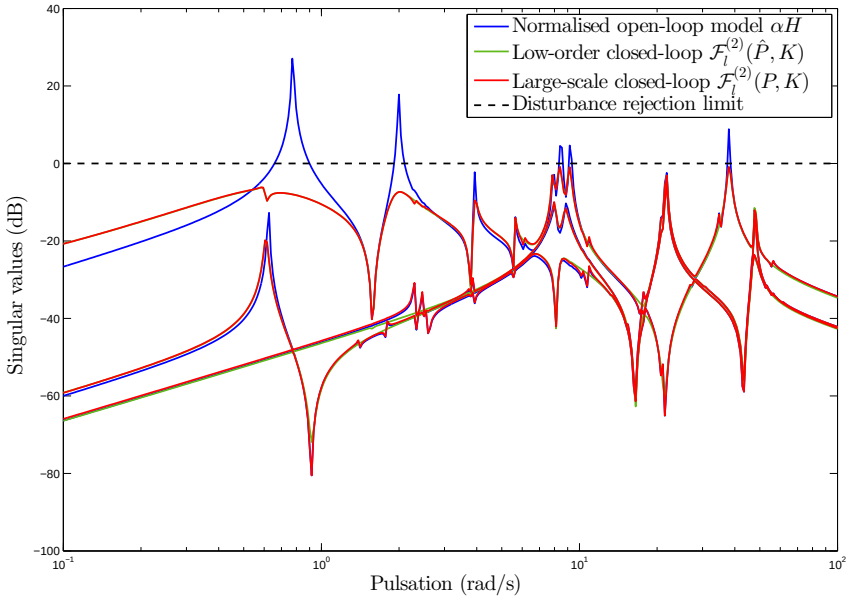


Figure 9.21: Singular values of the normalised open-loop model $\mathbf{H}$, of the low-order closed loop $\mathcal{F}_l^{(2)}(\hat{P}, K)$ and of the large-scale closed-loop $\mathcal{F}_l^{(2)}(P, K)$.

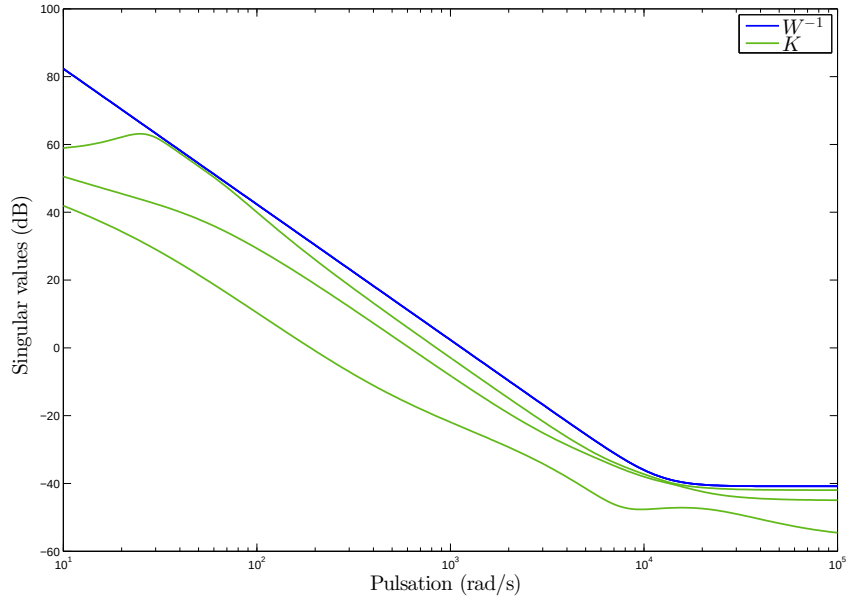


Figure 9.22: Singular values of the 6-th order controller $K(s)$ synthesised on problem (9.13) and of the inverse of the roll-off filter $W^{-1}(s)$.

Conclusion

The poles-residues formulation of the $\mathcal{H}_{2,\Omega}$ approximation error between a large-scale model and a reduced-order one together with its gradient, developed in Chapter 8, have been used here in a complex-domain, quasi-Newton optimisation algorithm. This optimisation algorithm relies on the BFGS update of the approximate inverse Hessian and on an inexact line-search method satisfying the strong Wolfe conditions. The resulting approximation algorithm **DARPO** produces a sequence of reduced-order model which yields a monotonically decreasing $\mathcal{H}_{2,\Omega}$ approximation error with the initial large-scale model and is expected to converge towards a local minimum.

DARPO is a local optimisation procedure and is therefore particularly dependent of the initial point it is given. The modal truncation is the most natural initial reduced-order model that can be used but other frequency-limited methods can also be used. Depending of the proximity of the initial reduced-order model to a local minimum, **DARPO** can bring substantial improvement with respect to the initial point.

DARPO has also been used in a simple process for designing a stabilising controller for a large-scale model. It consists in building a frequency template upper bounding the approximation error with the method proposed in Section 7.3.4 and to consider it as an additive uncertainty on the reduced-order model. A robust controller can then be designed on this low-order uncertain model and is guaranteed to stabilise the initial large-scale model.

In the following Chapter, **DARPO** is used together with **FL-ISTIA** on an industrial application as part of a global process used to design control laws on large-scale aircraft models.

Chapter 10

Industrial aeronautical use case

The study detailed in this chapter has been motivated by the increasing dimension of the models representing civilian aircraft which prevents standard control design methods from being directly used. An approximation step is mandatory and this step must not weaken the robustness and performance properties of the controllers that are designed. Indeed, in the aeronautical context, strong constraints can exist on the control laws in order to satisfy safety specifications. Several steps have been taken towards this goal,

- in a first time, the performances of several model approximation methods, including **FL-ISTIA**, have been evaluated on the large-scale models representing the longitudinal behaviour of a long-range flexible civilian aircraft in [Vuillemin et al., 2012a, 2013b],
- then **FL-ISTIA** and **DARPO** have been used to enable the design of *(i)* an anti-vibration control law for a business jet aircraft [Poussot-Vassal et al., 2013] and *(ii)* a control law ensuring flight performance and load clearance in presence of input saturation on a longitudinal aircraft model [Burlion et al., 2014].
- Model approximation has also been used as the starting point in a modelling process aimed at creating a low-order uncertain, parameter varying model from a set of large-scale dynamical models [Poussot-Vassal and Vuillemin, 2013; Poussot-Vassal et al., 2014],
- thanks to this process, the global stability and performance of the set of controlled large-scale models representing the business jet aircraft subject to actuator saturation have been proven by taking into account the error induced by the approximation and the interpolation of the initial large-scale models [Vuillemin et al., 2014a].

In this Chapter, the global process used for the design and validation of an anti-vibration control law for a business jet aircraft is detailed. In particular, in Section 10.1, the preliminary approach used to design an anti-vibration control law for one model representing a business jet aircraft at one single flight operating condition is described. The stability of the controller obtained from the generalisation of this approach to the set of models representing the aircraft at different Mach-numbers is then assessed in Section 10.2 in presence of input saturation.

Note that this process involves way more methods and tools than model approximation only which are out of the scope of this study. Hence all the steps cannot be covered in depth, but adequate references are pointed out when required.

Contents

10.1 Vibration control for one business jet aircraft model 160
10.1.1 General industrial framework 160
10.1.2 Problem formulation and business jet aircraft model approximation . 161
10.1.3 Anti-vibration control design 162
10.1.4 Numerical results 166
10.2 Global stability validation of a parameter varying business jet aircraft 167
10.2.1 Problem statement 167
10.2.2 Model approximation & error bounding 168
10.2.3 Low-order LFR model creation 170
10.2.4 IQC-based global stability analysis 173

10.1 Vibration control for one business jet aircraft model

The motivations and general industrial framework of this study are described in Section 10.1.1. The general control problem is then formulated in Section 10.1.2 and addressed in Section 10.1.3. Finally, the numerical results are presented in Section 10.1.4.

10.1.1 General industrial framework

In aeronautics, engineers have often to cope with many technical and practical problems for the design of a control laws caused by the complexity of the underlying models which are referred as *aeroservoelastic models*. Such models take into account

- the physic involved in the aeroelastic phenomena such as the structural loads (*e.g.* manoeuvres) or the aerodynamics loads (*e.g.* gusts),
- the flight control system behaviour, *i.e.* the dynamics of the actuators (aerodynamic control surfaces) and the sensors (*e.g.* accelerometers and gyrometers), eventual measurements delays, etc.
- uncertainties concerning the characteristics and state of the aircraft such as its mass (which varies during the flight due to the tanks filling), its speed, its flight altitude, etc.

For industrial civilian aircraft, these models are generally composed by a set of large-scale models representing the aircraft at different flight operating conditions.

Based on these aeroservoelastic models, *nominal flight control laws* are designed such that *(i) the robustness* (stability and performances) of the aircraft over the whole parametric domain (speed, altitude, etc.) is guaranteed, *(ii) flying quality* specifications (*i.e.* fast and smooth response to pilot commands and rejection of gust disturbances) are satisfied.

In addition, *comfort control laws*[1] are designed in order to attenuate the undesirable vibrations which are likely to appear due to aerodynamics disturbances. For instance, disturbances on the rear part of the aircraft can generate vibrations in the cockpit or in the cabin. These *anti-vibration* control laws must also satisfy some constraints, in particular *(i)* they are required to reduce the vibrations to a specified level, *(ii)* they must have a negligible effect on the nominal control law and *(iii)* they must be of low complexity.

Here, the preliminary study for the design of such an anti-vibration control law for an industrial business jet aircraft (denoted BizJet in the sequel), in the context of a collaboration between Onera and Dassault-Aviation, is presented. More specifically, and with reference to Figure 10.1, the objective is to design an anti-vibration control law (in red) for a large-scale BizJet model controlled by its nominal flight control law (in blue) which respects the specifications mentioned above. This objective is challenging for several reasons,

- the large dimension of the model ($n \approx 600$) prevents from applying directly standard control tools,
- the undesirable vibrations should be attenuated over a specific frequency range only (related to passenger and pilot comfort specifications),
- the nominal flight control law performances must be preserved and high frequency lightly damped modes must remain sufficiently stable,
- the anti-vibration control is based on the deflection of existing control surfaces which means that no dedicated actuators are available to achieve the objective.

[1] Actually, these control-laws also enable to decrease the loads on the structure of the aircraft.

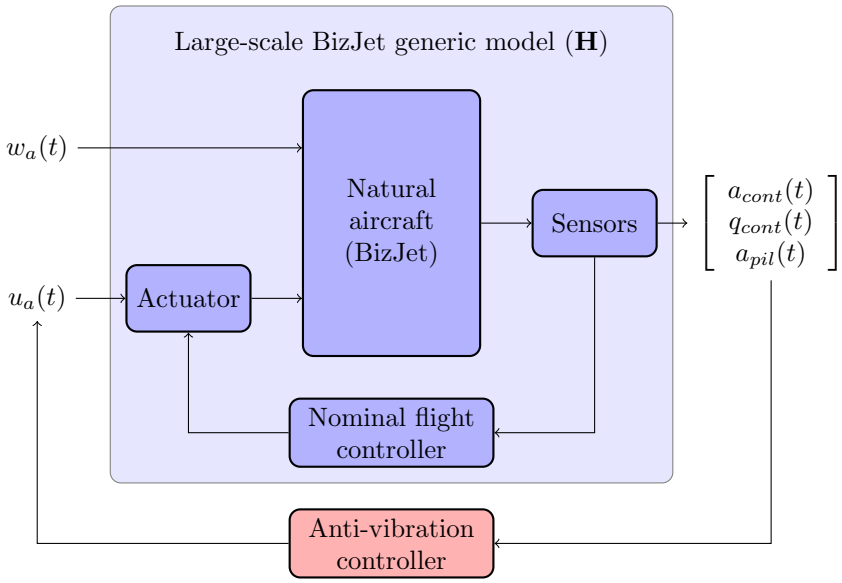


Figure 10.1: BizJet model (**H**) and anti-vibration controller.

The approach used to achieve that objective is in three steps *(i)* first, the large-scale model is reduced, *(ii)* then, a general control problem is formulated by using suitable weighting filters which capture the frequency-limited nature of the objectives, *(iii)* finally, a structured controller based on non-smooth $\mathcal{H}_\infty$ synthesis is designed. At this stage, the approximation step does not offer any kind of guarantee for the controller designed on the reduced-order model, hence, the anti-vibration control law must be validated experimentally on the large-scale BizJet model. This issue is alleviated in Section 10.2 by a rigorous global stability analysis.

10.1.2 Problem formulation and business jet aircraft model approximation

Control problem formulation

The initial BizJet aircraft model is of order $n = 597$, has $n_u = 2$ inputs,

- $w_a(t)$, the external disturbance input representing the aerodynamic disturbance affecting the rear part of the aircraft,
- $u_a(t)$, the anti-vibration control output, *i.e.* the input of the elevator on the rear part,

and $n_y = 3$ outputs,

- $a_{cont}(t)$, the vertical acceleration output, used by the nominal flight controller,
- $q_{cont}(t)$, the pitch rate of the aircraft, used by the nominal flight controller,
- $a_{pil}(t)$, the acceleration sensor located close to the pilot's seat, which has to be monitored for comfort issues.

Let us denote by $T_{w_a \to a_{pil}}$ and $T_{w_a \to a_{cont}}$ the transfers, in the nominal case (*i.e.* without the anti-vibration control law) from the disturbance input to the acceleration of the pilot and to the vertical acceleration, respectively. An let us denote by $\mathcal{F}_l(H,K)_{w_a \to a_{pil}}$ and $\mathcal{F}_l(H,K)_{w_a \to a_{cont}}$ the same transfers when the anti-vibration loop is closed with the controller K.

Let us reformulate the objectives and constraints related to the desired controller,

- **Objectives :** the controller should reduce the pilot acceleration response $a_{pil}(t)$ to an aerodynamic disturbance $w_a(t)$ over $\Omega_1 = [\omega_1\ \omega_2]$ while keeping this transfer as unchanged as possible in $\Omega_2 = [0, \omega 1] \bigcup [\omega_2, \omega_3]$, $(\omega_1 < \omega_2 < \omega_3)$.

- **Constraints :** the controller must not modify the transfer $T_{w_a \to a_{cont}}$ of more than 20% of its nominal value for frequencies inferior to ω_c and it must be asymptotically stable (strong stability). Additionally, the controller should be of the lowest achievable complexity[2].

These objectives and constraints are stated more formally by using the frequency-limited $\mathcal{H}_2$-norm as the problem of finding the controller K which solves

$$\begin{array}{rcll} K & = & \arg\min \|\mathcal{F}_l(H, \tilde{K})_{w_a \to a_{pil}}\|_{\mathcal{H}_{2,\Omega_1}} & \\ & & \text{s.t.} & \\ & & (i) \quad \max_{\omega \in [0, \omega_c]} \left| \frac{\mathcal{F}_l(H,\tilde{K})_{w_a \to a_{cont}}(j\omega) - T_{w_a \to a_{cont}}(j\omega)}{T_{w_a \to a_{cont}}(j\omega)} \right| < 0.2 & \\ & & (ii) \quad \max_i \mathbf{Re}\left(\lambda_i(\tilde{K})\right) < 0 & \\ & & (iii) \quad \|\mathcal{F}_l(H, \tilde{K})_{w_a \to a_{pil}} - T_{w_a \to a_{pil}}\|_{\mathcal{H}_{2,\Omega_2}} \leq \beta & \end{array} \tag{10.1}$$

While the constraints *(i)* and *(ii)* are required, the third one is looser, that is why *(iii)* has been mentioned as an objective before, *i.e.* one would like the closed-loop transfer from w_a to a_{pil} to remain as unchanged as possible over Ω_2.

Note that to the author's knowledge, there is no tool to solve a $\mathcal{H}_{2,\Omega}$ control problem, hence in the sequel, a $\mathcal{H}_\infty$ control problem which mimics problem (10.1) is built and addressed instead. To alleviate the computational complexity induced by the large dimension of the aircraft model, a preliminary approximation step is performed.

Model approximation

To chose a suitable reduced-order model, the relative approximation errors obtained by approximating it with the sub-optimal $\mathcal{H}_2$ model approximation method **ISRKA** (see Section 4.2.2) and the **BT** (see Section 3.1.2) are computed for varying approximation orders r going from 2 to 50 and reported in Figure 10.2.

The first observation that one can make is that the **ISRKA** leads generally to a better reduced-order model than the **BT** on this model. Yet, what is really important to notice here is the relative errors achieved by the approximation methods. For $r = 2$, the relative approximation errors are between 80% and 100% and fall below 10% only for $r = 50$. The relative errors are significantly higher than what could be achieved on the simpler benchmarks considered in the other chapters of this thesis. It highlights the complexity of the underlying model which cannot be reduced efficiently with a very low order model.

In the sequel, the large-scale model $\mathbf{H}$ is replaced by the 50-th order approximation $\hat{\mathbf{H}}$ obtained with the **ISRKA**.

10.1.3 Anti-vibration control design

To design anti-vibration controllers, engineers generally rely on their expertise of the physical system. Yet, due to the complexity of the design problems, this approach often requires several trials and errors. Here, a more generic approach based on the reformulation of the problem into a $\mathcal{H}_\infty$ optimisation one is considered.

[2]Limiting the complexity of the controller enables to simplify the implementation process and the eventual re-tuning performed during flight tests.

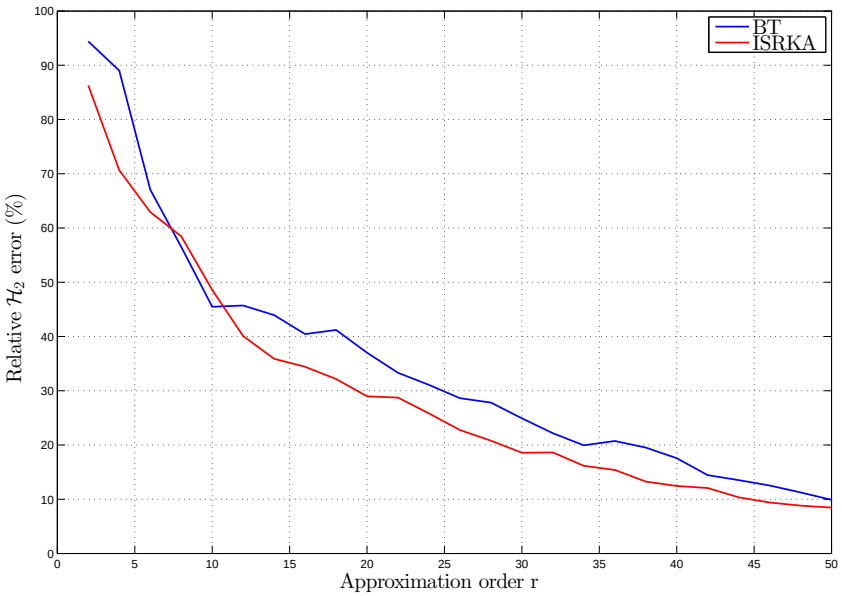


Figure 10.2: Relative $\mathcal{H}_2$ approximation error (in %) between the large-scale BizJet aircraft model and its approximation for varying orders r obtained with the **ISRKA** and the **BT**.

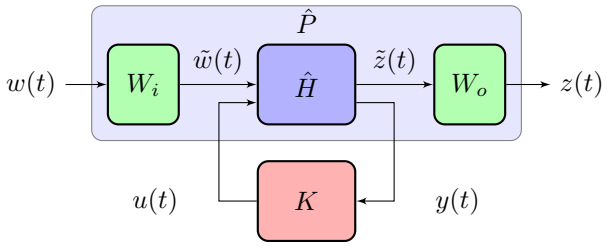


Figure 10.3: General control problem.

General control problem

The first step in $\mathcal{H}_\infty$-synthesis is to express the control plant $\hat{P}$ involved in the general control problem formulation (see for instance [Zhou et al., 1995; Apkarian and Noll, 2006; Burke et al., 2006]). This general plant is illustrated in Figure 10.3 and represents the interconnection of the open-loop model $\hat{\mathbf{H}}$ (in violet) and the performances objectives which are characterised by weighting functions (in green). The $\mathcal{H}_\infty$ control problem then consists in finding a controller $K \in \mathcal{K}$ which solves

$$K = \arg \min_{\substack{\tilde{K} \in \mathcal{K} \\ \tilde{K} \text{ stabilises } \hat{P}}} \|\mathcal{F}_l(\hat{P}, \tilde{K})\|_{\mathcal{H}_\infty}. \tag{10.2}$$

Here, the signals used to build the standard control problem are the following,

- $\tilde{w}(t) = w(t) = w_a(t)$ $(W_i = 1)$, representing the only exogenous input,
- $u(t) = u_a(t)$, representing the only control input,

- the measured output vector $y(t)$ is composed of $a_{cont}(t)$, $q_{cont}(t)$ and $a_{pil}(t)$,
- $z(t) = W_o \tilde{z}(t)$, which gathers the output performances.

More specifically, the performance signal $z(t)$ is linked to *(i)* the input signal $u_a(t)$, *(ii)* the acceleration of the pilot $a_{pil}(t)$ and *(iii)* to a fictive error signal $\bar{a}_{cont}(t) - a_{cont}(t)$ where $\bar{a}_{cont}(t)$ is a reference signal allowing to fulfil constraint *(i)* from (10.1). In other words, $z(t)$ can be written as

$$z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \underbrace{\begin{bmatrix} W_e & 0 & 0 \\ 0 & W_u & 0 \\ 0 & 0 & W_{av} \end{bmatrix}}_{W_o} \begin{bmatrix} \bar{a}_{cont}(t) - a_{cont}(t) \\ u_a \\ a_{pil} \end{bmatrix},$$

where

- W_e is designed as a n_e-th order invertible low-pass filter aimed at penalising the error $\bar{a}_{cont} - a_{cont}$ in low frequency,

$$W_e = \frac{G_{\omega_c}}{G_{le}} \left(\frac{\frac{1}{\omega_{e2}} s + 1}{\frac{1}{\omega_{e1}} s + 1} \right)^{n_e}$$

  where
  - G_{le} enables to set the maximum low frequency gain that the controller is allowed to reach for that transfer,
  - $\omega_{e1} < \omega_{e2}$ are the cut-off pulsations,
  - G_{w_c} is the inverse of the gain of the uncontrolled transfer from w_a to a_{cont} evaluated at w_c, *i.e.* $|T_{w_a \to a_{cont}}(j\omega_c)|^{-1}$. It enables to normalise the performance channel.
- W_u is an invertible n_u-th order high-pass filter aimed at penalising the command in high frequency,

$$W_u = \frac{1}{G_{lu}} \left(\frac{\frac{1}{\omega_{u1}} s + 1}{\frac{1}{\omega_{u2}} s + 1} \right)^{n_u}$$

  with $\omega_{u1} < \omega_{u2}$. This filter allows a signal amplification of gain G_{lu} in low frequencies up to w_{u1} and add a roll-off in higher frequencies.
- the filter W_{av} is a first-order low-pass filter aimed at preventing the controller to modify $T_{w_a \to a_{pil}}$ in low frequency,

$$W_{av} = G_{av} \frac{1}{\frac{1}{\omega_{av}} s + 1}.$$

  The gain G_{av} is also normalised by the highest gain of the nominal transfer from w_a to a_{pil}.

With reference to problem (10.1), the filter W_e translates the constraint *(i)*, the filters W_u and W_{av} translate together the main objective and the constraint *(iii)*. Note that even if those filters involved several tuning parameters, they can easily be linked with engineers specifications.

The reference signal $\bar{a}_{cont}$ must represent the transfer $T_{w_a \to a_{cont}}$ in low frequency, *i.e.* below the pulsation ω_c. Since this transfer is of high order, it cannot be used directly, instead, it is replaced by its frequency-limited approximation $\hat{H}_{\omega_c}$ over $[0, \omega_c]$ obtained with the **FL-ISTIA** and the reference signal is chosen as

$$\bar{a}_{cont} = \hat{H}_{\omega_c} w_a,$$

The frequency responses of both models are plotted in Figure 10.4 and one can observe that $\hat{H}_{\omega_c}$ matches well the behaviour of $T_{w_a \to a_{cont}}$ for $\omega \leq \omega_c$.

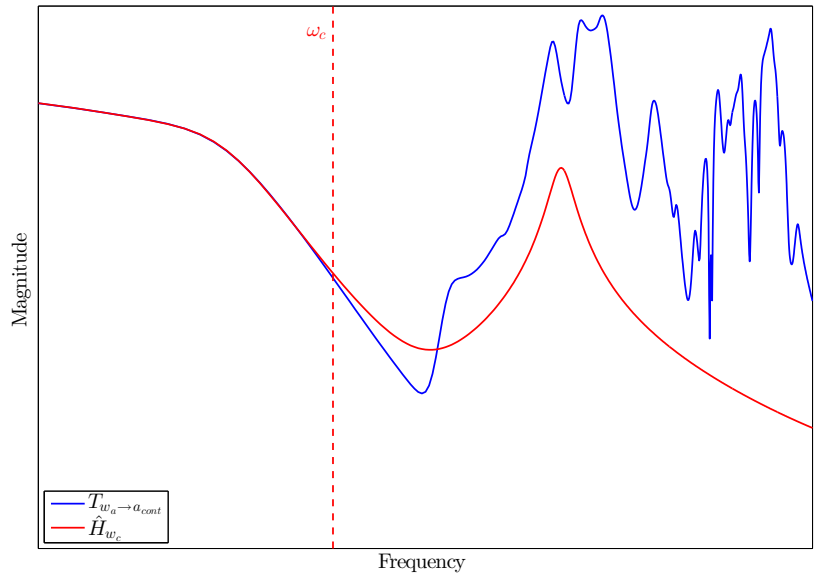


Figure 10.4: Frequency responses of $T_{w_a \to a_{cont}}$ and $\hat{H}_{\omega_c}$.

Structured $\mathcal{H}_\infty$ design

Now that the standard control problem has been formulated, it can be solved. Clearly, since problem (10.1) requires the controller to be as simple as possible, full-order design must not be considered. Instead, structured $\mathcal{H}_\infty$ synthesis is used [Apkarian and Noll, 2006; Burke et al., 2006]. Such a synthesis allows to choose the order of the controller as well as its structure.

Overall process

Since the approximation step does not offer any guarantee on the synthesised controller, a trial and error process cannot completely be avoided for the design of this anti-vibration control law. However, the overall process presented here has the advantage to assist the engineers by making the approximation and control steps more systematic. It can be summarised as follows,

(1) compute a reduced-order model $\hat{\mathbf{H}}$ from the large-scale BizJet aircraft model,

(2) generate a candidate controller K by solving the $\mathcal{H}_\infty$ control problem (10.2),

(3) perform an analysis step on the reduced-order model,

- if it is not satisfactory, then modify (relax) the gains of the weighting filters or increase the order of the desired controller and return to step (2),
- otherwise, perform an analysis on the large-scale model,
  - if it is not satisfactory, then modify the weighting filters or increase the order of the controller and return to step (2) or increase the approximation order and return to step (1),
  - otherwise, stop.

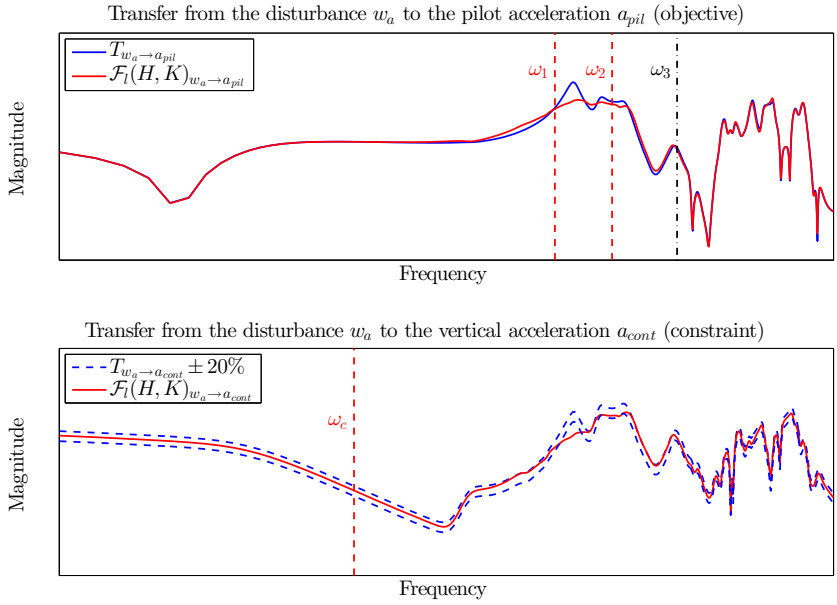


Figure 10.5: Open and closed loops of the objective (top) and constraint (bottom) transfers.

10.1.4 Numerical results

In this case, a controller K of order $n_k = 6$ which solves the $\mathcal{H}_\infty$ control problem (10.2) (with $\gamma_K = 0.97 < 1$) is obtained with the previous process and increasing further the order does not help improving the performances anymore.

With this controller, the mean vibration attenuation achieved in the frequency interval $\Omega_1 = [\omega_1, \omega_2]$, measured through the decrease of the $\mathcal{H}_{2,\Omega}$-norm over this interval, is of 27%, *i.e.*

$$\frac{\|T_{w_a \to a_{pil}}\|_{\mathcal{H}_{2,\Omega_1}} - \|\mathcal{F}_l(H,K)_{w_a \to a_{pil}}\|_{\mathcal{H}_{2,\Omega_1}}}{\|T_{w_a \to a_{pil}}\|_{\mathcal{H}_{2,\Omega_1}}} = 27\%. \tag{10.5}$$

The attenuation can also be measured by comparing the highest gains in Ω_1 reached in open and closed-loop by using Γ_{Ω_1} (see Section 7.3) to compute

$$\frac{\Gamma_{\Omega_1}(T_{w_a \to a_{pil}}) - \Gamma_{\Omega_1}(\mathcal{F}_l(H,K)_{w_a \to a_{pil}})}{\Gamma_{\Omega_1}(T_{w_a \to a_{pil}})} = 51\%. \tag{10.6}$$

Since the models in (10.6) are SISO, the bounds Γ_{Ω_1} coincide with the maximum singular values of the corresponding transfer functions in Ω_1. These results can be visualised in Figure 10.5 (top) where the frequency responses of the open and closed loops of the performance transfer are plotted.

One can also check that the transfer from w_a to a_{pil} has not been too modified by the anti-vibration controller by computing the mean error between the open and closed loops over $\Omega_2 = [0, \omega_1] \bigcup [\omega_2, \omega_3]$,

$$\frac{\|T_{w_a \to a_{pil}} - \mathcal{F}_l(H,K)_{w_a \to a_{pil}}\|_{\mathcal{H}_{2,\Omega_2}}}{\|T_{w_a \to a_{pil}}\|_{\mathcal{H}_{2,\Omega_2}}} = 28\%. \tag{10.7}$$

Note that (10.5) and (10.7) does not represent the same quantities at all. Indeed, with reference to Figure 10.5 (top), (10.5) represents the proportion of the area between the blue and red curves

in Ω_1 in comparison to the area below the blue curve in the same interval while (10.7) represents the proportion of the mean absolute gap between the red and blue curves over Ω_2 in comparison to the area below the blue curve over Ω_2. The high value of (10.7) comes from the gap between the open and closed loops which appears just before ω_1 and just after ω_2 (the scale is logarithmic in Figure 10.5). Yet, the error has been judged acceptable and the controller validated.

Besides, in Figure 10.5 (bottom), one can observe that for frequencies $\omega \leq \omega_c$, the closed-loop $\mathcal{F}_l(H, K)_{w_a \to a_{cont}}$ remains in the envelope formed by $\pm 20\%$ of the open-loop transfer $T_{w_a \to a_{cont}}$.

Conclusion

In this section, the approach that has been followed in order to design an anti-vibration control law for one large-scale BizJet aircraft model has been presented. Although it is not completely generic, the approach has greatly simplified the process of designing a controller subject to various constraints.

The design process presented here represents only a preliminary study to what has finally been used in practice. Indeed, here, only one model representing the aircraft at one single flight operating condition has been considered. The final controller K has been obtained by generalising this control process through a *multi-model* $\mathcal{H}_\infty$ synthesis (not detailed here). The global stability of this final controller is assessed in the next section through a global stability analysis.

10.2 Global stability validation of a parameter varying business jet aircraft

In this Section, the anti-vibration controller K generated by the multi-model extension of the process described in Section 10.1.3 is considered. Its stability must be assessed over the full parametric domain in presence of actuator saturation. The problem is stated more formally in Section 10.2.1 while Sections 10.2.2, 10.2.3 and 10.2.4 describe the three steps involved to address this problem.

10.2.1 Problem statement

Let us consider the set of large-scale models $\{\mathbf{H}_1, \ldots, \mathbf{H}_N\}$ (in blue in Figure 10.6), which represent the BizJet aircraft model at different flight operating conditions, interconnected with the controller K (in purple), which output is subject to a magnitude saturation (in red). The objective here is to prove the stability of this large-scale controlled aircraft model *over the continuum of parameter variation* when the actuator is *subject to a saturation*. In this study, the Mach-number is the only varying parameter considered and there are $N = 3$ large-scale models.

To this aim, a three steps methodology is followed

(i) first, in Section 10.2.2, each large-scale model is approximated over a frequency-limited interval and a frequency template is built to bound the worst approximation error[3],

(ii) then, the set of reduced-order models are interpolated into an LFR with a bound on the interpolation error in Section 10.2.3,

(iii) finally, the stability validation is performed in Section 10.2.4 with an irrational multipliers-based Integral Quadratic Constraint (IQC) approach.

[3]The robust controller designed by multi-model $\mathcal{H}_\infty$ synthesis considered in this section is based on these frequency-limited reduced-order models.

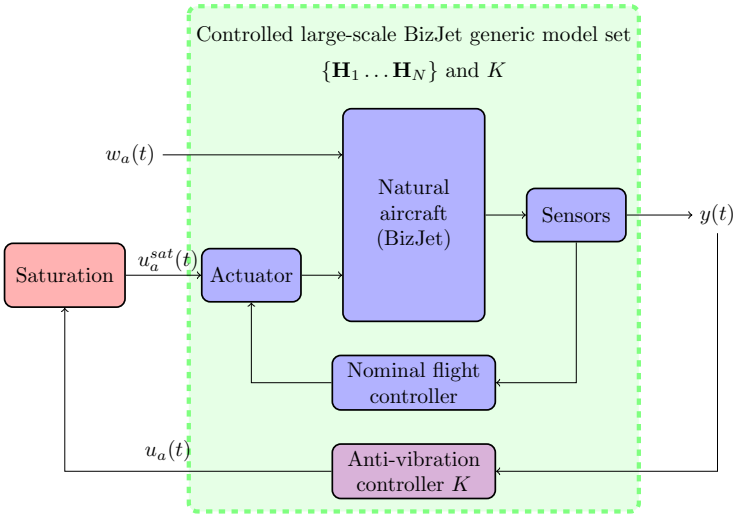


Figure 10.6: Controlled large-scale models $\mathbf{H}_i$ $(i = 1, \dots, N)$ with a saturation in input.

Together, steps *(i)* and *(ii)* are aimed at transforming the initial set of controlled large-scale models (green block in Figure 10.6) into a low-order uncertain and parameter varying model $\hat{\mathbf{H}}(\Delta_r, \delta_M, \delta_{err})$ as represented in Figure 10.7 where $\Delta_r(s)$ and δ_{err} embed the approximation and interpolation errors, respectively. These errors bounds are finally taken into account in step *(iii)* with the saturation to check if necessary conditions for stability are fulfilled.

Note that an in-depth description of the techniques and tools involved in steps *(ii)* and *(iii)* is out of the scope of this manuscript. For further information on the method used for interpolating a set of models into a LFR, refer to [Ferreres, 2011]. The irrational multipliers-based approach is an extension of the method presented in [Demourant, 2013]. Additional useful information and references can also be found on the site of the SMAC Toolbox, `http://w3.onera.fr/smac/`, which gathers the tools used for steps *(ii)* and *(iii)*.

10.2.2 Model approximation & error bounding

Approximation of the large-scale models

Here, the reduced-order models are aimed at being interpolated in place of the large-scale ones. For this interpolation process to be as simple as possible and for the resulting parameter varying model to have a low-complexity with respect to the parameters, the dimension of the reduced-order models to be interpolated must be as low as possible. Hence, unlike in Section 10.1 where the controller has been synthesised, one cannot consider 50-th order approximations.

To limit the complexity of the reduced-order models, the initial large-scale models are reduced over the bounded frequency interval $\Omega = [0, \omega_3]$ which is where the main dynamics of the models are located. Discarding the higher frequency dynamics makes sense since *(i)* the dynamics they represent are not perfectly known, hence discarding them during the approximation process only implies a higher uncertainty at high frequency for the low-order uncertain model and *(ii)* since the controller K has been designed with a roll-off filter preventing it from disturbing high frequency, it is unlikely that unstability in closed-loop comes from these dynamics.

The approximation order is set to $r = 16$ and the reduced-order models $\hat{\mathbf{G}}_i$ $(i = 1, \dots, N)$ are obtained with **DARPO**. The relative errors

$$E_i = \frac{\|H_i - \hat{G}_i\|_{\mathcal{H}_{2,\Omega}}}{\|H_i\|_{\mathcal{H}_{2,\Omega}}},$$

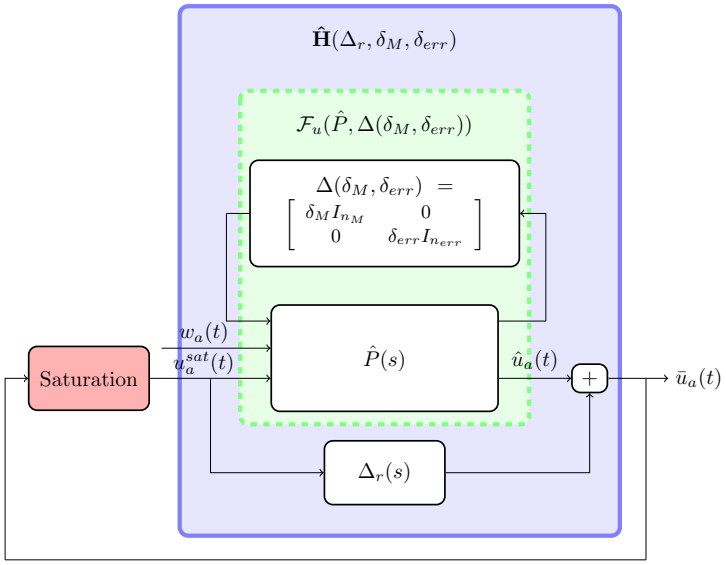


Figure 10.7: Low-order uncertain and parameter varying model $\hat{\mathbf{H}}(\Delta_r, \delta_M, \delta_{err})$ which encompasses all the the possible trajectories of the set of large-scale models in Figure 10.6.

i	1	2	3
E_i (%)	2.86	2.39	2.49

Table 10.1: Relative $\mathcal{H}_{2,\Omega}$ approximation errors between $\mathbf{H}_i$ and $\hat{\mathbf{G}}_i$ ($i = \{1, 2, 3\}$).

achieved by these models are reported in Table 10.1 and the frequency responses of $\mathbf{H}_1$ and $\hat{\mathbf{G}}_1$ are plotted in Figure 10.8.

One can observe from Table 10.1 that limiting the approximation interval enables to greatly decrease the approximation error. Indeed, with an approximation order of $r = 16$, a relative error inferior to 3% is obtained while the relative error was superior to 30% for the same order when approximating over $[0, \infty)$ (see Figure 10.2). This was expected since the high frequency dynamics require a complex model to be accurately represented while the low frequency behaviour, which contains the rigid behaviour and the first flexible modes, can be caught more easily.

Bounds on the approximation errors

For the uncertain low-order model $\hat{\mathbf{H}}$ from Figure 10.7 to encompass the error induced by the approximation step, a worst-case operator $\Delta_r(s)$ must be constructed. Due to the fact that the controller K is included in $\hat{\mathbf{H}}$, $\Delta_r(s)$ is not built to bound directly the approximation error but rather the error between the open-loop transfer functions from the input u_a^{sat} to the outputs of the controller K, *i.e.* u_a and $\hat{u}_a$, respectively. More specifically, by denoting $F_i(s) = K(s)H_i(s)_{u_a^{sat} \to y}$ and $\hat{F}_i(s) = K(s)\hat{G}_i(s)_{u_a^{sat} \to \hat{y}}$, $\Delta_r(s)$ must statisfies

$$\|\Delta_r(j\omega)\|_F \geq \max_{i=1,\dots N} \left\| F_i(j\omega) - \hat{F}_i(j\omega) \right\|_F,$$

for all $\omega \in \mathbb{R}_+$. Note that these open-loops are SISO, thus the Frobenius norm reduces to the absolute value.

Since $\Delta_r(s)$ is not involved in the interpolation step and the method used to perform the IQC analysis is not limited by the dimension of the block, the order of $\Delta_r(s)$ can be fixed freely.

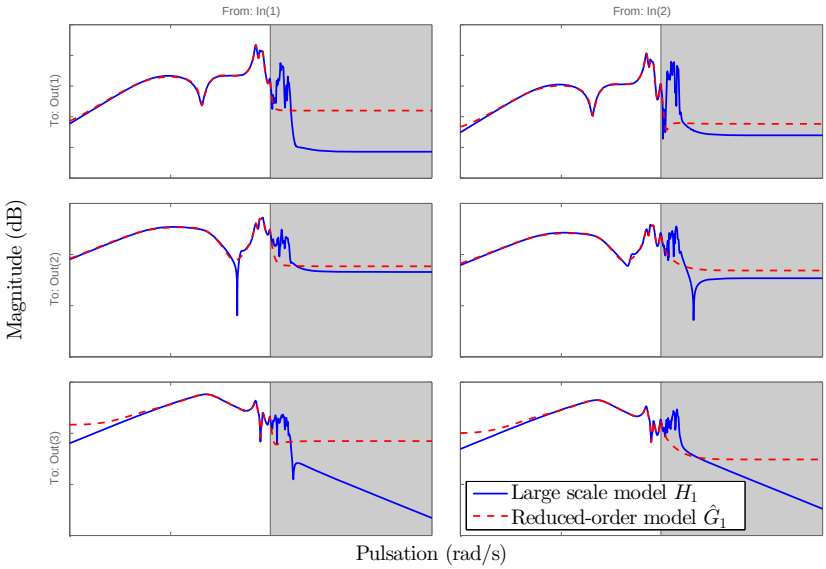


Figure 10.8: Frequency responses of the large scale model $\mathbf{H}_1$ and of the 16-th order reduced-order model $\hat{\mathbf{G}}_1$ obtained with **DARPO** for $\Omega = [0, \omega_3]$.

In this case, an order 25 is chosen and the frequency template $\Delta_r(s)$ is built with an extension of the method presented in Section 7.3.4 which provides less conservatism, especially considering the shape of the error here[4]. The resulting frequency template is plotted in Figure 10.9.

10.2.3 Low-order LFR model creation

Given the set of reduced-order models obtained in Section 10.2.2, the objective is here to derive a low-order uncertain and parameter varying model that can be exploited by the IQC framework in Section 10.2.4. More precisely, and with reference to Figure 10.7, the aim is to compute the low-order interconnection plant $\hat{P}(s)$, such that the *continuum* of parametrised transfer functions

$$T_{u_a^{sat} \to \hat{u}_a}(s) = \mathcal{F}_u(\hat{P}(s), \Delta(\delta_M, \delta_{err})),$$

where $\Delta(\delta_M, \delta_{err}) = \mathbf{diag}(\delta_M I_{n_M}, \delta_{err} I_{n_{err}})$, covers all possible Mach cases in the admissible range when the normalised parameter δ_M evolves in $[-1, 1]$.

The plant $\hat{P}(s)$ is obtained by polynomial interpolation from the set of low-order open-loop plants $\hat{F}_i(s)$ associated with a selection of Mach numbers obtained in the previous section. To limit the complexity of the LFR, low-complexity polynomials are used which yields interpolation errors that have to be encompassed by the uncertainty $\delta_{err} I_{n_{err}}$ which creation is detailed thereafter. Together, the LFR $\mathcal{F}_u(\hat{P}(s), \Delta(\delta_M, \delta_{err}))$ and the approximation error $\Delta_r(s)$ enable the nonlinear closed-loop model in Figure 10.7 to encompass all possible trajectories of the parametrised family of large-scale models with a reasonable complexity. The creation of the plant $\hat{P}(s)$ consists in three steps detailed below.

[4]This extension is described quickly in the perspectives in Section 11.3

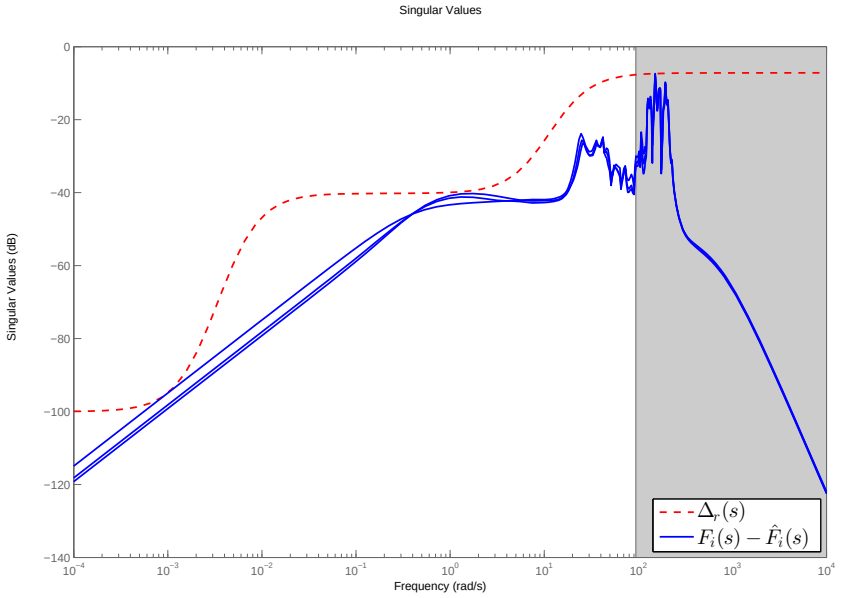


Figure 10.9: Frequency template $\Delta_r(s)$ of the approximation error.

Preliminary transformation of the open-loop models

The reduced-order open-loop transfer functions $\hat{F}_i(s)$ $(i = 1, \ldots, N)$ obtained in Section 10.2.2 have now 22 states (the controller has 6 states and the reduced-order model has 16 states) with 4 pairs of poorly damped but rather high frequency modes. Those transfers are available for three Mach number ranging from 0.8 to 0.9 and includes a fixed part mainly due to the actuator whose model is independent of the Mach number. Hence, each transfer can be factorised as

$$\hat{F}_i(s) = A_c(s)L_i(s),$$

where $A_c(s) = s(s+a)^{-1}(s+b)^{-1}$ $(a, b > 0)$ is a constant third order transfer functions while the 19-th order Mach-number dependent transfers $L_i(s) = C_i(sI_{19} - A_i)^{-1}B_i + D_i$ are rewritten here in a rescaled state-space companion form (see Remark 22) as,

$$\left(\begin{array}{c|c} A_i & B_i \\ \hline C_i & D_i \end{array}\right) = \left(\begin{array}{cccc|c} 0 & \tau_1 & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & \tau_{18} & 0 \\ a_{1,i} & a_{2,i} & \ldots & a_{19,i} & b_i \\ \hline c_{1,i} & c_{1,i} & \ldots & c_{19,i} & d_i \end{array}\right) \tag{10.8}$$

Remark 22 (About the rescaled companion form)**.** *Note that the chosen state-space representation (10.8) is a straightforward generalisation of the standard controllable canonical form which is easily recovered with $\tau_1 = \tau_2 = \ldots = \tau_{18} = 1$. Here, these parameters are tuned in order to improve the condition number of A_i.*

In the context of LFR modeling the above description is of very high interest since the varying state-space parameters to be interpolated all appear only in the last two lines of the model. Thus, assuming that every parameter is approximated by a q-th order polynomial, the size of the Δ-block $\delta_M I_{n_M}$ capturing Mach-number induced variations is limited to $n_M = 2q$.

For further details on the rescaled state-space companion form, refer to [Ferreres, 2011].

Polynomial interpolation with guaranteed error bounds

Let us denote $V^{(i)} = V(\delta_{M_i}) \in \mathbb{R}^{2\times 20}$ the last two lines of the matrix (10.8) associated with its varying part. The objective here is to find a polynomial approximation of the finite set of matrices $\{V^{(i)}\}_{i=1\dots N}$ with guaranteed error bounds which will constitute the uncertainty δ_{err}.

Given q, the order of the polynomial, the problem reduces to the determination of a set of matrices $\{V_l\}_{l=0\dots q}$ such that the nonnegative entries of the error matrix E are minimised under the following linear constraints (with $j = 1, 2$ and $k = 1 \dots 20$),

$$\left| \left[V_0 + \sum_{l=1}^{q} \delta_{M_i}^l V_l - V^{(i)} \right]_{j,k} \right| \leq E_{j,k}\,,\ i = 1 \dots N.$$

The above optimisation problem is easily solved by standard linear programming solvers. Note that the order q of the polynomial must be carefully chosen. Indeed, low values might lead to a rough approximation which might result in a too conservative LFR model. Conversely, high values may help improve the accuracy of the interpolation but the resulting LFR might be too complex. Besides, undesirable oscillations are likely to appear between the interpolation points if the difference between the number of points and the order of the polynomial, *i.e.* $N - q$, becomes too small. This is typically the case here, where $N = 3$.

To alleviate this issue, additional models for intermediate Mach numbers must be considered to enrich the set $\{V^{(i)}\}_{i=1\dots N}$. Here, since there is no additional models available, fictitious models must be created. For simplicity, this is done through a linear interpolation process which provides stable models with smooth evolution of the closed-loop poles for two new Mach numbers 0.825 and 0.875. With these additional points, the number of models available increases from $N = 3$ to $n_M = 5$ which is sufficient to avoid any oscillation in the interpolation.

The result of the interpolation process with $q = 2$ is illustrated in Figure 10.10 where the evolution of the coefficient c_{19} (chosen for its significant variations) is plotted. The two red triangles represent the two fictitious points which have been added. The nominal Mach-dependent polynomial function is represented by the solid blue line while its uppers and lower bounds are represented by the dotted red lines. For this coefficient, the error is bounded by $E_{2,19} = 0.04$ and its three nominal values for $\delta_M = -1$, $\delta_M = 0$ and $\delta_M = 1$ (blue circles) all appear on the upper-bound.

Quite interestingly, with $n = 2$, a similar property as the one observed above is true for any of the 40 parameters whose nominal values on the initial Mach grid either all belong to the upper-bound or the lower-bound of its polynomial approximation. Consequently, a matrix $\tilde{E}$ whose entries satisfy $\tilde{E}_{j,k} = \pm E_{j,k}$ is easily deduced from E. Hence, for any value of $\delta_M = \delta_{M_i}$ on the initial grid, there exists a scalar $\delta_{err} \in [-1,\ 1]$ such that

$$V^{(i)} = V_0 + \sum_{l=1}^{2} \delta_{M_i}^l V_l + \delta_{err}\tilde{E}. \tag{10.9}$$

LFR modeling

Based on (10.9) and using standard LFR manipulations (see [Magni, 2006]), one can easily verify the existence of a matrix $\Sigma_V \in \mathbb{R}^{8\times 48}$ such that $\forall(\delta_M, \delta_{err}) \in [-1,\ 1]^2$,

$$V(\delta_M, \delta_{err}) = \mathcal{F}_u(\Sigma_V, \mathbf{diag}(\delta_M I_5, \delta_{err} I_2)), \tag{10.10}$$

where $\mathcal{F}_u()$ represent the upper LFR. Then, there simply remains to insert the static LFR object (10.10) into the state-space description (10.8) and to multiply the result by the fixed part $A_c(s)$[5]. Finally, one obtains the interconnection $\mathcal{F}_u(\hat{P}, \Delta(\delta_M, \delta_{err}))$, with $\Delta(\delta_M, \delta_{err}) = \mathbf{diag}(\delta_M I_{n_M}, \delta_{err} I_{n_{err}})$, depicted in Figure 10.7 with $n_M = 5$ and $n_{err} = 2$.

[5]These operations are easily handled by the LFR Toolbox [Magni, 2006].

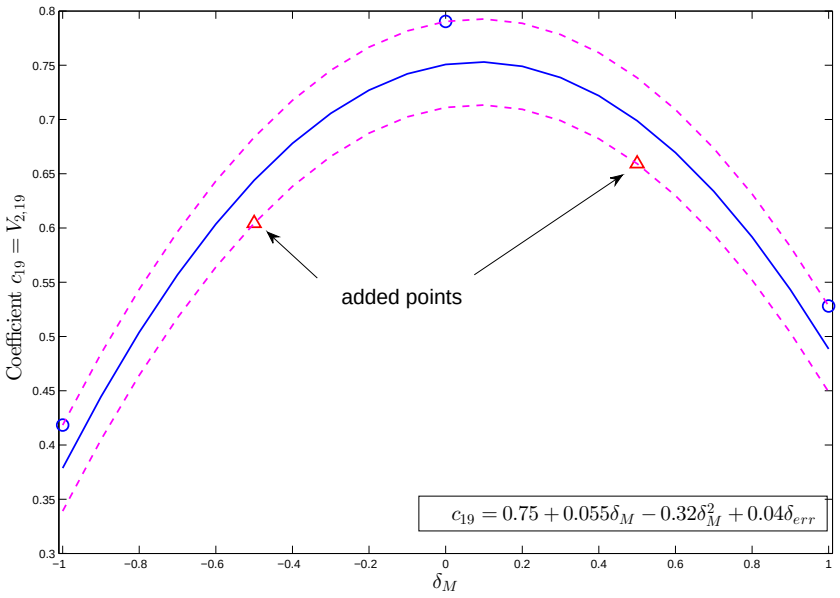


Figure 10.10: Illustration of a 2^{nd} order polynomial interpolation result with minimized guaranteed error bound for the coefficient $c_{19} = V_{2,19}$.

Preliminary validation tests on the linear closed-loop

Before the global validation step, one can already perform some preliminary tests without the saturation nor the approximation error Δ_r by looking at the closed-loop poles for $(\delta_M, \delta_{err}) \in [-1, 1]^2$. If unstable poles are detected, then the polynomial approximation must be refined by increasing the interpolation order q.

From Figure 10.11, one can observe that all the closed-loops remain stable here. Hence, it does not seem that an higher order polynomial approximation is required here. In order to verify that no worst-case parametric combination has been missed, a μ analysis with respect to the parametric uncertainties δ_M, δ_{err} and the worst-case LTI operator Δ_r must be performed (see [Biannic and Ferreres, 2005; Roos et al., 2011] for further information about μ analysis and the tools used here). This step (not detailed here) enables to show that the continuum of closed-loop models remains stable for an admissible uncertainty. It remains to check the stability in presence of the saturation which cannot be achieved by μ analysis anymore.

10.2.4 IQC-based global stability analysis

Generally speaking, IQC methods enable to analyse the stability and performance properties of an interconnection $M(s)/\Delta$ of a LTI operator $M(s)$ with a structured model uncertainty Δ which can gather non-linearities (such as saturation), LTI or Linear time-varying parameters, delays, etc. Due to their versatility, IQC methods can be used on a wide range of problems and are perfectly suited here due to the presence of the input saturation.

General principle

Let us consider two signals, v and w, square integrable over $[0, \infty)$ and which represent the input and output signals of the uncertainty block Δ, *i.e.* $w = \Delta v$. These signals are said to satisfy

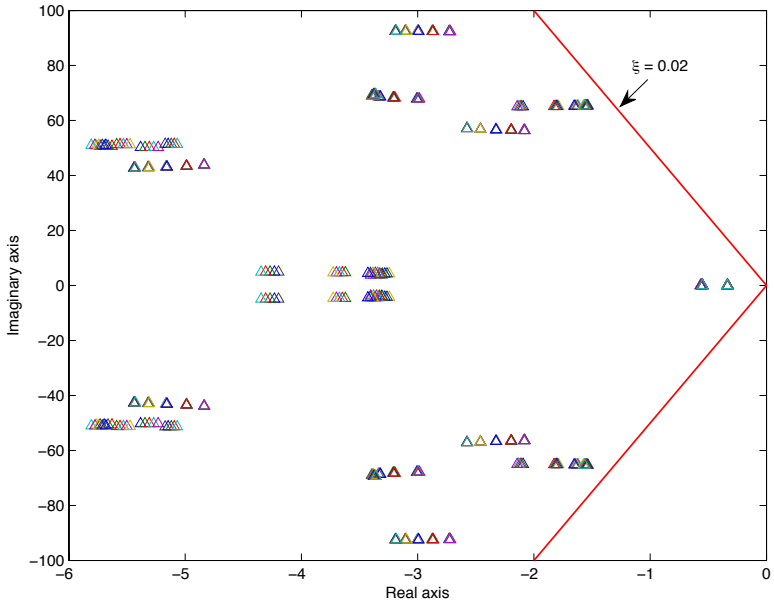


Figure 10.11: Zoom near the origin on the closed-loop poles evolution without saturation for $\delta_M \in [-1, 1]$ and $\delta_{err} \in [-1, 1]$ and $\Delta_r(s) = 0$.

the IQC defined by Π if and only if their Fourrier transforms $\tilde{v}$ and $\tilde{w}$ satisfy

$$\int_{-\infty}^{\infty} \begin{bmatrix} \tilde{v}(j\omega) \\ \tilde{w}(j\omega) \end{bmatrix}^H \Pi(j\omega) \begin{bmatrix} \tilde{v}(j\omega) \\ \tilde{w}(j\omega) \end{bmatrix} \leq 0.$$

Π can be any measurable Hermitian matrix function of ω and is called a *multiplier*. To analyse the stability of the interconnection of $M(s)$ with the bounded operator Δ (through a positive feedback) which input and output signals satisfy the IQC defined by Π, one must verify that[6], $\forall \omega \in \mathbb{R}_+$,

$$\begin{bmatrix} M(j\omega) \\ I \end{bmatrix}^H \Pi(j\omega) \begin{bmatrix} M(j\omega) \\ I \end{bmatrix} \prec 0. \tag{10.11}$$

Condition (10.11) can be turned into a Linear Matrix Inequality (LMI) feasibility problem dependent on the frequency ω. Then, the standard approach to address the latter problem is to use the Kalman-Yakubovitch-Popov lemma to transform it into another LMI feasibility problem which does not depend on the frequency ω anymore (see [Megretski and Rantzer, 1997] for further information). This approach enables to transform an infinite number of LMI feasibility problems into one single LMI feasibility problem.

Yet, the number of decision variables involved in the last LMI problem grows quadratically with the closed-loop order which might make it quickly untractable in practice if the model is of large dimension. In addition, the parametrisation of the multiplier Π must be chosen a priori which might lead to a tedious trial and error process[7]. Instead here, an approach based on frequency inequalities on a gridding is used.

[6] Somme additional mild assumptions are actually required, see [Megretski and Rantzer, 1997] for further information

[7] Indeed, if the feasibility problem is not satisfied for a parametrisation, one cannot conclude on the stability of the closed-loop since there may be another parametrisation which would imply the feasibility.

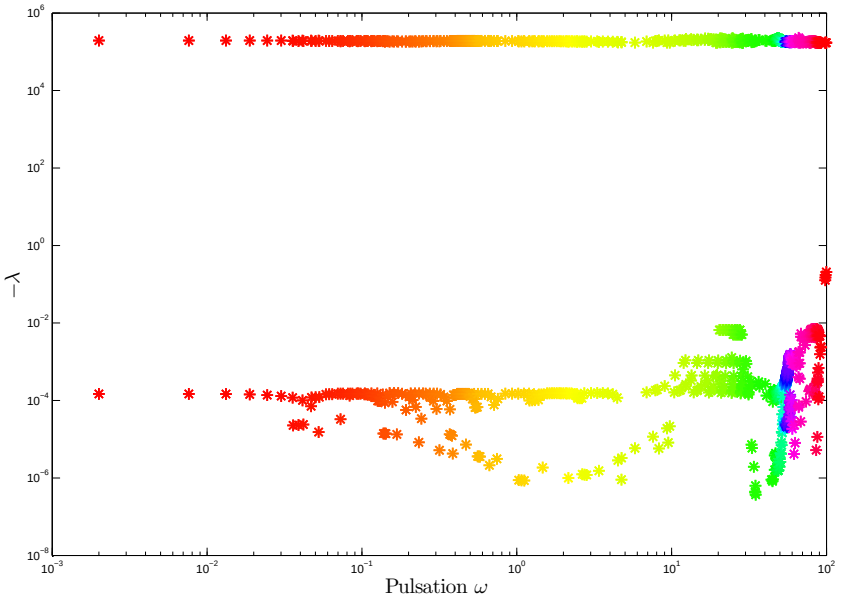


Figure 10.12: Opposite of the eigenvalues of (10.11) between 0 and 100 rad/s for the plant $\hat{P}(s)$ interconnected with Φ.

Frequency gridding approach for the IQC

The technique has been proposed in [Sideris and Sanchez Pena, 1990] and adapted to the problem considered here in [Demourant, 2013]. The idea of the method relies on the fact that if (10.11) is true for some frequency ω_i, then it holds for a small interval Ω_i around ω_i which upper and lower bounds can be characterised. Hence, by iteratively adding new frequency points ω_i for which new multipliers $\Pi_i = \Pi(\omega_i)$ are found, one can extend the interval where (10.11) is verified. In particular, if the intersection of all the intervals of validity Ω_i covers $\mathbb{R}_+$, *i.e.* if $\bigcup_{i=1,\dots,N_\Omega} \Omega_i = [0, \infty)$, then it means that (10.11) holds true for all $\omega \in \mathbb{R}_+$ which implies the stability of the interconnection of $M(s)$ and Δ. If at some frequency, no multiplier is found, then the stability is not guaranteed.

Note that with such an approach, one actually obtains a set of multipliers Π_i associated with each interval Ω_i, *i.e.* $\forall \omega \in \Omega_i$, $\Pi(j\omega) = \Pi_i(j\omega)$.

Global stability validation for the anti-vibration controller

Here, the method described above is used to analyse the stability of the interconnection of $\hat{P}(s)$ with $\Phi = \mathbf{diag}(\phi, \Delta_r(s), \delta_M I_{n_M}, \delta_{err} I_{n_{err}})$ where ϕ represents the saturation. See [Vuillemin et al., 2014a] for a precise description of the parametrisation of the multiplier.

After 8 iterations, the methods stops with $N_\Omega = 104$ valid frequency intervals Ω_i which cover $\mathbb{R}_+$. Hence, the solution is validated on the whole frequency domain. This is illustrated in Figure 10.12 where the opposite of the eigenvalues of $\begin{bmatrix} \hat{P}(j\omega) & I \end{bmatrix}^* \Pi(j\omega) \begin{bmatrix} \hat{P}(j\omega) & I \end{bmatrix}^T$ are plotted for pulsations going from 0 to 100 rad/s. Note that these eigenvalues are necessarily real since $\Pi(j\omega)$ is Hermitian and that $\Pi(j\omega)$ may change for different pulsations since it is piece-wise continuous. For each pulsation, there are 4, possibly multiple, eigenvalues which are plotted with the same colour. One can observe that the opposite of the eigenvalues remains positive over $\mathbb{R}_+$.

This enables to conclude on the stability of the low-order uncertain and parameter varying model $\hat{\mathbf{H}}(\Delta_r, \delta_M, \delta_{err})$ subject to the input saturation. Since the trajectories of this model also encompass all the possible trajectories of the parametrised family of large-scale models $\mathbf{H}_i$ $(i = 1, \ldots, N)$, one can conclude on the global stability of the large-scale parameter varying model subject to input saturation for all normalised Mach numbers $\delta_M \in [-1, 1]$.

Conclusion

In this chapter, the global process followed for the design and global stability validation of an anti-vibration control law for an industrial business jet aircraft represented by a family of large-scale LTI models subject to input saturation has been presented. Model approximation merely represents a preliminary step in both the design and validation process, but still a mandatory one.

Indeed, in the design step, model approximation has enabled to consider a control problem with complex constraints which would have been difficult to address directly with the initial large-scale model. Similarly, in the analysis step, model approximation has enabled to greatly decrease the complexity of the family of large-scale models which could not have been interpolated otherwise. In addition, by bounding the approximation error, the stability of the initial large-scale models can still be guaranteed.

This chapter concludes the contributions of this thesis and the next (and last) part concludes this thesis.

Part IV

Conclusion

Chapter 11

Discussion

In this Chapter, the main contributions of the thesis are recalled and their limitations are discussed. In addition, solutions or, at least hints, are proposed to alleviate these drawbacks and to extend the range of application of the results.

Contents

11.1 FL-ISTIA . . . 179
11.2 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm . . . 180
11.2.1 Numerical robustness . . . 180
11.2.2 Scalability . . . 182
11.3 Bounds on the $\mathcal{H}_\infty$-norm and its applications . . . 183
11.3.1 Improvement of the construction of frequency templates . . . 183
11.3.2 Other possible applications for the $\mathcal{H}_\infty$ bound Γ_Ω . . . 185
11.4 First-order optimality conditions for the $\mathcal{H}_{2,\Omega}$ approximation problem . . . 185
11.5 DARPO . . . 186
11.5.1 Issues related to the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm . . . 186
11.5.2 Optimisation scheme . . . 188
11.5.3 Approximation for a fixed error . . . 189
11.5.4 Model approximation for the control of large-scale models . . . 190

11.1 FL-ISTIA

In Chapter 6, a heuristic method for frequency-limited model approximation has been proposed. The core of the method is the same as the sub-optimal $\mathcal{H}_2$ model approximation method **ISRKA** presented in Section 4.2.2, excepted that frequency-limited gramians are used in place of the infinite gramians. This modification has been inspired by the way the **FL-BT** has been obtained from the **BT** (see Sections 5.2.1 and 3.1.2, respectively).

In spite of its simplicity, this modification has led to improve the overall performances of the **ISRKA** in terms of the $\mathcal{H}_{2,\Omega}$-norm of the error. Yet, two fundamentals elements of the method are directly inherited from the **ISRKA** and have not been modified

- the Krylov subspaces involved in the interpolation,
- the selection of the new interpolation points as the mirror images of the current reduced-order model poles.

More specifically, the Krylov subspaces and the interpolation points are such that at convergence, the reduced-order model tangentially interpolates the large-scale model from the left (or the right) at the mirror images of its poles, *i.e.*

$$H(-\hat{\lambda}_i)\hat{\mathbf{b}}_\mathbf{i}^T = \hat{H}(-\hat{\lambda}_i)\hat{\mathbf{b}}_\mathbf{i}^T,\ (i = 1, \ldots, r), \tag{11.1}$$

which is one of the first order optimality condition of the optimal $\mathcal{H}_2$ approximation problem (see Section 4.1).

However, as it has been shown in Chapter 8, the optimality conditions for the optimal $\mathcal{H}_{2,\Omega}$ approximation problem are not equivalent to the $\mathcal{H}_2$ ones, hence fulfilling (11.1) is clearly irrelevant for the approximation over a bounded frequency range. Hence, **FL-ISTIA** could greatly be improved by considering meaningful Krylov subspaces and interpolation points.

To this aim, one should determine whether the interpolation conditions expressed for the optimal $\mathcal{H}_{2,\Omega}$ approximation problem in Chapter 8 can be fulfilled through projection on Krylov subspaces. If not possible, an approximation of these interpolation conditions may be used instead, or, simply another interpolation framework such as the Loewner framework [Antoulas et al., 2012; Ionita, 2013].

11.2 Poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm

In Chapter 7, the $\mathcal{H}_{2,\Omega}$-norm of a LTI dynamical model has been formulated through the use of the poles and residues of its transfer function both when the model has semi-simple poles and when it has high order poles. This formulation naturally extends the one existing for the $\mathcal{H}_2$-norm (see Section 2.2) and enables to express the $\mathcal{H}_{2,\Omega}$-norm of a model in a closed form involving only elementary functions as soon as the poles and residues of the model are available. It has proven to be numerically very efficient for the computation of the $\mathcal{H}_{2,\Omega}$-norm of models with semi-simple poles only and has enabled to derive the most significant results in this thesis.

11.2.1 Numerical robustness

The poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm is mainly limited by two elements in the large-scale framework,

- the eigenvalue decomposition is not particularly numerically robust,
- the formulation is not convenient when the model has Jordan blocks. Indeed, the computation of the Jordan decomposition is only possible on "toy models" and should not be considered in larger settings.

When Jordan blocks are involved, two workarounds can be considered,

(i) if the residues of a transfer function can be efficiently computed knowing the eigenvalues but not the eigenvectors, then the $\mathcal{H}_{2,\Omega}$-norm may be computed by evaluating the transfer function and its derivative as in the $\mathcal{H}_2$-case (see Section 7.2.3). Evaluating the derivative of the transfer function is not particularly simple either but remains preferable than computing the Jordan canonical form.

(ii) It is well-known that matrices with Jordan blocks are extremely sensitive to perturbations, however, one may want to consider the impact of this sensitivity for the computation of the $\mathcal{H}_{2,\Omega}$-norm. As illustration, let us consider a n-th order model $\mathbf{H} = (A, B, C)$ where

$$A = \begin{bmatrix} -1 & 1 & & \\ & -1 & \ddots & \\ & & \ddots & 1 \\ & & & -1 \end{bmatrix} \in \mathbb{R}^{n\times n},\ C = B^T = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{1\times n},$$

i.e. $\lambda = -1$ is an eigenvalue of order n. Let us also consider the perturbed model $\mathbf{H}_\epsilon = (A_\epsilon, B, C)$, where

$$A_\epsilon = \begin{bmatrix} -1 & 1 & & \\ & -1 & \ddots & \\ & & \ddots & 1 \\ \epsilon & & & -1 \end{bmatrix} \in \mathbb{R}^{n\times n},$$

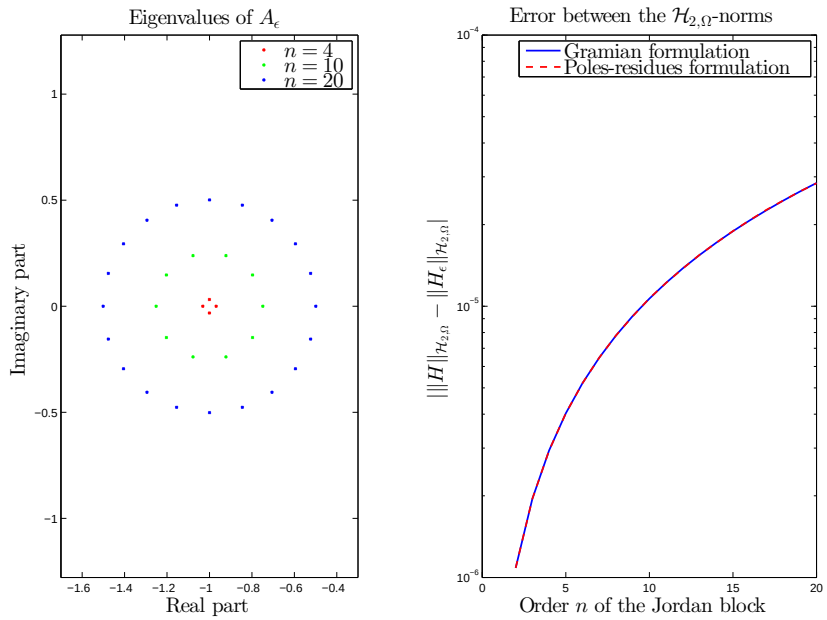


Figure 11.1: Eigenvalues of the perturbed matrix A_ϵ (left) and error between the $\mathcal{H}_{2,\Omega}$ norms of $\mathbf{H}$ and $\mathbf{H}_\epsilon$ (right) for $\Omega = [0, 5]$ and $\epsilon = 10^{-6}$.

The eigenvalues of A_ϵ are located on a circle centred at -1 and of radius $\epsilon^{\frac{1}{n}}$ as illustrated in Figure 11.1 (left) and are no longer multiple, *i.e.* A_ϵ is not defective anymore. Let us compute the $\mathcal{H}_{2,\Omega}$-norm of $\mathbf{H}$ and of $\mathbf{H}_\epsilon$ for $n = 2, \ldots, 20$, and check the error between the two norms

$$\left| \|H\|_{\mathcal{H}_{2,\Omega}} - \|H_\epsilon\|_{\mathcal{H}_{2,\Omega}} \right|,$$

where $\|H_\epsilon\|_{\mathcal{H}_{2,\Omega}}$ is calculated both with the poles-residues and gramian formulations for $\Omega = [0, 5]$ and $\epsilon = 10^{-6}$. The errors are plotted in Figure 11.1 (right) with respect to n.

One can observe that the error is increasing from ϵ to $\sim 3 \times 10^{-5}$ which remains reasonably low. In particular, when comparing this error to the $\mathcal{H}_{2,\Omega}$-norm of $\mathbf{H}$ for $n = 20$ which is equal to 51.66 one obtains a relative error of $10^{-4}\%$.

Note that with this value of ϵ both the gramian and poles-residues formulations give the same value for the $\mathcal{H}_{2,\Omega}$-norm of $\mathbf{H}_\epsilon$ but this might not always be the case. Indeed, depending on the value of ϵ, the eigenvalues of A_ϵ are more or less close to each other which can impact the eigenvalue computation process.

This simple example is not a proper perturbation study but simply highlights the fact that perturbing the models which have Jordan blocks may be a satisfactory approach to make the poles-residues formulation viable to compute their $\mathcal{H}_{2,\Omega}$-norm to a satisfactory accuracy. The main difficulty lies in the fact that, in practice, the matrix A is not given in its Jordan canonical form and determining which element must be disturbed and by which gain might not be obvious. In addition, A_ϵ remains a *nearly defective* matrix which eigen-decomposition must be computed with precautions.

A more radical solution would consist in replacing the eigen-decomposition by another matrix decomposition which is numerically more robust and which is not modified if multiple eigenvalues appear. Hessenberg or tridiagonal forms are interesting candidates since a somehow explicit

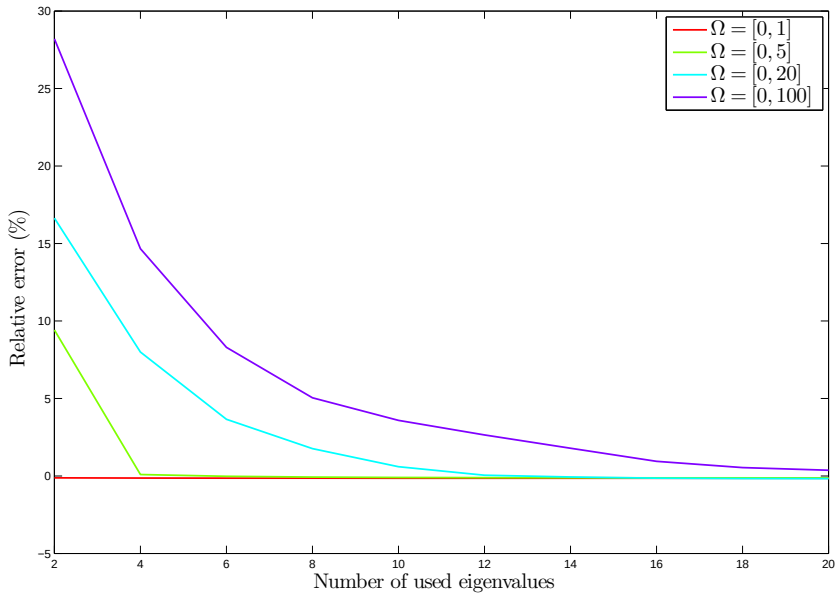


Figure 11.2: Relative error between the real $\mathcal{H}_{2,\Omega}$-norm and the approximate $\mathcal{H}_{2,\Omega}$-norm considered with a increasing number of eigenvalues on the ISS model for several frequency interval Ω.

formulae exists to invert them [Mallik, 2001]. Whether it could be exploited to compute efficiently the $\mathcal{H}_{2,\Omega}$-norm is not obvious though.

11.2.2 Scalability

Independently of the formulation chosen for the $\mathcal{H}_{2,\Omega}$-norm, its computation for very large-scale models is difficult. With the poles-residues formulation, one possible solution relies on the assumption that on an interval Ω, the $\mathcal{H}_{2,\Omega}$-norm of a model $\mathbf{H}$ can be well approximated by considering only some of its poles λ_i (and associated residues Φ_i), *i.e.*

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = -\sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i H(-\lambda_i)^T\right) \mathbf{a}_{\omega,\lambda_i} \approx -\sum_{i=1}^{m} \mathbf{tr}\left(\Phi_i H(-\lambda_i)^T\right) \mathbf{a}_{\omega,\lambda_i},$$

where $m < n$. The initial assumption is likely to be verified with models which have poorly damped dynamics.

As an illustration, let us consider the ISS model ($n = 270$) which approximate $\mathcal{H}_{2,\Omega}$-norm is computed by using an increasing number of poles ordered by decreasing values of $\mathbf{Re}\left(-\mathbf{tr}\left(\Phi_i H(-\lambda_i)^T\right) \mathbf{a}_{\omega,\lambda_i}\right)$ and compared to the true $\mathcal{H}_{2,\Omega}$-norm. In Figure 11.2, the relative error between the real $\mathcal{H}_{2,\Omega}$-norm and the approximate one is plotted for different frequency intervals. One can see that with this model, the approximate error between the approximate $\mathcal{H}_{2,\Omega}$-norm and the real one decreases quickly. Obviously, the wider the interval Ω is, the more eigenvalues are needed. However, with 20 eigenvalues used among 270, the error is quite low even on the widest interval $[0, 100]$ which embeds all the dynamics of this model. The validity of the approach needs to be experimented on less resonant models though.

From a practical point of view, the poles and residues that must be retained could be determined with an algorithm similar to the Dominant Poles Algorithm [Rommes and Martins,

2006]. The main question raised by this approach is to know whether the error induced by this approximation can be quantified without the discarded poles and residues.

11.3 Bounds on the $\mathcal{H}_\infty$-norm and its applications

In Section 7.3, the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm is used to derive two upper bounds on the $\mathcal{H}_\infty$-norm of a LTI dynamical model $\mathbf{H}$,

$$\max_{\omega\in\Omega}\sigma_{max}(H(j\omega)) \leq \Gamma_\Omega(H) \leq \bar{\Gamma}_\Omega(H).$$

From a practical point of view, computing $\bar{\Gamma}_\Omega$ requires to find the maximum of several first-order rational functions while computing Γ_Ω requires to find the maximum of a sum of univariate rational functions. These upper bounds have proven to be, quite unexpectedly, not too conservative.

Obviously, since these bounds are based on the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm, they inherit its limitations. In this case however, it is clear that the issue related to the Jordan blocks can be readily alleviated for the computation of Γ_Ω by looking for the maximum of $\|H(j\omega)\|_F$ over Ω without exploiting the eigen-decomposition of the matrix A.

These bounds have been particularly useful in the context of model approximation through the construction of frequency templates bounding the approximation error. An improvement of the algorithm developed to build the frequency templates as well as more general applications are discussed thereafter.

11.3.1 Improvement of the construction of frequency templates

The method proposed to build frequency template in Section 7.3.4 is simple and efficient but could be improved, either by using the same idea with a more elaborate template structure or by completely reformulating the problem.

A quick improvement of the frequency templates construction

The current method for building frequency templates is in two steps. First, a set of relevant pulsations at which the gain of the transfer function is higher than the gain reached for any smaller pulsations ω is determined by computing the upper bounds Γ_Ω with $\Omega = [0, \omega]$ when ω varies between 0 and the pulsation ω_∞ where the $\mathcal{H}_\infty$-norm is reached.

Then the frequency template is built by successive relaxation of its pole and gain until its Frobenius norm upper bounds the Frobenius norm of the initial transfer function at the pulsations determined previously.

The frequency template could easily be refined by enriching its structure as

$$G(s) = \underbrace{K_1\left(\frac{s-z_1}{s-p_1}\right)^{r_1}}_{G_1(s)}\underbrace{K_2\left(\frac{s-z_2}{s-p_2}\right)^{r_2}}_{G_2(s)}, \tag{11.2}$$

where the poles p_1, p_2 and the zeros z_1, z_2 are all located in the left half-plane. Mainly two interesting frequency templates can be build from (11.2),

- if $|z_1| < |p_1| < |z_2| < |p_2|$, then a frequency template with stairs can be obtained. This is illustrated in Figure 11.3. One way to obtain such a frequency template is to *(i)* divide the set of relevant pulsation $\{w_1, \ldots, w_k\}$ in two sets $\mathcal{W}_1$ and $\mathcal{W}_2$ such that the minimum pulsation of $\mathcal{W}_2$ is superior the maximum value of $\mathcal{W}_1$, *(ii)* construct G_1 as in Section 7.3.4 using only the pulsations contained in $\mathcal{W}_1$, *(iii)* construct G_2 in a similar way to G_1 with the pulsations of $\mathcal{W}_2$ but instead of fixing z_2, the pole p_2 is fixed and adjusted for each value of z_2, K_2 so that $G_2(0) = 1$ in order no to disturb the first filter G_1. A frequency

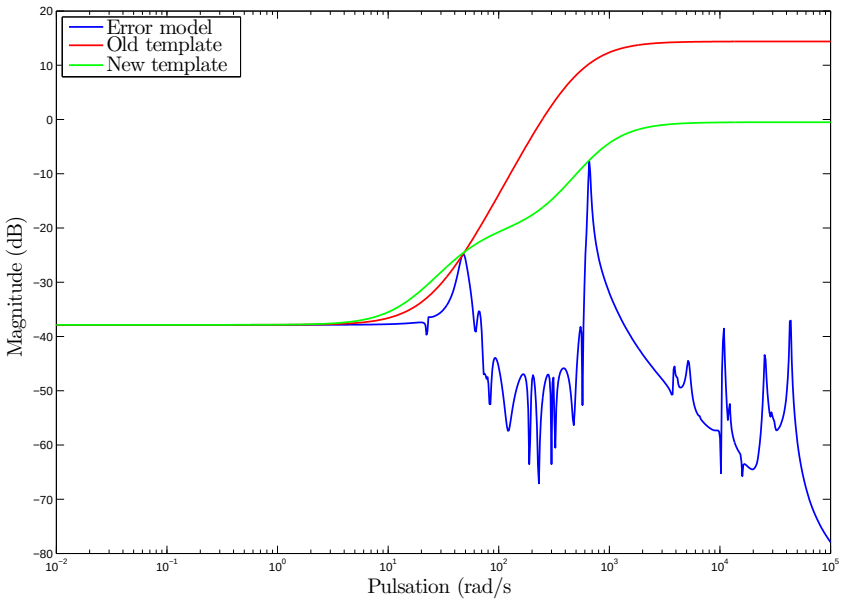


Figure 11.3: Illustration of the frequency template (11.2) when $|z_1| < |p_1| < |z_2| < |p_2|$.

template of this form has actually been used in [Vuillemin et al., 2014a] (see Section 10.2), but in its current state, the algorithm still requires some tuning.

- If $|z_1| < |p_1| < |p_2| < |z_2|$, then the frequency template is a band-pass filter. This structure requires more work, in particular the relevant pulsations located after the pulsation where the $\mathcal{H}_\infty$-norm is reached have to be known. This can be obtained through Algorithm 8 of Section 7.3.4 by modifying the successive intervals of research in order to look for higher pulsations than the first one instead of lower ones. Then, G_1 and G_2 can be designed separately in a similar way to what is done in Section 7.3.4 and the final frequency template G could be adjusted at the end in order to decrease the disturbances caused by G_2 on G_1 in low frequency and conversely.

Obviously, both structures can be combined and additional layers can be added to the frequency template (11.2), in that case however, a more rigorous approach for the determination of the parameters may be required. This is discussed in the next section.

Alternative formulation for the construction of frequency templates

A more rigorous approach of the problem, based on optimisation, could be considered and may lead to a much more tight frequency templates.

In particular, the $\mathcal{H}_{2,\Omega}$-norm ban be used to translate the fact that the frequency template $\mathbf{G}$ must be tight in some frequency interval Ω (possibly $[0, \infty)$). The problem would then consists in finding the frequency template $\mathbf{G}$ whose transfer matrix $G(s)$ is the solution of

$$\begin{array}{rcl} G & = & \arg\min \|\tilde{G}\|^2_{\mathcal{H}_{2,\Omega}} \\ & & \text{s.t.} \\ & & \tilde{G} \in \mathcal{G} \\ & & \left\|\tilde{G}(j\omega)\right\|_F \geq \|H(j\omega)\|_F \,, \forall \omega \in \mathbb{R}_+, \end{array} \tag{11.3}$$

where $\mathcal{G}$ represents the space of frequency templates with a defined structure (stable, fixed order, minimum phase, etc.). In addition, problem (11.3) may be completed with a constraint on the $\mathcal{H}_\infty$-norm of $\tilde{G}(s)$ in order to avoid any unwanted local resonance or unbounded behaviour outside of Ω. Note that if $\tilde{G}$ in equation (11.3) is given the same structure as in Section 7.3.4, then the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm for models with high order poles developed in Section 7.2.1 may be used and since the Jordan block is directly known, it does not represent an issue here.

The inequality constraint on the Frobenius norm might be difficult to fulfil. Indeed, unlike in Section 7.3.4 and depending on the structure chosen for $G(s)$, it cannot necessarily be restricted to some relevant pulsations ω_k in order to be true for all $\omega \in \mathbb{R}_+$ especially if resonant modes are involved.

A valid frequency template could certainly be obtained by solving an adequate $\mathcal{H}_\infty$ problem, yet in an approximation context, $\mathbf{H}$ is of large-scale, hence (11.3) seems to be a more modest and reachable objective. Alternatively, the bound Γ_Ω could be used instead of the $\mathcal{H}_{2,\Omega}$-norm in problem (7.3.4).

11.3.2 Other possible applications for the $\mathcal{H}_\infty$ bound Γ_Ω

Since Γ_Ω upper bounds the $\mathcal{H}_\infty$-norm and is faster to compute than the latter, it could be used as a surrogate when the computation of the $\mathcal{H}_\infty$-norm becomes an issue. In particular

- for model approximation, one could look for the reduced-order model $\hat{\mathbf{H}}$ whose transfer matrix $\hat{H}(s)$ solves
$$\hat{H} = \arg \min_{rank(G)=r} \Gamma_\Omega \left(H(j\omega) - G(j\omega) \right),$$

- for robust control of large-scale models, whether Γ_Ω can represent a valid replacement of the $\mathcal{H}_\infty$-norm in tools such as HIFOO [Burke et al., 2006] should be investigated. This would enable large-scale control problem such as the one presented in Section 9.2.3 to be handled directly.

Some preliminary experiments have been conducted during this study concerning the second point but they do not enable to conclude on the validity of the approach, yet.

11.4 First-order optimality conditions for the $\mathcal{H}_{2,\Omega}$ approximation problem

In Chapter 8, the $\mathcal{H}_{2,\Omega}$-norm of the approximation error between a large-scale model $\mathbf{H}$ and a reduced-order one $\hat{\mathbf{H}}$ has been expressed with the poles and residues of both models under the assumption that both models have semi-simple poles only. Then, the first-order optimality conditions of the optimal $\mathcal{H}_{2,\Omega}$ approximation problem have been derived with respect to the poles and residues of the reduced-order model.

Both the approximation error and its gradient can efficiently be computed through elementary operations once the poles and residues of both models are available. This point is particularly interesting in an optimisation context which usually involves many evaluations of the objective function and its gradient.

Moreover, it has also been shown how these first-order optimality conditions can be formulated as interpolation conditions similar to those arising in the optimal $\mathcal{H}_2$ approximation problem (see Section 4.1.2). Indeed, if $\hat{\mathbf{H}}$ is a r-th order minimum of the $\mathcal{H}_{2,\Omega}$ approximation

problem, then, for $m = 1, \ldots, r$,

$$\begin{array}{rcl} T(-\hat{\lambda}_m)\hat{\mathbf{b}}_\mathbf{m}^T & = & \hat{T}(-\hat{\lambda}_m)\hat{\mathbf{b}}_\mathbf{m}^T \\ \hat{\mathbf{c}}_\mathbf{m}T(-\hat{\lambda}_m) & = & \hat{\mathbf{c}}_\mathbf{m}\hat{T}(-\hat{\lambda}_m) \\ \hat{\mathbf{c}}_\mathbf{m}T'(-\hat{\lambda}_m)\hat{\mathbf{b}}_\mathbf{m}^T & = & \hat{\mathbf{c}}_\mathbf{m}\hat{T}'(-\hat{\lambda}_m)\hat{\mathbf{b}}_\mathbf{m}^T, \end{array} \tag{11.4}$$

where $\hat{\lambda}_m$, $\{\hat{\mathbf{c}}_\mathbf{m}, \hat{\mathbf{b}}_\mathbf{m}\}$ are the poles and associated residues of the reduced-order model and

$$T(s) = H_\omega(s) + H(s)\mathbf{a}_{\omega,-s}, \qquad \hat{T}(s) = \hat{H}_\omega(s) + \hat{H}(s)\mathbf{a}_{\omega,-s}.$$

with

$$H_\omega(s) = \sum_{i=1}^{n} \frac{\Phi_i}{s - \lambda_i}\mathbf{a}_{\omega,\lambda_i}, \qquad \hat{H}_\omega(s) = \sum_{i=1}^{r} \frac{\hat{\Phi}_i}{s - \hat{\lambda}_i}\mathbf{a}_{\omega,\hat{\lambda}_i},$$

and $\mathbf{a}_{\omega,s} = \frac{2}{\pi}\mathbf{atan}(\frac{\omega}{s}) := \frac{1}{j\pi}(\mathbf{log}(1 + j\frac{\omega}{s}) - \mathbf{log}(1 - j\frac{\omega}{s}))$.

The interpolation conditions (11.4) have not been exploited in this study and their meaning is still not clear but they represent an interesting lead for future researches in frequency-limited model approximation. In particular, the elements that need to be clarified are

- the link between the interpolation conditions (11.4) and the ones derived for the optimal frequency-weighted $\mathcal{H}_2$ model approximation problem in [Anić et al., 2013] (see Section 5.1.2),
- whether conditions (11.4) can be fulfilled efficiently. Krylov subspaces are unlikely to be applicable since this approach would require a realisation for $T(s)$. The Loewner framework [Antoulas et al., 2012; Ionita, 2013] seems more suited since it enables to build a model of fixed order from frequency data. Still, it would only provide an approximate realisation for $\hat{T}(s)$ which is not the reduced-order model transfer function $\hat{H}(s)$.

11.5 DARPO

In Chapter 9, a descent algorithm aimed at finding a local minimum for the optimal $\mathcal{H}_{2,\Omega}$ approximation problem has been developed. The algorithm relies on a quasi-Newton method based on the BFGS update to find poles and residues of a reduced-order model which satisfy the first-order optimality conditions.

The approximation method has been illustrated on various examples and in particular, it has been used to demonstrate one process that can be followed to address the control of a large-scale model using the robust control framework.

Since **DARPO** is based on the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm, it inherits its limitations concerning *(i)* the Jordan block structure of the models and *(ii)* the applicability of the method to very large-scale models. Both points are discussed in Section 11.5.1. Several improvement can also be considered for the optimisation process, they are discussed in Section 11.5.2. Then in Section 11.5.3, an alternative use of **DARPO** is presented. Finally, the future developments concerning the application of **DARPO** for the control of large-scale models are presented in Section 11.5.4.

11.5.1 Issues related to the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm

Presence of Jordan blocks in the reduced-order model

The presence of high order eigenvalues in the large-scale model has been discussed in Section 11.2.1, here the presence of Jordan blocks in the reduced-order model is discussed.

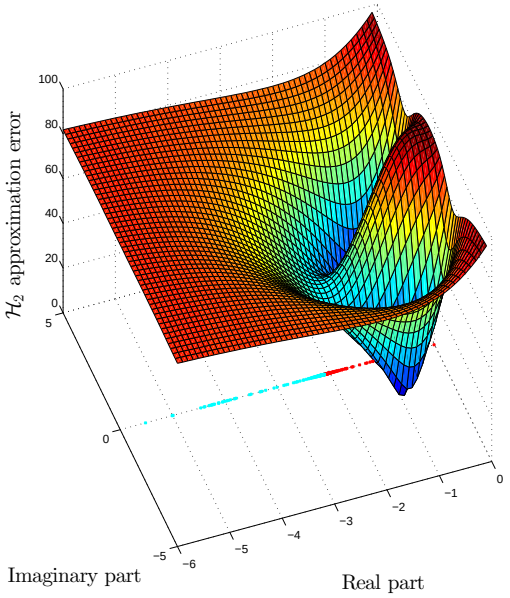


Figure 11.4: Apparition of a Jordan block in the reduced-order model with **DARPO**.

Jordan blocks can appear in the reduced-order model when two poles cross each other. In general, this does not happen, but one can build examples in which it should happen. In particular, if a second-order model with a pair of complex conjugate poles is approximated to an order 2 and that **DARPO** is initialised with a second-order model with real poles. The corresponding $\mathcal{H}_2$ approximation error is represented in Figure 11.4 where the successive positions of the poles are represented by blue and red dots.

The expected behaviour is that the poles move along the real axis until they reach each other and at this point, they should split as complex conjugate poles. When the poles are at the exact same location, the semi-simple poles-residues formulation of the approximation error and its gradient are no longer valid. Instead, the generalised formulation must be used.

Here, what happens in practice is that both poles stop moving before they reach the same location. The exact explanation is still not clear, but it is likely that the line-search algorithm does not find a suitable step-length along the real axis.

In theory, using the poles-residues formulation generalised to model with high-order poles would be sufficient to alleviate the issue. Yet in practice, it might be difficult to determine when some poles are close enough to switch (in the algorithmic procedure) to this formulation.

Very-large scale models

In order for the algorithm **DARPO** to be applicable, the *full* eigenvalue decomposition of the initial large-scale model is required. This is an important limitation in the context of model approximation since it is expensive to compute. In addition, the eventual sparsity of the matrices C and B of the large-scale model is lost when changing to the modal realisation of the large-scale model. These issues are inherent to the framework used in the study and cannot fully be alleviated but still, some hints are presented here in order to decrease the complexity associated with the algorithm.

The idea here relies on the same assumption as in Section 11.2.2, *i.e.* that the $\mathcal{H}_{2,\Omega}$-norm of a model can be well estimated by computing only some of its poles and associated residues.

Ω	**DARPO**	**sDARPO**	Modal trunc.
[0 10]	7.87	13.48(6)	13.68
[0 20]	16.74	17.07(12)	28.99
[0 34]	12.74	13.21(18)	15.33
[0 60]	19.43	19.48(36)	22.26

Table 11.1: Relative $\mathcal{H}_{2,\Omega}$ errors (in %) obtained with **DARPO**, **sDARPO** and the modal truncation used as initial point in the former case after reduction of the LAH model to an order 6 for several Ω.

Hence, instead of approximating directly the large-scale model $\mathbf{H}$, one would approximate an intermediate model $\mathbf{H}_i$ which is the m-th order ($r < m < n$) modal truncation of $\mathbf{H}$ such that

$$\|H_i\|_{\mathcal{H}_{2,\Omega}} \approx \|H\|_{\mathcal{H}_{2,\Omega}}.$$

As mentioned in Section 11.2.2, such an intermediate model could be obtained with an algorithm similar to the Dominant Poles Algorithm [Rommes and Martins, 2006] which takes advantage of the sparsity of matrices.

To illustrate the idea, let us approximate the LAH model with **DARPO** over $\Omega = [0, \omega]$ to an order 6 using either all the eigenvalues and their corresponding eigenvectors or only the eigenvalues which have a magnitude below ω. When only some eigenvalues are used, the algorithm is denoted **sDARPO** for **sparse DARPO**. As the error given by **sDARPO** is not exact, it is recomputed afterwards.

The relative errors obtained with *(i)* the modal approximation used as initial reduced-order model in **DARPO**, *(ii)* **DARPO** and *(iii)* **sDARPO**, are reported in Table 11.1, where the number of eigenvalues used with **sDARPO** is indicated in parenthesis. Note that the modal truncation used as initial reduced-order model represents the reduced-order model that would have been obtained if a modal truncation had been performed directly to the order r. It is presented here to illustrate that using an intermediate model is beneficial for the final reduced-order model.

One can observe that even if only some eigenvalues are used, the optimisation process can still improve the reduced-order model in comparison to a direct modal truncation. For some frequency intervals, the error obtained with **sDARPO** is even comparable to the one obtained with **DARPO**.

Such an approaches raises two main questions,

- first, and as in Section 11.2.2, how the error induced by the intermediate approximation step can be quantified?
- secondly, does minimising the $\mathcal{H}_{2,\Omega}$ approximation error between the reduced-order model and the intermediate one necessarily leads to a decrease of the error between the reduced-order model and the initial large-scale one?

11.5.2 Optimisation scheme

The optimisation scheme used in **DARPO** is already efficient but several improvements can be considered :

- in order to decrease the number of optimisation variables, the idea of optimising separately the poles and the residues as it is done in [Marmorat and Olivi, 2002] and [Beattie and Gugercin, 2009] should be used.
- In order to formally guarantee the stability of the reduced-order model, constrained optimisation should be used instead of unconstrained optimisation. This implies to determine

how the constrained optimisation methods existing for the optimisation of real functions of real variables are modified when complex variables are considered.

Using constrained optimisation would also enable to add additional constraints to the $\mathcal{H}_{2,\Omega}$ approximation problem such as *(i)* matching the static gain of the large-scale model when $\Omega = [0, \omega]$ *(ii)* preventing its behaviour outside of Ω from exploding, or *(iii)* preventing the reduced-order model poles to move too far where the approximation error becomes flat (see Example 17).

- As in the majority of non-linear and non-convex optimisation problems, the initialisation step is extremely important in **DARPO**. While the current one based on the modal truncation of the large-scale model is simple and quite efficient in general, it tends to perform poorly when the frequency of interval does not contains enough relevant poles. Indeed, in those cases, some poles outside the interval are selected while they do not necessarily enable to decrease the approximation error. Hence the initialisation process should be investigated more in depth. A possible lead it to use the Loewner framework [Ionita, 2013] to build a reduced-order model that interpolates the initial one in the considered frequency interval.

More radically, an approach different from a local optimisation algorithm may be worth considering. Due to the non-linearity of the problem and the presence of many local minima, evolutionary, stochastic or hybrid methods could be used, either partially, *e.g.* in the initialisation step, or completely for determining the poles of the reduced-order model.

11.5.3 Approximation for a fixed error

As it has been formulated in this study, the model approximation problem consists in finding a r-th order model $\hat{\mathbf{H}}$ from a large-scale model $\mathbf{H}$ of order $n \gg r$. Yet, from a practical point of view, the approximation order r is merely a constraint and what is really interesting is rather to know whether the reduced-order model is sufficiently accurate, or if it can be used safely for control purpose, etc.

Concerning the accuracy of the reduced-order model, **DARPO** can actually be used in a constructive way, the approximation order r being increased until some objective $\mathcal{H}_{2,\Omega}$ approximation error is reached. The way the feature can be used is shown in Appendix C.9.

For instance, let us consider the LAH model, one would like to find a reduced-order model $\hat{\mathbf{H}}$ such that the relative $\mathcal{H}_{2,\Omega}$ approximation error over $\Omega = [0, 20]$ is below 5%. The decrease of the relative error and the frequency responses of the large-scale and final reduced-order models are plotted in Figure 11.5. Note that since the LAH model has only complex poles, the approximation order is increased by steps of 2.

One can observe that **DARPO** stops after 490 iterations (achieved in $\sim$ 4 sec). The final approximation order is 14 and the final error is 0.051%. Indeed, the optimisation for the current approximation order r is not stopped even if the objective is met since the reduced-order model can be improved even further with that order. Here, it means that an approximation order of 12 has led to an error above 5% while the order 14 enables to decrease the error to 0.051%. This is probably due to the new parameters selection strategy considered here.

Indeed, each time the approximation order is increased, new poles and residues must be added to the current model. This is done by taking into account the initial ranking of the poles obtained with the initialisation criterion $\mathcal{J}_{init}^{(2)}$ from Section 9.1.3 which does not take into account the current state of the reduced-order model and whether the new poles enable to decrease the approximation error or not.

This approach was also developed in order to see if starting from a small order and increasing it by steps could lead, at the end, to a better reduced-order model than the one obtained by directly approximating at the final order. Yet this is not the case in the current state of the approach. Indeed, the model obtained here is the same as the one that would have been obtained

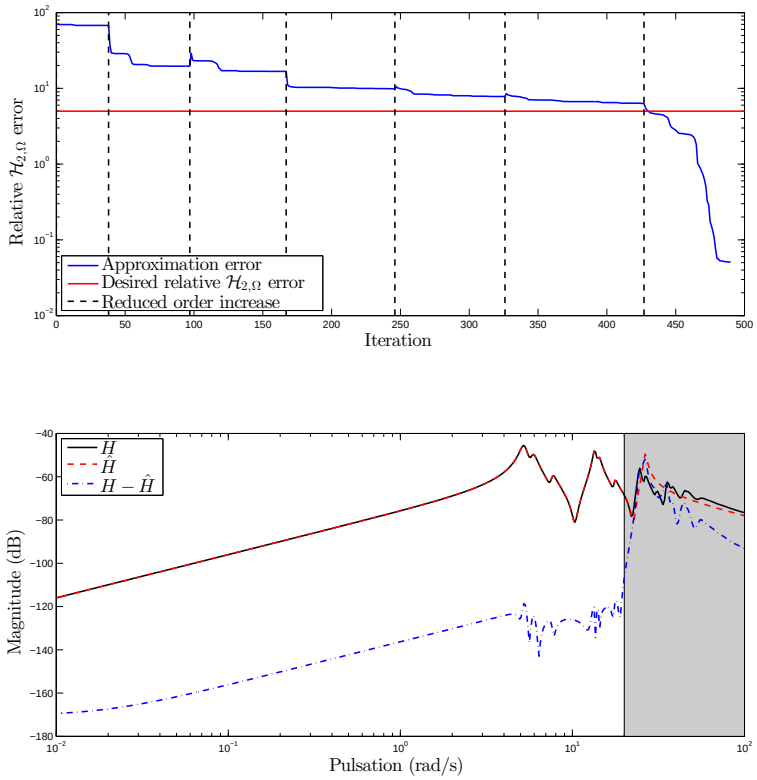


Figure 11.5: Approximation of the LAH model over $\Omega = [0, 20]$ for a fixed relative approximation error of 5%.

by approximating the LAH model directly to an order 14. This is probably due to the selection strategy used for the new poles and residues which could greatly be improved.

A constructive process is also used in [Marmorat and Olivi, 2002], and for each increase of the approximation order, several guesses are built from the current reduced-order model. A similar idea should be developed here.

11.5.4 Model approximation for the control of large-scale models

In Section 9.2.3, **DARPO** has been used in order to design a simple controller to perform disturbance rejection on a large-scale model. The underlying process consists in completing the reduced-order model by a low complexity uncertainty which represents (and bounds) the approximation error and to use the robust control framework to design a controller on this low-order uncertain model.

This approach is promising but several points need to be clarified, especially

- the influence of the approximation order r is not clear and an indication of the right order required for the design of a controller is missing,
- the influence of the approximation interval Ω is not clear either. Indeed, the impact of the frequency interval lies mainly in the initial reduced-order model selection which set the

dynamics that will be preserved. Hence, the real, non-trivial, question is to determine the dynamics that matter for the closed-loop.

A completely different approach would consists in using the gap or ν-gap metrics [Vidyasagar, 1984; Vinnicombe, 1992] as criterion for the approximation as it has been done in [Sootla, 2011]. Yet, the resulting approximation problem is difficult and whether it can be solved efficiently on large-scale models is not clear due to the complexity of the evaluation of this norm which is based on the $\mathcal{H}_\infty$ one.

Conclusion

The main limitation of the proposed model approximation method lies in the use of the poles-residues formulation of the models and of the approximation error. Indeed, this formulation requires the full eigen-decomposition of the large-scale model and becomes ill-conditioned if high order poles are involved. Hence, a formulation of the state-space may appear more suitable, however *(i)* the computational complexity does not vanish since the eigenvalue decomposition is replaced by the need to solve various Sylvester equations, which is a complex task when performed exactly, *(ii)* in the frequency-limited context, the logarithm of a matrix as well as its derivative are required, which are not trivial to compute either. The core of the problem when using an optimisation algorithm to perform large-scale model reduction lies in the computation of the approximation error and its gradient which reduces here to the computation of the inverse $(sI_n - A)^{-1}$ arising in the transfer function (and thus in the $\mathcal{H}_{2,\Omega}$-norm).

As pointed out in [Moler and Van Loan, 1978], a trade-off must be considered between the eigen-decomposition (or Jordan form) which offers the easiest, but most ill-conditioned, way to compute a matrix function, and another decomposition which is well-conditioned but does not yield a particularly easy computation of the function (such as the Schur decomposition). In [Moler and Van Loan, 1978], a block diagonal matrix where nearly confluent eigenvalues are clustered is presented as an interesting candidate. Such a decomposition may be considered for the computation of the $\mathcal{H}_{2,\Omega}$-norm and in the optimal $\mathcal{H}_{2,\Omega}$ approximation problem in future researches.

Besides, the other interesting lead for further development lies in the first-order optimality conditions expressed as interpolation conditions. Indeed, generalising interpolation methods such as the **IRKA** to the optimal $\mathcal{H}_{2,\Omega}$ approximation would be particularly interesting given the performance and simplicity of the approach. In the next chapter, some longer term perspectives are pointed out.

Chapter 12

Outlook

The perspectives of this thesis can be considered along two major lines : model approximation and control of large-scale models. The former concerns the extension and further development of the proposed methods while the latter focuses rather on the adaptation and application of these methods for the control of large-scale models.

More specifically, concerning model approximation, and in addition of the technical points raised in the previous chapter :

- the different contributions of this thesis mainly result of the formulation of the $\mathcal{H}_{2,\Omega}$-norm with the poles and residues of the transfer function associated with a LTI dynamical model. Whether this formulation can be extended to other types of models should be investigated. In particular, it may be extended to descriptor models by using the *Weierstrass canonical form* of a matrix pencil and to some time-delays models by using the *Lambert function* to compute the poles of the transfer function.
- More generally, $\mathcal{H}_{2,\Omega}$ model approximation represents the natural generalisation of the $\mathcal{H}_2$ model approximation and it should be extended to other types of models such as LPV models, bilinear models (such as in [Shaker and Tahavori, 2014]), etc.

Concerning the control of large-scale models :

- the global process for the control of large-scale model used in Chapter 10 may be simplified by merging the approximation and interpolation steps. To that goal, multi-model approximation, *i.e.* low-order LPV modelling, could be investigated.
- Alternatively, a constructive approach which successively include additional dynamics of the large-scale model until a satisfactory controller is found may also be considered.
- In this thesis, only robust control has been considered, but other control methods such as predictive or optimal control may also benefit from the use of model approximation methods to alleviate the numerical burden associated with large-scale models.

The model approximation problem is far from being solved and the increasing use of numerical tools for the modelling of complex dynamical systems is likely to make this subject a central element for the practical implementation of theoretical tools.

Part V

Appendix

Appendix A

Proofs related to the poles-residues expression of the $\mathcal{H}_{2,\Omega}$-norm

A.1 Poles-residues expression of the $\mathcal{H}_{2,\Omega}$-norm in the semi-simple case

Let us consider the Definition of the $\mathcal{H}_{2,\Omega}$-norm of $\mathbf{H}$ with $\Omega = [0, \omega]$,

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} := \mathbf{tr}\left(\frac{1}{2\pi}\int_{-\omega}^{\omega} H(j\nu)H(-j\nu)^T d\nu\right). \tag{A.1}$$

The assumption is made that $H(s)$ can be written as,

$$H(s) = \sum_{i=1}^{n} \frac{\phi_i}{s-\lambda_i} + D, \tag{A.2}$$

where $D \in \mathbb{R}^{n_y \times n_u}$, $\phi_i \in \mathbb{C}^{n_y \times n_u}$ and $\lambda_i \in \mathbb{C}$ $(i = 1, \ldots, n)$. In other words, we assume that H has only semi-simple poles. By noting $\tilde{H}(s) = \sum_{i=1}^{n} \frac{\phi_i}{s-\lambda_i}$ and replacing H by its decomposition (A.2), in (A.1), one obtains

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \frac{1}{2\pi}\mathbf{tr}\left(\int_{-\omega}^{\omega} DD^T + \tilde{H}(j\nu)D^T + D\tilde{H}(-j\nu)^T + \tilde{H}(j\nu)\tilde{H}(-j\nu)^T d\nu\right). \tag{A.3}$$

Each term of this integral is then considered separately:

(i) Considering the first term, it follows that

$$\frac{1}{2\pi}\int_{-\omega}^{\omega} DD^T d\nu = \frac{\omega}{\pi} DD^T. \tag{A.4}$$

(ii) Greater attention should be given to the following (second) integral

$$\frac{1}{2\pi}\int_{-\omega}^{\omega} \tilde{H}(j\nu)D^T d\nu = \frac{1}{2\pi}\sum_{i=1}^{n}\int_{-\omega}^{\omega} \frac{\phi_i}{j\nu-\lambda_i} D^T d\nu.$$

Indeed, if $\mathbf{Re}\,(\lambda_i) = 0$, then $\frac{\phi_i}{j\nu-\lambda_i}$ is integrable over $[-\omega, \omega]$ if and only if $\omega < |\lambda_i|$. Hence, in the sequel, the following assumption is made:

$$\omega < \min\left\{|\lambda_i| / \mathbf{Re}\,(\lambda_i) = 0\right\}. \tag{A.5}$$

It implies that $\frac{\phi_i}{j\nu-\lambda_i}$ is integrable over $[-\omega, \omega]$ for all $i = 1, \ldots, n$. Under this assumption and based on the definition of the complex arctangent (Definition 8), one obtains

$$\begin{aligned}\int_{-\omega}^{\omega} \frac{\phi_i}{j\nu-\lambda_i} D^T d\nu &= j\phi_i\left[\mathbf{log}\,(-j\omega-\lambda_i) - \mathbf{log}\,(j\omega-\lambda_i)\right] D^T \\ &:= -2\phi_i \mathbf{atan}\left(\frac{\omega}{\lambda_i}\right) D^T.\end{aligned}$$

Thus

$$\frac{1}{2\pi}\int_{-\omega}^{\omega}\tilde{H}(j\nu)D^T d\nu = -\frac{1}{\pi}\sum_{i=1}^{n}\phi_i D^T \mathbf{atan}\left(\frac{\omega}{\lambda_i}\right). \tag{A.8}$$

(iii) In a similar way

$$\frac{1}{2\pi}\int_{-\omega}^{\omega}D\tilde{H}(-j\nu)^T d\nu = -\frac{1}{\pi}\sum_{i=1}^{n}D\phi_i^T \mathbf{atan}\left(\frac{\omega}{\lambda_i}\right). \tag{A.9}$$

(iv) Regarding the last term of (A.3), one has

$$\tilde{H}(j\nu)\tilde{H}(-j\nu)^T = \sum_{i=1}^{n}\sum_{k=1}^{n}\frac{\phi_i\phi_k^T}{(j\nu-\lambda_i)(-j\nu-\lambda_k)}$$

and it follows that

$$\frac{1}{2\pi}\int_{-\omega}^{\omega}\tilde{H}(j\nu)\tilde{H}(-j\nu)^T d\nu = \sum_{i=1}^{n}\sum_{k=1}^{n}\underbrace{\frac{1}{2\pi}\int_{-\omega}^{\omega}\frac{\phi_i\phi_k^T}{(j\nu-\lambda_i)(-j\nu-\lambda_k)}d\nu}_{a_{i,k}}.$$

From here, two cases must be considered :

(a) If $\lambda_i+\lambda_k\neq 0$, then

$$a_{i,k} = \frac{1}{2\pi}\int_{-\omega}^{\omega}\frac{\phi_i\phi_k^T}{(j\nu-\lambda_i)(-j\nu-\lambda_k)}d\nu = \frac{1}{2\pi}\int_{-\omega}^{\omega}\left(\frac{p_{i,k}}{j\nu-\lambda_i}+\frac{p_{i,k}}{-j\nu-\lambda_k}r\nu\right),$$

with $p_{i,k} = -\frac{\phi_i\phi_k^T}{\lambda_i+\lambda_k}$. Because of assumption (A.5), each term can then be integrated as previously and one gets

$$a_{i,k} = \frac{1}{\pi}\frac{\phi_i\phi_k^T}{\lambda_i+\lambda_k}\mathbf{atan}\left(\frac{\omega}{\lambda_i}\right)+\frac{1}{\pi}\frac{\phi_i\phi_k^T}{\lambda_i+\lambda_k}\mathbf{atan}\left(\frac{\omega}{\lambda_k}\right).$$

Since $p_{i,k} = p_{k,i}^T$, the sums can be reordered as follows

$$a_{i,k} = \frac{1}{\pi}\frac{\phi_i\phi_k^T+\phi_k\phi_i^T}{\lambda_i+\lambda_k}\mathbf{atan}\left(\frac{\omega}{\lambda_i}\right). \tag{A.10}$$

(b) If $\lambda_i+\lambda_k=0$, then

$$a_{i,k} = \frac{1}{2\pi}\int_{-\omega}^{\omega}\frac{\phi_i\phi_k^T}{(j\nu-\lambda_i)(-j\nu-\lambda_k)}d\nu = \frac{1}{2\pi}\int_{-\omega}^{\omega}-\frac{\phi_i\phi_k^T}{(j\nu-\lambda_i)^2}d\nu.$$

After integration, one obtains

$$\begin{aligned} a_{i,k} &= \frac{1}{2\pi}\left(-j\frac{\phi_i\phi_k^T}{j\omega-\lambda_i}+j\frac{\phi_i\phi_k^T}{-j\omega-\lambda_i}\right) \\ &= -\frac{1}{\pi}\frac{\omega\phi_i\phi_k^T}{w^2+\lambda_i\lambda_i}. \end{aligned} \tag{A.11}$$

Taking the trace of (A.4), (A.8), (A.9) and (A.10) (or (A.11)) and adding them up leads to the result of Theorem 10.

A.2 Poles-residues expression of the $\mathcal{H}_{2,\Omega}$-norm for models with higher order poles

Let us consider an asymptotically stable and strictly proper MIMO LTI dynamical model $\mathbf{H}$ with n_b eigenvalues λ_i of multiplicity n_i described by its transfer function $H(s)$. The $\mathcal{H}_{2,\Omega}$-norm of $\mathbf{H}$ is defined as,

$$\|H\|^2_{\mathcal{H}_{2,\Omega}} = \frac{1}{2j\pi}\int_{-j\omega}^{j\omega} \mathbf{tr}\left(H(s)H(-s)^T\right) ds. \tag{A.12}$$

By replacing $H(s)$ by its partial fraction expansion (2.2), one obtains

$$\mathbf{tr}\left(H(s)H(-s)^T\right) = \sum_{i=1}^{n_b}\sum_{k=1}^{n_b}\sum_{l=1}^{n_i}\sum_{m=1}^{n_j} \frac{\mathbf{tr}\left(\Phi_i^{(l)}\Phi_k^{(m)T}\right)}{(s-\lambda_i)^l\,(-s-\lambda_k)^m}.$$

The integral (A.12) comes down to the following integrals for each i, k, l and m,

$$\int_{-j\omega}^{j\omega} \frac{\mathbf{tr}\left(\Phi_i^{(l)}\Phi_k^{(m)T}\right)}{(s-\lambda_i)^l\,(-s-\lambda_k)^m} ds = \mathbf{tr}\left(\Phi_i^{(l)}\Phi_k^{(m)T}\right) \underbrace{\int_{-j\omega}^{j\omega} f_{iklm}(s) ds}_{I_{lm}(\lambda_i,\lambda_k,\omega)}.$$

By noticing that the functions f_{iklm} has two poles, λ_i and $-\lambda_k$, of order l and m, respectively, their partial fraction decomposition are given by

$$f_{iklm}(s) = \sum_{p=1}^{l}\frac{a_p}{(s-\lambda_i)^p} + \sum_{q=1}^{m}\frac{b_q}{(-s-\lambda_k)^q}, \tag{A.13}$$

where

$$a_p = \frac{1}{(l-p)!}\frac{d^{l-p}}{ds^{l-p}}\left.(s-\lambda_i)^l\, f_{iklm}(s)\right|_{s=\lambda_i}$$

for $p = 1,\ldots,l$ and

$$b_q = (-1)^{m-q}\frac{1}{(m-q)!}\frac{d^{m-q}}{ds^{m-q}}\left.(-s-\lambda_k)^m\, f_{iklm}(s)\right|_{s=-\lambda_k},$$

for $q = 1,\ldots,m$. Note that the sign $(-1)^{m-q}$ is introduced due to the specific form of the partial fraction decomposition (A.13) which uses $\frac{1}{(-s-\lambda_k)^q}$ instead of $(-1)^q\frac{1}{(s+\lambda_k)^q}$.

The residues a_p and b_q can be written in similar forms. Indeed $a_p = r_{m,l-p}(\lambda_i,\lambda_k)$ and $b_q = r_{l,m-q}(\lambda_i,\lambda_k)$ where

$$r_{u,v}(\lambda_i,\lambda_k) = (-1)^{(u+v)}\begin{pmatrix} u+v-1 \\ v \end{pmatrix}\frac{1}{(\lambda_i+\lambda_k)^{u+v}}.$$

Since the system $\mathbf{H}$ is assumed to be asymptotically stable, each integral composing $I_{lm}(\lambda_i,\lambda_k,\omega)$ can be directly calculated. Indeed,

$$\begin{aligned}\int_{-j\omega}^{j\omega}\frac{a_1}{s-\lambda_i} &= a_1\left[\mathbf{log}\,(s-\lambda_i)\right]_{-j\omega}^{j\omega},\\ \int_{-j\omega}^{j\omega}\frac{a_2}{(s-\lambda_i)^2} &= a_2\left[-\tfrac{1}{s-\lambda_i}\right]_{-j\omega}^{j\omega},\end{aligned}$$

and so on for each value of $p = 1,\ldots,l$ and $q = 1,\ldots,m$. Finally, the resulting functions of ω can be written in a more convenient way as $W_{p-1}(j\omega,\lambda_i)$ and $W_{q-1}(j\omega,\lambda_k)$ where

$$W_u(j\omega,\lambda) = \frac{1}{u!}\frac{d^u}{dy^u}\left.\left(\mathbf{log}\,(-j\omega-y) - \mathbf{log}\,(j\omega-y)\right)\right|_{y=\lambda}.$$

A.3 Poles-residues expression of the $\mathcal{H}_{2,\Omega}$ approximation error

The proof comes from writing the $\mathcal{H}_{2,\Omega}$-norm of the error in terms of the residues $\tilde{\Phi}_i$ and poles $\tilde{\lambda}_i$, $i = 1, \ldots, n+r$ of transfer function $\tilde{H}$ of the error system $\tilde{\mathbf{H}} = \mathbf{H} - \hat{\mathbf{H}}$,

$$\mathcal{J}_{\mathcal{H}_{2,\Omega}} = \sum_{i=1}^{n+r}\sum_{k=1}^{n+r} \frac{\mathbf{tr}\left(\tilde{\Phi}_i \tilde{\Phi}_k^T\right)}{\tilde{\lambda}_i + \tilde{\lambda}_k} \mathbf{a}_{\omega,\tilde{\lambda}_i} - \sum_{i=1}^{n+r} \mathbf{tr}\left(\tilde{\Phi}_i \tilde{D}^T\right) \mathbf{a}_{\omega,\tilde{\lambda}_i} + \frac{1}{\pi}\omega \mathbf{tr}\left(\tilde{D}\tilde{D}^T\right). \tag{A.14}$$

The poles/residues of $\tilde{H}$ are composed of the poles/residues of H and $-\hat{H}$. By supposing they are ordered as

$$\begin{aligned}
\tilde{\lambda}_i &= \begin{cases} \lambda_i & \text{if } i = 1, \ldots, n \\ \hat{\lambda}_i & \text{if } i = n+1, \ldots, n+r, \end{cases} \\
\tilde{\Phi}_i &= \begin{cases} \Phi_i & \text{if } i = 1, \ldots, n \\ -\hat{\Phi}_i & \text{if } i = n+1, \ldots, n+r, \end{cases}
\end{aligned}$$

the sums in (A.14) can then be divided as follow,

$$\begin{aligned}
\mathcal{J}_{\mathcal{H}_{2,\Omega}} = & \sum_{i=1}^{n}\sum_{k=1}^{n} \frac{\mathbf{tr}\left(\Phi_i \Phi_k^T\right)}{\lambda_i + \lambda_k} \mathbf{a}_{\omega,\lambda_i} - \sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i \tilde{D}^T\right) \mathbf{a}_{\omega,\lambda_i} \ldots \\
& + \sum_{i=1}^{n}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(-\Phi_i \hat{\Phi}_k^T\right)}{\lambda_i + \hat{\lambda}_k} \mathbf{a}_{\omega,\lambda_i} + \sum_{i=1}^{r}\sum_{k=1}^{n} \frac{\mathbf{tr}\left(-\hat{\Phi}_i \Phi_k^T\right)}{\hat{\lambda}_i + \lambda_k} \mathbf{a}_{\omega,\hat{\lambda}_i} \ldots \\
& + \sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\hat{\Phi}_i \hat{\Phi}_k^T\right)}{\hat{\lambda}_i + \hat{\lambda}_k} \mathbf{a}_{\omega,\hat{\lambda}_i} + \sum_{i=1}^{r} \mathbf{tr}\left(\hat{\Phi}_i \tilde{D}^T\right) \mathbf{a}_{\omega,\hat{\lambda}_i} \ldots \\
& + \frac{1}{\pi}\omega \mathbf{tr}\left(\tilde{D}\tilde{D}^T\right).
\end{aligned}$$

Finally, regrouping the cross sums leads to the result,

$$\begin{aligned}
\mathcal{J}_{\mathcal{H}_{2,\Omega}} = & \sum_{i=1}^{n}\sum_{k=1}^{n} \frac{\mathbf{tr}\left(\Phi_i \Phi_k^T\right)}{\lambda_i + \lambda_k} \mathbf{a}_{\omega,\lambda_i} - \sum_{i=1}^{n} \mathbf{tr}\left(\Phi_i \tilde{D}^T\right) \mathbf{a}_{\omega,\lambda_i} \ldots \\
& - \sum_{i=1}^{n}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\Phi_i \hat{\Phi}_k^T\right)}{\lambda_i + \hat{\lambda}_k} \left(\mathbf{a}_{\omega,\lambda_i} + \mathbf{a}_{\omega,\hat{\lambda}_k}\right) + \frac{1}{\pi}\omega \mathbf{tr}\left(\tilde{D}\tilde{D}^T\right) \ldots \\
& + \sum_{i=1}^{r}\sum_{k=1}^{r} \frac{\mathbf{tr}\left(\hat{\Phi}_i \hat{\Phi}_k^T\right)}{\hat{\lambda}_i + \hat{\lambda}_k} \mathbf{a}_{\omega,\hat{\lambda}_i} + \sum_{i=1}^{r} \mathbf{tr}\left(\hat{\Phi}_i \tilde{D}^T\right) \mathbf{a}_{\omega,\hat{\lambda}_i}.
\end{aligned}$$

Appendix B

Elements about the principal value of the complex arctangent

B.1 Difference between the two definitions of the inverse tangent function

The principal value of the complex inverse tangent function can either be defined as

$$\mathbf{atan}(z) = \frac{1}{2j}\left(\mathbf{log}\left(1+jz\right) - \mathbf{log}\left(1-jz\right)\right), \tag{B.1}$$

or as

$$\mathbf{atan}(z) = \frac{1}{2j}\left(\mathbf{log}\left(\frac{1+jz}{1-jz}\right)\right). \tag{B.2}$$

The two definitions yield the same branch cuts $(-j\infty, -j]\cup[j, j\infty)$ but do not coincide on $(-j\infty, -j[$ (see [Haber, 2012] for the proof). This is illustrated in Figure B.1 where the real part, imaginary part and absolute value of both functions are plotted with respect to the real and imaginary parts of $z \in \mathbb{C}$ and where the red (respectively black) line represents the imaginary (respectively real) axis.

B.2 Limit of atan $\left(\frac{\omega}{\lambda}\right)$ as ω tends towards infinity

By definition, $\mathbf{atan}\left(\frac{\omega}{\lambda}\right)$, $\lambda \in \mathbb{C}^*$ can be written in terms of the principal value of the complex logarithm, *i.e.*

$$\begin{aligned}\mathbf{atan}\left(\frac{\omega}{\lambda}\right) &:= \frac{1}{2j}\left[\mathbf{log}\left(1+j\frac{\omega}{\lambda}\right) - \mathbf{log}\left(1-j\frac{\omega}{\lambda}\right)\right] \\ &:= \frac{1}{2j}\left[\mathbf{ln}\left(\left|1+j\frac{\omega}{\lambda}\right|\right) - \mathbf{ln}\left(\left|1-j\frac{\omega}{\lambda}\right|\right) + j\left(\mathbf{arg}\left(1+j\frac{\omega}{\lambda}\right) - \mathbf{arg}\left(1-j\frac{\omega}{\lambda}\right)\right)\right]\end{aligned} \tag{B.3}$$

By denoting $\lambda = x + jy$, the imaginary part of (B.3) becomes

$$\begin{aligned}\mathbf{ln}\left(\left|1+j\frac{\omega}{\lambda}\right|\right) - \mathbf{ln}\left(\left|1-j\frac{\omega}{\lambda}\right|\right) &= \mathbf{ln}\left(\left|\frac{\lambda+j\omega}{\lambda-j\omega}\right|\right) \\ &= \frac{1}{2}\mathbf{ln}\left(\frac{\omega^2+2\omega y+y^2-x^2}{\omega^2-2\omega y+y^2+x^2}\right),\end{aligned}$$

which tends towards 0 as ω tends towards ∞.

Let us now consider the real part of (B.3) and set the following notations :

$$\begin{aligned}z_1(\omega) &= 1+j\frac{\omega}{\lambda} = \underbrace{1+\frac{\omega}{|\lambda|^2}y}_{a_1(\omega)} + j\underbrace{\frac{\omega}{|\lambda|^2}x}_{b_1(\omega)}, \\ z_2(\omega) &= 1-j\frac{\omega}{\lambda} = \underbrace{1-\frac{\omega}{|\lambda|^2}y}_{a_2(\omega)} + j\underbrace{\frac{-\omega}{|\lambda|^2}x}_{b_2(\omega)}.\end{aligned}$$

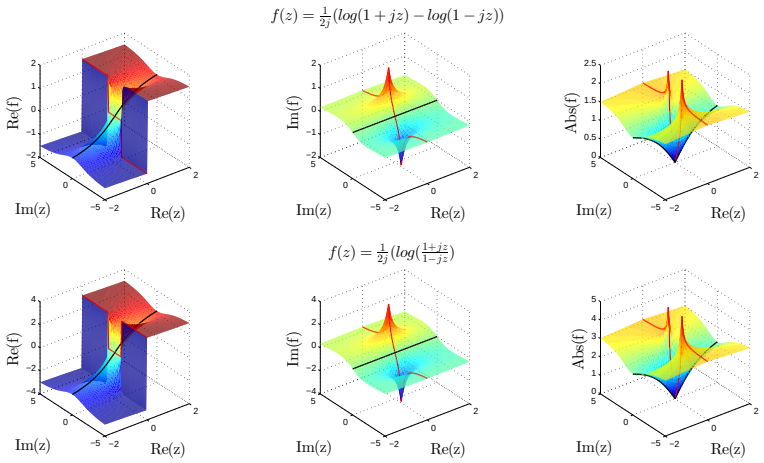


Figure B.1: Real, imaginary and absolute value of the complex inverse tangent function defined as (B.1) (top) or as (B.2) (bottom).

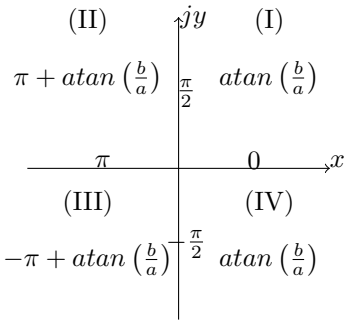


Figure B.2: Principal value of the complex argument $\arg z$ of $z = a + jb$ through the four quadrant inverse tangent function (often denoted *atan2*).

Note that $\forall \omega > 0$, $b_2(\omega) = -b_1(\omega)$ and that $\forall \omega > W = \frac{x^2+y^2}{|y|}$, $\mathbf{sign}\left(a_1(\omega)\right) = -\mathbf{sign}\left(a_2(\omega)\right)$. Hence, for sufficiently large values of ω, $z_1(\omega)$ and $z_2(\omega)$ lie in symmetric (with respect to the origin) quadrants of the complex plane.

Let us denote by (I), (II), (III) and (IV) the four quadrants of the complex plane as represented in Figure B.2. The principal value of the complex argument $\arg z$ can be expressed in different ways depending on the complex quadrant where z lies (see Figure B.2) and since

$$\lim_{\omega\to\infty} atan\left(\frac{b_1(\omega)}{a_1(\omega)}\right) = \lim_{\omega\to\infty} atan\left(\frac{b_2(\omega)}{a_2(\omega)}\right),$$

then one obtains

$$\lim_{\omega\to\infty} \arg z_1(\omega) - \arg z_2(\omega) = \begin{cases} \pi & \text{if } \mathbf{Im}\left(z_1(\omega)\right) > 0 \text{ or } z_1(\omega) \in \mathbb{R}_- \\ -\pi & \text{if } \mathbf{Im}\left(z_1(\omega)\right) < 0 \text{ or } z_1(\omega) \in \mathbb{R}_+ \end{cases}.$$

Note that when $z_1(\omega)$ lies in the upper half-plane (without the real axis) then by applying a conjugation and a clockwise rotation of $\frac{\pi}{2}$, one can see that λ belongs to the right half-plane and if $z_1(\omega)$ lies in the lower half-plane then λ is in the left half-plane, *i.e.*

$$\begin{aligned} \mathbf{Im}\left(z_1(\omega)\right) > 0 &\quad\Leftrightarrow\quad \mathbf{Re}\left(\lambda\right) > 0, \\ \mathbf{Im}\left(z_1(\omega)\right) < 0 &\quad\Leftrightarrow\quad \mathbf{Re}\left(\lambda\right) < 0. \end{aligned}$$

Thus

$$\lim_{\omega\to\infty} \mathbf{atan}\left(\frac{\omega}{\lambda}\right) = \begin{cases} \frac{\pi}{2} & \text{if } \mathbf{Re}\,(\lambda) > 0 \\ -\frac{\pi}{2} & \text{if } \mathbf{Re}\,(\lambda) < 0 \end{cases}$$

To determine the limits on the imaginary axis, let us assume that $\lambda = j\gamma$ with $\gamma \in \mathbb{R}$, then

$$\arg z_1(\omega) - \arg z_2(\omega) = arg(1 + \frac{\omega}{\gamma}) - arg(1 - \frac{\omega}{\gamma}),$$

which tends towards $-\pi$ is $\gamma > 0$ and to π if $\gamma < 0$ as ω tends towards ∞. Finally, the limits of $\mathbf{atan}(\frac{\omega}{\lambda})$ are given by

$$\lim_{\omega\to\infty} \mathbf{atan}\left(\frac{\omega}{\lambda}\right) = \begin{cases} -\frac{\pi}{2} & \text{if } \mathbf{Re}\,(\lambda) < 0 \\ \frac{\pi}{2} & \text{if } \mathbf{Re}\,(\lambda) > 0 \\ \frac{\pi}{2} & \text{if } \mathbf{Re}\,(\lambda) = 0 \text{ and } \mathbf{Im}\,(\lambda) < 0 \\ -\frac{\pi}{2} & \text{if } \mathbf{Re}\,(\lambda) = 0 \text{ and } \mathbf{Im}\,(\lambda) > 0 \end{cases}$$

which is coherent with the limits of $\mathbf{acot}(z) = \mathbf{atan}(\frac{1}{z})$ derived in [Haber, 2012].

Appendix C

Code samples

In this appendix, some code samples are provided so that the reader can reproduce some of the examples presented in this thesis and also see how the routines developed during this thesis can be called through the MORE Toolbox [Poussot-Vassal and Vuillemin, 2012]. Here are the steps to follow for these code samples to work :

- first, download the limited version of the MORE toolbox from `w3.onera.fr/more`. It requires to create an account.
- Then, download the benchmark collection $COMPl_eib$ from `www.complib.de/`. This library contains, among others, the LAH, the CBM and the ISS models that are used as benchmarks and which can be used in the limited version of the MORE Toolbox.
- Get the files `lah.mat`, `cbm.mat` and `iss1_2.mat` from $COMPl_eib$ and add the library as well as the toolbox to the Matlab® path.

The call of the routines are relatively self-explanatory with the embedded commentaries and are therefore not described in depth here, for that purpose, please refer to the documentation of the toolbox on the website. Note that the most recent algorithms developed during this thesis might not be fully documented yet.

C.1 $\mathcal{H}_{2,\Omega}$-norm computation

This code enables to plot Figure 2.3 of Section 2.2.2 that illustrates the behaviour of the $\mathcal{H}_{2,\Omega}$-norm on the LAH model.

```
%% Illustration H2W-norm
% initial model
load lah;
H       = ss(A,B,C,0);
h2      = norm(H);
[M,Ph,W] = bode(H);
% Computation of the H2w norm for varying w
opt.compMethod = 1; % set which method is used to computed the H2W-norm
                    % 1 : the poles-residues formulation (default)
                    % 2 : gramian formulation
ww = logspace(0,2,200);
for i = 1:length(ww);
  % frequency interval Omega = [0 w(i)]
  opt.freqBand = [0,ww(i)];
  % computation of the frequency-limited H2 norm over Omega
  h2w(i)       = moreNorm(H,2,opt);
end
xl = [1,100];
figure
subplot(211)
loglog(W,20*log(abs(M(:))),'b','lineWidth',3)
set(gca,'xlim',xl);
xlabel('Pulsation (rad/s)');
ylabel('Gain (dB)');
title('Frequency response');
subplot(212)
semilogx([ww(1),ww(end)],[h2,h2],'b','lineWidth',3)
```

```
hold on
semilogx(ww,h2w,'r--','lineWidth',3);
set(gca,'xlim',xl);
l = legend('H_2-norm','H_{2,\Omega}-norm');
xlabel('Pulsation (rad/s)');
ylabel('Value of norms')
title('Evolution of the H_{2,\Omega}-norm wrt to \omega')
```

C.2 Example 11 : non-linearity of the optimal $\mathcal{H}_2$ approximation problem

This is the code that has been used to produce the figure of Example 11 aimed at illustrating the non-linearity of the optimal $\mathcal{H}_2$ approximation problem

```
%% Illustration of the non-convexity of the optimal H2 approximation pb.
% load large-scale model
load lah;
H     = ss(A,B,C,0);
normH = norm(H);
% real and imaginary parts of the reduced-order model poles
rp    = linspace(-3,-0.1,30);
ip    = linspace(1,40,40);
% preallocating
error          = zeros(length(rp),length(ip));
squaredNormHr = error;
for i = 1:length(rp)
  for j = 1:length(ip)
    % fix the poles of the reduced-order model
    P(i,j)   = complex(rp(i),ip(j));
    p        = P(i,j);
    % determine the associated optimal residues
    M        = [1/(-2*p)        , 1/(-2*real(p));
                1/(-2*real(p)) , 1/(-2*p')];
    rhs      = [freqresp(H,-p);
                freqresp(H,-p')];
    phi_opt = M\rhs;
    phi      = phi_opt(1);
    % compute the H2 approximation error
    squaredNormHr(i,j)  = -abs(phi)^2/real(p)-real(phi^2/(p));
    error(i,j)          = sqrt(normH^2-squaredNormHr(i,j));
  end
end
surf(real(P),imag(P),error/normH*100)
zlabel('Relative H2 error')
xlabel('Real part')
ylabel('Im. part')
view(55,34)
```

C.3 Use of the FL-ISTIA

In this example, the LAH model is reduced to an order $r = 10$ over $\Omega = [10, 20]$ with the **FL-ISTIA**. The frequency responses of the large-scale model, the reduced order model and the error model are plotted (see Figure C.1) and the relative $\mathcal{H}_{2,\Omega}$-norm of the approximation error computed.

```
%% Example FL-ISTIA
%% Model and approximation parameters
load lah;
H                = ss(A,B,C,0);       % initial model
```

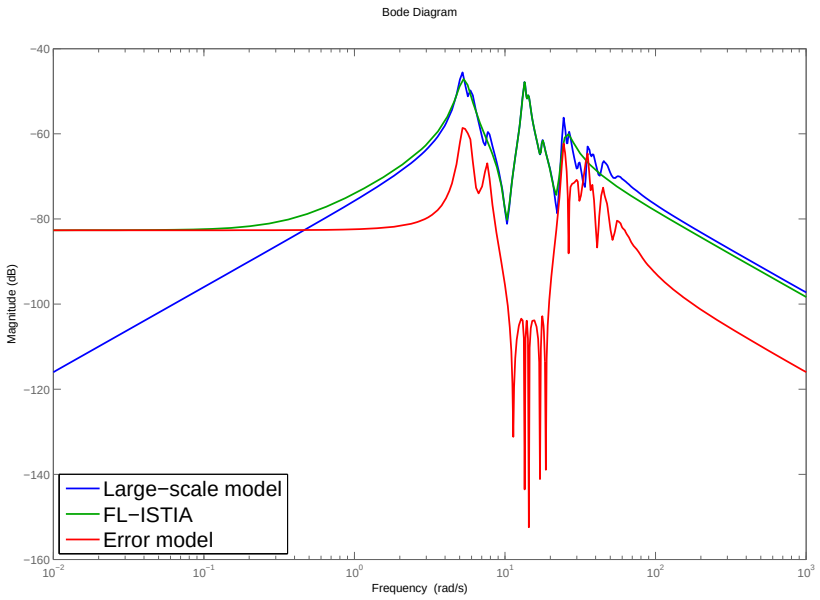


Figure C.1: Frequency responses of the large-scale model, the reduced-order model and the error model.

```
r                = 10;                % approximation order
fb               = [10 20];           % frequency interval
%% Optional parameters
% default values are assigned if left blank
opt.shiftPoint   = linspace(fb(1),fb(2),r);% initial shift points
opt.gram         = 'o';                % gramian used
opt.checkH2      = 'no';               % error monitoring
opt.restart      = 0;                  % number of restart
opt.maxIter      = 30;                 % maximum number of iterations
opt.display      = 1;                  % verbosity of the algorithm
opt.freqBand     = fb;                 % frequency band
%% Approximation
% routine to call ISTIA on the LAH model
[ H_flistia ,out1] = moreLTI('LAH',r,'ISTIA',opt);          % FL−ISTIA
%% Error
bodemag(H,H_flistia,H−H_flistia)
legend('Large−scale model','FL−ISTIA','Error model','Location','SouthWest');
optNorm.freqBand = fb;  % frequency band for the norm
% relative error computed with moreNorm
relError = 100*moreNorm(H−H_flistia,2,optNorm)/moreNorm(H,2,optNorm);
sprintf('Norm of the relative error : %2.2f (%%)',relError)
```

C.4 Example 15 : computation of the bounds on the $\mathcal{H}_\infty$-norm

This code shows how the bounds Γ and $\bar{\Gamma}$ can be called (Example 15 from Section 7.3).

```
%% Illustration of the uppber bounds on the Hinf norm
load lah
H = ss(A,B,C,0);
```

```
% Computation of the bounds over [0 inf)
[GammaBar,Gamma,winf] = moreHinfOmegaBound(H);
% and over [10 20]
opt.freqBand = [20 30];
[GammaBarOmega,GammaOmega,winfOmega] = moreHinfOmegaBound(H,opt);
%
bodemag(H,'k')
hold on
loglog(get(gca,'xlim'),20*log10(GammaBar)*ones(1,2),'b')
loglog(get(gca,'xlim'),20*log10(Gamma)*ones(1,2),'r')
loglog(get(gca,'xlim'),20*log10(GammaBarOmega)*ones(1,2),'c')
loglog(get(gca,'xlim'),20*log10(GammaOmega)*ones(1,2),'m')
loglog(opt.freqBand(1)*ones(1,2),get(gca,'ylim'),'k--')
loglog(opt.freqBand(2)*ones(1,2),get(gca,'ylim'),'k--')
h=legend('$H$','$\bar{\Gamma}$','$\Gamma$','$\bar{\Gamma}_{\Omega}$','$\Gamma_\Omega$','$\
    Omega$');
set(h,'Fontsize',18,'interpreter','latex','location','southEast');
```

C.5 Example 16 : construction of a frequency template

This code shows how a frequency template can be built to bound the Frobenius norm of the approximation error between the ISS model and a reduced-order model obtained with **DARPO** (Example 16 from Section 7.3.4).

```
%% Illustration of the frequency template
load iss1_2
H = ss(A,B,C,0);
% Reduction of the ISS model over [0 10] to an order 20 with DARPO
r =20;
opt.freqBand = [0 10];
[Hr,out] = moreLTI('ISS1',r,'DARPO',opt);
% Construction of the frequency template with different tuning parameters
tempOrder = 2;
optTemp.alpha = (1-1e-4); % decrease factor of the pole
optTemp.beta = (1+1e-3); % increase factor of the gain
optTemp.minStaticG = 1e-6;% minimal allowed static gain of the template
G1 = moreTemplate(H-Hr,tempOrder,optTemp);
%
optTemp.alpha = (1-1e-3); % decrease factor of the pole
optTemp.beta = (1+1e-4); % increase factor of the gain
optTemp.minStaticG = 1e-6;% minimal allowed static gain of the template
G2 = moreTemplate(H-Hr,tempOrder,optTemp);
%
sigma(H-Hr,G1,G2)
legend('H-Hr','G1','G2','Location','northWest')
```

C.6 Example 17 : behaviour of the $\mathcal{H}_{2,\Omega}$ approximation error when approximating to a second-order model

This code enables to reproduce Example 17. It is interesting to make the frequency bound ω vary to see how the shape of the approximation error is modified.

```
%%
load lah;
H = ss(A,B,C,0);
% computation of the residues and poles of the large-scale model that
% are required to build the affine term in the quadratic expression of the
% approximation error
[X, D]  = eig(H.A);
```

```
eign     = diag(D);
c        = C*X;
b        = X\B;
for i = 1:length(H.A)
  phi(i,1) = c(:,i)*b(i,:) ;
end
omega = 20; % upper bound of the frequency interval Omega
rp    = linspace(-3,3,50); % real part of the reduced order poles
rp    = [rp 0]; rp    = sort(rp); % adding the imaginary axis
ip    = linspace(1,30,50); % imaginary part of the reduced-order poles
% H2w norm of the large-scale model
opt.freqBand = [0 omega];
h2wNormOfH   = moreNorm(H,2,opt);
% the inverse tangent function
myatan = @(z) 1/(2*1i)*(log(1+1i*z)-log(1-1i*z));
% function that computes : a_{omega,l1} + a_{omega,l2}
sumAtan = @(omega,l1,l2) 2/pi*(feval(myatan,omega./l1)+feval(myatan,omega./l2));
% function used to build the affine term h
Hw      = @(omega,l) sum(phi./(eign + l).*feval(sumAtan,omega,eign,l));
%%
h2wError =zeros(length(rp),length(ip));
for i = 1:length(rp)
  for j = 1:length(ip)
    % poles of the reduced-order model
    P(i,j) = complex(rp(i),ip(j));
    p1 = P(i,j);p2 = conj(p1);
    if rp(i) == 0 && ip(j)<=omega
      % if the poles are on the imaginary axis and inside [0, omega]
      % the error is, by definition, infinite
      h2wError(i,j) = inf;
    elseif rp(i)==0 && ip(j)>omega
      % if the poles are on the imaginary axis and outside [0, omega]
      L = [ 2/pi*feval(myatan,omega/p1)/p1, -2*omega/pi*1/(omega^2-p1*p2);
            -2*omega/pi*1/(omega^2-p2*p1)   , 2/pi*feval(myatan,omega/p2)/p2];
      rhs = [feval(Hw,omega,p1);% rhs = -h
             feval(Hw,omega,p2)];

      redPhi_opt =L\rhs; % optimal residues for these poles
      squaredh2wNormOfHr = 1/2*real(redPhi_opt.'*L*redPhi_opt);
      h2wError(i,j) = sqrt(h2wNormOfH^2-squaredh2wNormOfHr);
    else
      % if the poles are not on the imaginary axis
      L = [ 2/pi*feval(myatan,omega/p1)/p1      , feval(sumAtan,omega,p1,p2)/(p1+p2);
            feval(sumAtan,omega,p2,p1)/(p2+p1), 2/pi*feval(myatan,omega/p2)/p2];

      rhs = [ feval(Hw,omega,p1); %rhs = -h
              feval(Hw,omega,p2)];
      redPhi_opt =L\rhs;% optimal residues for these poles
      squaredh2wNormOfHr = 1/2*real(redPhi_opt.'*L*redPhi_opt);
      h2wError(i,j) = sqrt(h2wNormOfH^2-squaredh2wNormOfHr);
    end
  end
end
%% 3D plot of the H2W approximation error
surf(real(P),imag(P),h2wError/h2wNormOfH*100)
view(-133,16)
xlabel('Real part')
ylabel('Im. part')
zlabel('Relative H2W error')
```

C.7 Use of DARPO

In this example, the LAH model is reduced to an order $r = 10$ over $\Omega = [10, 20]$ with the **DARPO** (see Chapter 9) and the frequency responses of the models are plotted together with the frequency response of the error.

```
%% Example FL-ISTIA
%% Model and approximation parameters
load lah;
H               = ss(A,B,C,0);      % initial model
r               = 10;               % approximation order
fb              = [10 20];          % frequency interval
%% Optional parameters
% default values are assigned if left blank
opt.maxIter     = 500;                  % maximum number of iterations
opt.display     = 1;                    % verbosity of the algorithm
opt.freqBand    = fb;                   % frequency band
opt.tol         = 1e-12;                % tolerance
%% Approximation
% routine to call DARPO on the LAH model
[H_darpo,out1] = moreLTI('LAH',r,'DARPO',opt);
%% Error
bodemag(H,H_darpo,H-H_darpo)
legend('Large-scale model','DARPO','Error model','Location','SouthWest');
% relative error (in %) given by the algorithm directly
relError = out1.h2wError(end);
sprintf('Norm of the relative error : %2.2f (%%)',relError)
```

C.8 Use of DARPO for the control of a large-scale model

This code enables to reproduce the example presented in Section 9.2.3. Note that for convenience, the $\mathcal{H}_\infty$ problems are solved with the method available in the Robust Control Toolbox of Matlab®, but one could also have used HIFOO which is available at `http://www.cs.nyu.edu/overton/software/hifoo/`.

```
clear all;close all;clc;
%% Illustration of the process for the control of large-scale model
%% initial model and parameters
load iss1_2
H_ol  = ss(A,B,C,0);
I3    = eye(3);
Z3    = zeros(3,3);
doLargeSynthesis = 0; % set to 1 to make the synthesis on the large-scale model
%% P for the hinf control problem
P_large     = ss(A,[B B],[C;C],0);
%
Xdb         = 30;
alpha       = (10^(Xdb/20))/norm(H_ol,inf);
% actuator
Act         = tf(20*20,[1 1.4*20 20*20])*I3;
% roll off filter
wrl         = 40;
W           = 1/(3*alpha)*tf([1/wrl^2 0 0],[1/(300*wrl)^2 1.4/(300*wrl) 1])*I3;
% hinf control problem with actuator
P_large_act = P_large*append(I3,Act);
% Static gain for the disturbance rejection on the large plant without
% actuators :
BF1         = alpha*lft(P_large,-alpha*eye(3));
% Static gain for the disturbance rejection on the large plant with
% actuators :
BFa1        = alpha*lft(P_large,-alpha*eye(3)*Act);
```

```
%% Control on the large plant
if doLargeSynthesis == 1
K6th1      = ltiblock.ss('K6_large',6,3,3);%
% LFR for the 'large' Hinf problem
CL_large   = alpha*lft(P_large_act,K6th1);
t          = cputime;
[CL_large_opt,gamma_large]=hinfstruct(blkdiag(CL_large,W*K6th1));
dt_large   = cputime-t;
K_large    = ss(CL_large_opt.blocks.K6_large);
% Checks that the controller synthesised on the large plant
% stabilises the large plant
isstable(lft(P_large_act,K_large))
% satisfies the disturbance rejection on the large plant
norm(alpha*lft(P_large_act,K_large),Inf)
end
%% Control on the low order plant with uncertainty
K6th_small          = ltiblock.ss('K6_small',6,3,3);%
r                   = 30; % approximation order
opt.freqBand        = [0 40]; % frequency interval of reduction
Hr                  = moreLTI('ISS1',r,'DARPO',opt); % reduction of the large-scale model
He                  = H_ol-Hr; % approximation error
optTemp.alpha       = 1-1e-5; % decrease of the pole of the template
optTemp.beta        = 1+1e-4; % increase of the gain of the template
optTemp.minStaticG  = 1e-8; % minimal static gain
G                   = moreTemplate(He,2,optTemp); % construction of the frequency template
G                   = G*I3;
%% Low order Hinf control problem
%
P_small             = ss(Hr.a,[Hr.b Hr.b],[Hr.c;Hr.c],0);
P_small_act         = P_small*append(I3,Act);
%
P_small2            = ss(Hr.a,[0*Hr.b Hr.b],[0*Hr.c;Hr.c],[Z3 I3;I3 Z3]);
P_small_error_act   = append(G,I3)*P_small2*append(I3,Act);
% LFRs for the low-order Hinf problem
CL_small_act        = alpha*lft(P_small_act,K6th_small); % disturbance rejection on the
                                                         % low-order plant
CL_small_error_act  = lft(P_small_error_act,K6th_small); % robust stability channel
%
t                   = cputime;
[CL_small_opt,gamma_small,info] = hinfstruct(blkdiag(CL_small_act,CL_small_error_act,W*K6th_small));
dt_small            = cputime-t;
K_small             = ss(CL_small_opt.blocks.K6_small);
% Checks that the controller synthesised on the low-order plant
% stabilises the large plant
isstable(lft(P_large_act,K_small))
% satisfies the disturbance rejection on the large-plant
norm(alpha*lft(P_large_act,K_small),Inf)
%% PLOT
WF      = sort([logspace(-1,2,400) linspace(30,100,100)]);
[svBO1] = sigma(alpha*H_ol,WF);

% Open loop vs closed loop without actuator
figure
green = 1/256*[97 186 26];
[svBF1] = sigma(BF1,WF);
p1 = semilogx(WF,20*log10(svBO1),'b');
hold on
p2 = semilogx(WF,20*log10(svBF1),'color',green);
p3 = semilogx(get(gca,'xlim'),zeros(1,2),'k--');
hh=legend([p1(1) p2(1) p3],'Normalised open-loop model $\alpha H$','Closed-loop model $\mathcal{F}_l({
    P},{K_1})$','Disturbance rejection limit');
set(hh,'interpreter','latex');
xlabel('Pulsation (rad/s)');ylabel('Singular values (dB)');
% Open loop vs closed loop with actuator
evBO1 = eig(H_ol.A);
K2 = -3*alpha*I3;
```

```
BF2 = alpha*lft(P_large_act,K2);
evBF2= eig(BF2.A);

figure
subplot(121)
[svBFa1] = sigma(BFa1,WF);
p1 =semilogx(WF,20*log10(svBO1),'b');
hold on
p2 = semilogx(WF,20*log10(svBFa1),'r');
set(gca,'xlim',[min(WF),max(WF)])
p3 = semilogx(get(gca,'xlim'),zeros(1,2),'k--');
hh=legend([p1(1),p2(1),p3(1)],'Normalised open-loop model $\alpha H$','Closed-loop model $\mathcal{F}_l
    ^{(2)}({P},{K_1})$','Disturbance rejection limit');
set(hh,'interpreter','latex');
xlabel('Pulsation (rad/s)');ylabel('Singular values (dB)');
subplot(122)
plot(real(evBO1),imag(evBO1),'bx');
hold on
plot(real(evBF2),imag(evBF2),'r.');
axis([-5 1 -1 60])
plot(get(gca,'xlim'),[0 0],'k')
plot([0 0],get(gca,'ylim'),'k')
hh=legend('Open-loop poles','Closed-loop poles');
set(hh,'interpreter','latex');
xlabel('Real part');ylabel('Imaginary part');

% Frequency template on the approximation error
WFe = logspace(-4,3,500);
figure
svError = sigma(He,WFe);
svG     = sigma(G(1,1),WFe);
p1 = semilogx(WFe,20*log10(svError),'b');
hold on
p2 = semilogx(WFe,20*log10(svG),'color',green);
h = legend([p1(1),p2(1)],'Error model $H-\hat{H}$','Frequency template $G$');
set(h,'Location','NorthWest','interpreter','latex');
xlabel('Pulsation (rad/s)');ylabel('Singular values (dB)');
set(gca,'xlim',[min(WFe),max(WFe)])

% Final closed loop
figure
svBFr = sigma(alpha*lft(P_small_act,K_small),WF);
svBFf = sigma(alpha*lft(P_large_act,K_small),WF);
p1 = semilogx(WF,20*log10(svBO1),'b');
hold on
p2 = semilogx(WF,20*log10(svBFr),'color',green);
p3 = semilogx(WF,20*log10(svBFf),'r');
p4 = semilogx(get(gca,'xlim'),zeros(1,2),'k--');
hh = legend([p1(1),p2(1),p3(1),p4(1)],'Normalised open-loop model $\alpha H$','Low-order closed-loop $\
    mathcal{F}_l^{(2)} (\hat{P},K)$','Large-scale closed-loop $\mathcal{F}_l^{(2)} (P,K)$','Disturbance
    rejection limit');
set(hh,'interpreter','latex');
xlabel('Pulsation (rad/s)');ylabel('Singular values (dB)');

% Controller dynamic
figure
WF2 = logspace(1,5,200);

[svW] =sigma(1/W,WF2);
[svK] = sigma(K_small,WF2);
p1 = semilogx(WF2,20*log10(svW),'b');
hold on
p2 = semilogx(WF2,20*log10(svK),'color',green);
hh=legend([p1(1),p2(1)],'$W^{-1}$','$K$');
set(hh,'interpreter','latex');
xlabel('Pulsation (rad/s)');ylabel('Singular values (dB)');
```

C.9 Use of DARPO for the approximation at a fixed error

This code illustrates how **DARPO** can be used to approximate a model at a given relative $\mathcal{H}_{2,\Omega}$ error instead of a given approximation order r (see Section 11.5.3).

```
%% Illustration of the use of DARPO at fixed approximation error
% Initial model
load lah
H               = ss(A,B,C,0);
% Model approximation
opt.freqBand  = [0 20]; % frequency interval
opt.h2wError  = 5; % Desired relative H2w approximation error (in %)
% The order r must be left blank when calling the routine
[Hr,out]        = moreLTI('LAH',[],'DARPO',opt);
%
figure
semilogy([1:length(out.h2wError)],out.h2wError)
hold on
semilogy(get(gca,'xlim'),5*[1 1],'r--')
yl = get(gca,'ylim');
for i = 1:length(out.rIncreased)
  semilogy(out.rIncreased(i)*[1 1],yl,'k--');
end
legend('Approximation error','Desired relative H2W error','Reduced order increase');
xlabel('Iteration');ylabel('Relative error');
figure
bodemag(H,'k',Hr,'r--',H-Hr,'b-.')
legend('H','Hr','H-Hr');
```

Bibliography

Abramowitz, M. and Stegun, I. (1964). *Hanbook of mathematical functions with formulas, graphs and mathematical tables*. US Government Printing Office.

Al-Mohy, A., Higham, N., and Relton, S. (2013). Computing the Fréchet derivative of the matrix logarithm and estimating the condition number. *SIAM Journal on Scientific Computing*, 35(4):C394 – C410.

Anderson, M. R., Emami-Naeni, A., and Vincent, J. (1991). Measures of merit for multivariable flight control. Technical report, Systems Control Technology Inc, Palo Alto, California, USA.

Anić, B., Beattie, C., Gugercin, S., and Antoulas, A. (2013). Interpolatory weighted-$\mathcal{H}_2$ model reduction. *Automatica*, 49(5):1275 – 1280.

Antoulas, A. (2005). *Approximation of large-scale dynamical systems*. Advances in Design and Control. SIAM.

Antoulas, A., Ionita, A., and Lefteriu, S. (2012). On two-variable rational interpolation. *Linear Algebra and its Applications*, 436(8):2889 – 2915. Special Issue dedicated to Danny Sorensen's 65th birthday.

Apkarian, P. and Noll, D. (2006). Nonsmooth $\mathcal{H}_\infty$ synthesis. *IEEE Transactions on Automatic Control*, 51(1):71 – 86.

Badía, J., Benner, P., Mayo, R., and Quintana-Ortí, E. (2006). Parallel algorithms for balanced truncation model reduction of sparse systems. In Dongarra, J., Madsen, K., and Waśniewski, J., editors, *Applied Parallel Computing. State of the Art in Scientific Computing*, volume 3732 of *Lecture Notes in Computer Science*, pages 267 – 275. Springer.

Balakrishnan, A. (1996). Robust stabilizing compensators for flexible structures with collocated controls. *Applied Mathematics and Optimization*, 33(1):35 – 60.

Baratchart, L. (1986). Existence and generic properties of $\mathcal{L}_2$ approximants for linear systems. *IMA Journal of Mathematical Control and Information*, 3(2-3):89 – 101.

Baratchart, L., Cardelli, M., and Olivi, M. (1991). Identification and rational $\mathcal{L}_2$ approximation : A gradient algorithm. *Automatica*, 27(2):413 – 417.

Beattie, C. and Gugercin, S. (2007). Krylov-based minimization for optimal $\mathcal{H}_2$ model reduction. In *Proceedings of the Conference on Decision and Control*, pages 4385 – 4390.

Beattie, C. and Gugercin, S. (2009). A trust region method for optimal $\mathcal{H}_2$ model reduction. In *proceedings of the Joint Conference on Decision and Control and Chinese Control Conference*, pages 5370 – 5375.

Beattie, C. and Gugercin, S. (2012). Realization-independent $\mathcal{H}_2$ approximation. In *Proceedings of the Conference on Decision and Control*, pages 4953 – 4958.

Beattie, C., Gugercin, S., and Wyatt, S. (2012). Inexact solves in interpolatory model reduction. *Linear Algebra and its Applications*, 436(8):2916 – 2943. Special Issue dedicated to Danny Sorensen's 65th birthday.

Bemporad, A. and Morari, M. (1999). Robust model predictive control: A survey. In Garulli, A. and Tesi, A., editors, *Robustness in identification and control*, volume 245 of *Lecture Notes in Control and Information Sciences*, pages 207 – 226. Springer.

Benner, P. and Breiten, T. (2012). Krylov-subspace based model reduction of nonlinear circuit models using bilinear and quadratic-linear approximations. In Günther, M., Bartel, A., Brunk, M., Schöps, S., and Striebel, M., editors, *Progress in Industrial Mathematics at ECMI 2010*, Mathematics in Industry, pages 153 – 159. Springer.

Benner, P., Claver, J., and Quintana-Ortí, E. (1999). Parallel distributed solvers for large stable generalized Lyapunov equations. *Parallel Processing Letters*, 9(01):147 – 158.

Benner, P., Mehrmann, V., and Sorensen, D. (2005). *Dimension reduction of large-scale systems*, volume 45. Springer.

Benner, P., Quintana-Ortí, E., and Quintana-Ortí, G. (2000). Balanced truncation model reduction of large-scale dense systems on parallel computers. *Mathematical and Computer Modelling of Dynamical Systems*, 6(4):383 – 405.

Biannic, J. and Ferreres, G. (2005). Efficient computation of a guaranteed robustness margin. In *Proceedings of the IFAC World Congress*.

Blondel, V. and Megretski, A., editors (2004). *Unsolved problems in mathematical systems and control theory*. Princeton University Press.

Brandwood, D. (1983). A complex gradient operator and its application in adaptive array theory. In *Proceedings of the Conference on Communications, Radar and Signal Processing*, volume 130, pages 11 – 16.

Breiten, T., Beattie, C., and Gugercin, S. (2014). Near-optimal frequency-weighted interpolatory model reduction. Available as arXiv:1309.0136.

Bugarin, F., Henrion, D., and Lasserre, J. (2011). Minimizing the sum of many rational functions. Available as arXiv:1102.49544.

Burke, J., Henrion, D., Lewis, A., and Overton, M. (2006). HIFOO-a Matlab package for fixed-order controller design and $\mathcal{H}_\infty$ optimization. In *Proceedings of the IFAC Symposium on Robust Control Design*.

Burlion, L., Poussot-Vassal, C., Vuillemin, P., Leitner, M., and Kier, T. (2014). Longitudinal manoeuvre load control of a flexible large-scale aircraft. In *Proceedings of the IFAC World Congress*, pages 3413 – 3418.

Chahlaoui, Y., Lemonnier, D., Vandendorpe, A., and Van Dooren, P. (2006). Second-order balanced truncation. *Linear Algebra and its Applications*, 415(2–3):373 – 384. Special issue on order reduction of large-scale systems.

Corless, R. M., Jeffrey, D. J., Watt, S. M., and Davenport, J. H. (2000). "According to Abramowitz and Stegun" or arccoth needn't be uncouth. *ACM SIGSAM Bulletin*, 34:58 – 65.

de Villemagne, C. and Skelton, R. (1987). Model reductions using a projection formulation. In *Proceeding of the Conference on Decision and Control*, volume 26, pages 461 – 466.

Demourant, F. (2013). New algorithmic approach based on integral quadratic constraints for stability analysis of high order models. In *Proceedings of the European Control Conference*, pages 359 – 364.

Dergham, G., Sipp, D., Robinet, J.-C., and Barbagallo, A. (2011). Model reduction for fluids using frequential snapshots. *Physics of Fluids*, 23(6).

Enns, D. (1984). Model reduction with balanced realizations: An error bound and a frequency weighted generalization. In *Proceedings of the Conference on Decision and Control*, pages 127 – 132.

Ferreres, G. (2011). Computation of a flexible aircraft LPV/LFT model using interpolation. *IEEE Transactions on Control Systems Technology*, 19(1):132 – 139.

Flagg, G., Beattie, C., and Gugercin, C. (2012). Convergence of the iterative rational Krylov algorithm. *Systems & Control Letters*, 61(6):688 – 691.

Flagg, G., Beattie, C., and Gugercin, S. (2013). Interpolatory $\mathcal{H}_\infty$ model reduction. *Systems & Control Letters*, 62(7):567 – 574.

Fletcher, R. (2000). *Practical methods of optimization*. John Wiley & Sons.

Fuhrmann, P. (1994). A duality theory for robust stabilization and model reduction. *Linear Algebra and its Applications*, 203-204(0):471 – 578.

Fulcheri, P. and Olivi, M. (1998). Matrix rational $\mathcal{H}_2$ approximation: A gradient algorithm based on Schur analysis. *SIAM Journal on Control and Optimization*, 36(6):2103 – 2127.

Gadient, R., Lavrtsky, E., and Wise, K. (2012). Very flexible aircraft control challenge problem. In *Proceedings of the AIAA guidance, Navigation and Control Conference*.

Gallivan, K., Vandendorpe, A., and Van Dooren, P. (2004a). Model reduction of MIMO systems via tangential interpolation. *SIAM Journal on Matrix Analysis and Applications*, 26(2):328 – 349.

Gallivan, K., Vandendorpe, A., and Van Dooren, P. (2004b). Sylvester equations and projection-based model reduction. *Journal of Computational and Applied Mathematics*, 162(1):213 – 229.

Garulli, A., Hansson, A., Pakazad, S., Masi, A., and Wallin, R. (2013). Robust finite-frequency analysis of uncertain systems with application to flight comfort analysis. *Control Engineering Practice*, 21(6):887 – 897.

Gawronski, W. (2004). *Advanced structural dynamics and active control of structures*. Springer.

Gawronski, W. and Juang, J. (1990). Model reduction in limited time and frequency intervals. *International Journal of Systems Science*, 21(2):349 – 376.

Glover, K. (1984). All optimal Hankel-norm approximations of linear multivariable systems and their $\mathcal{L}_\infty$-error bounds. *International Journal of Control*, 39(6):1115 – 1193.

Grigoriadis, K. (1995). Optimal $\mathcal{H}_\infty$ model reduction via linear matrix inequalities: continuous- and discrete-time cases. *Systems & Control Letters*, 26(5):321 – 333.

Grimme, E. (1997). *Krylov projection methods for model reduction*. PhD thesis, University of Illinois at Ubarna-Champaign.

Gugercin, S. (2007). An iterative SVD-Krylov based method for model reduction of large-scale dynamical systems. *Linear Algebra and its Applications*, 428(8-9):1964 – 1986.

Gugercin, S. and Antoulas, A. (2004). A survey of model reduction by balanced truncation and some new results. *International Journal of Control*, 77(8):748 – 766.

Gugercin, S., Antoulas, A., and Beattie, C. (2006). A rational Krylov iteration for optimal $\mathcal{H}_2$ model reduction. In *Proceedings of the International Symposium on Mathematical Theory of Networks and Systems*, pages 1665 – 1667.

Gugercin, S., Antoulas, A., and Beattie, C. (2008). $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. *SIAM Journal on Matrix Analysis and Applications*, 30(2):609 – 638.

Gugercin, S. and Li, J. (2005). Smith-type methods for balanced truncation of large sparse systems. In Benner, P., Sorensen, D., and Mehrmann, V., editors, *Dimension reduction of large-scale systems*, volume 45 of *Lecture Notes in Computational Science and Engineering*, pages 49 – 82. Springer.

Gugercin, S., Stykel, T., and Wyatt, S. (2013). Model reduction of descriptor systems by interpolatory projection methods. Available as arXiv:1301.4524.

Gumussoy, S., Henrion, D., Millstone, M., and Overton, M. (2009). Multiobjective robust control with HIFOO 2.0. In *Proceedings of the 6th IFAC symposium on robust control design*.

Haber, H. (2012). The complex inverse trigonometric and hyperbolic functions. `http://scipp.ucsc.edu/~haber/webpage/arc.pdf`. University of California.

Hammarling, S. (1982). Numerical solution of the stable, non-negative definite Lyapunov equation. *IMA Journal of Numerical Analysis*, 2(3):303 – 323.

Heath, M., Laub, A., Paige, C., and Ward, R. (1986). Computing the singular value decomposition of a product of two matrices. *SIAM Journal on Scientific and Statistical Computing*, 7(4):1147 – 1159.

Hesse, H. and Palacios, R. (2014). Reduced-order aeroelastic models for dynamics of maneuvering flexible aircraft. *AIAA Journal*, 52(8):1717 – 1732.

Higham, N. (2008). *Functions of matrices : Theory and computation*. SIAM.

Himpe, C. (2014). emgr - Empirical Gramian Framework. `http://gramian.de`.

Hyland, D. and Bernstein, D. (1985). The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton, and Moore. *IEEE Transactions on Automatic Control*, 30(12):1201 – 1211.

Ionita, A. (2013). *Lagrange rational interpolation and its application to approximation of large-scale dynamical systems*. PhD thesis, Rice University.

Kahan, W. (1987). Branch cuts for complex elementary functions, or much ado about nothing's sign bit. In *Proceedings of the Joint IMA/SIAM Conference on the State of the Art in Numerical Analysis*, pages 165 – 211.

Kim, S., Anderson, B., and Madievski, A. (1995). Error bound for transfer function order reduction using freqeuncy weighted balanced truncation. *Systems & Control Letters*, 24(3):183 – 192.

Kokotovic, P., Khalil, H., and O'Reilly, J. (1986). *Singular perturbation methods in control: analysis and design*. SIAM.

Korvink, J. and Rudnyi, E. (2005). Oberwolfach benchmark collection. In Benner, P., Sorensen, D., and Mehrmann, V., editors, *Dimension reduction of large-scale systems*, volume 45 of *Lecture Notes in Computational Science and Engineering*, pages 311 – 315. Springer.

Krajewski, W., Lepschy, A., Redivo-Zaglia, M., and Viaro, U. (1995). A program for solving the $\mathcal{L}_2$ reduced-order model problem with fixed denominator degree. *Numerical Algorithms*, 9(2):355 – 377.

Kreutz-Delgado, K. (2009). The complex gradient operator and the CR-calculus. Available as arXiv:0906.4835.

Laub, A., Heath, M., Paige, C., and Ward, R. (1987). Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms. *IEEE Transactions on Automatic Control*, 32(2):115 – 122.

Leblond, J. and Olivi, M. (1998). Weighted $\mathcal{H}_2$ approximation of transfer functions. *Mathematics of Control, Signals and Systems*, 11(1):28 – 39.

Leibfritz, F. and Lipinski, W. (2003). Description of the benchmark examples in $COMPl_eib$ 1.0. Technical report, University of Trier.

Li, D. and Fukushima, M. (2001). On the global convergence of the BFGS method for nonconvex unconstrained optimization problems. *SIAM Journal on Optimization*, 11(4):1054 – 1064.

Li, J. and Kamon, M. (2005). PEEC model of a spiral inductor generated by Fasthenry. In Benner, P., Sorensen, D., and Mehrmann, V., editors, *Dimension reduction of large-scale systems*, volume 45 of *Lecture Notes in Computational Science and Engineering*, pages 373 – 377. Springer Berlin Heidelberg.

Li, J. and White, J. (2002). Low rank solution of Lyapunov equations. *SIAM Journal on Matrix Analysis and Applications*, 24(1):260 – 280.

Lienemann, J., Rudnyi, E., and Korvink, J. (2006). MST MEMS model order reduction: Requirements and benchmarks. *Linear Algebra and its Applications*, 415(2–3):469 – 498. Special issue on order reduction of large-scale systems.

Lin, C.-A. and Chiu, T.-Y. (1990). Model reduction via frequency weighted balanced realization. In *American Control Conference*, pages 2069 – 2070.

Liu, Y. and Anderson, B. D. (1989). Singular perturbation approximation of balanced systems. In *Proceedings of the Conference On Decision and Control*, pages 1355 – 1360.

Magni, J. (2006). Linear fractional representation toolbox for use with Matlab. Available at http://w3.onera.fr/smac/.

Mallik, R. (2001). The inverse of a tridiagonal matrix. *Linear Algebra and its Applications*, 325(1-3):109 – 139.

Marmorat, J. and Olivi, M. (2002). The rarl2 software: Realizations and rational approximation in $\mathcal{L}_2$-norm. `http://www-sop.inria.fr/apics/RARL2/rarl2.html`. Project APICS, INRIA, Sophia Antipolis.

Marmorat, J.-P., Olivi, M., Hanzon, B., and Peeters, R. (2002). Matrix rational $\mathcal{H}_2$ approximation: A state-space approach using Schur parameters. In *Proceedings of the Conference on Decision and Control*, volume 4, pages 4244 – 4249.

Mayo, A. and Antoulas, A. (2007). A framework for the solution of the generalized realization problem. *Linear Algebra and its Applications*, 425(2–3):634 – 662. Special Issue in honor of Paul Fuhrmann.

Megretski, A. and Rantzer, A. (1997). System analysis via integral quadratic constraints. *IEEE Transactions on Automatic Control*, 42(6):819 – 830.

Mehrmann, V. and Stykel, T. (2005). Balanced truncation model reduction for large-scale systems in descriptor form. In Benner, P., Sorensen, D., and Mehrmann, V., editors, *Dimension Reduction of Large-Scale Systems*, volume 45 of *Lecture Notes in Computational Science and Engineering*, pages 83–115.

Meier, L. and Luenberger, D. (1967). Approximation of linear constant systems. *IEEE Transactions on Automatic Control*, 12(5):585 – 588.

Moler, C. and Van Loan, C. (1978). Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review*, 20:801 – 836.

Moore, B. (1981). Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, 26(1):17 – 32.

Nocedal, J. and Wright, S. (1999). *Numerical optimization*. Springer.

Nocedal, J. and Yuan, Y. (1998). Combining trust region and line search techniques.

Ostrowski, A. and Schneider, H. (1962). Some theorems on the inertia of general matrices. *Journal of Mathematical Analysis and Applications*, pages 72 – 84.

Panzer, H., Hubele, J., Eid, R., and Lohmann, B. (2009). Generating a parametric finite element model of a 3d cantilever timoshenko beam using Matlab. Technical report, Technische Universität München.

Penzl, T. (1999). A cyclic low-rank Smith method for large sparse Lyapunov equations. *SIAM Journal on Scientific Computing*, 21(4):1401 – 1418.

Pernebo, L. and Silverman, L. (1982). Model reduction via balanced state space representations. *IEEE Transactions on Automatic Control*, 27(2):382 – 387.

Petersson, D. (2013). *A nonlinear optimization approach to $\mathcal{H}_2$-optimal modeling and control*. PhD thesis, Linköping University.

Petersson, D. and Löfberg, J. (2014). Model reduction using a frequency-limited $\mathcal{H}_2$-cost. *Systems & Control Letters*, 67(0):32 – 39.

Poussot-Vassal, C. (2011). An iterative SVD-tangential interpolation method for medium-scale MIMO systems approximation with application on flexible aircraft. In *Proceedings of the Conference on Decision and Control and European Control Conference*, pages 7117 – 7122.

Poussot-Vassal, C., Loquen, T., Vuillemin, P., Cantinaud, O., and Lacoste, J. (2013). Business jet large-scale model approximation and vibration control. In *Proceedings of the IFAC International Workshop on Adaptation Learning in Control and Signal Processing*, pages 199 – 204.

Poussot-Vassal, C. and Roos, C. (2012). Generation of a reduced-order LPV/LFT model from a set of large-scale MIMO LTI flexible aircraft models. *Control Engineering Practice*, 20(9):919 – 930.

Poussot-Vassal, C., Ross, C., Vuillemin, P., Cantinaud, O., and Lacoste, J. (2014). Control-oriented aeroelastic bizjet low-order LFT modeling. In *Control-oriented modeling and identification : theory and practice*. The Institution of Engineering and Technology. To appear.

Poussot-Vassal, C. and Vuillemin, P. (2012). Introduction to MORE: a MOdel REduction Toolbox. In *Proceedings of the Multi-conference on Systems and Control*, pages 776 – 781.

Poussot-Vassal, C. and Vuillemin, P. (2013). An iterative eigenvector tangential interpolation algorithm for large-scale LTI and a class of LPV model approximation. In *Proceedings of the European Control Conference*, pages 4490 – 4495.

Remmert, R. (1991). *Theory of complex functions*, volume 122. Springer.

Rommes, J. and Martins, N. (2006). Efficient computation of multivariable transfer function dominant poles using subspace acceleration. *IEEE Transactions on Power Systems*, 21:1471 – 1483.

Roos, C., Lescher, F., Biannic, J., Doll, C., and Ferreres, G. (2011). A set of μ-analysis based tools to evaluate the robustness properties of high-dimensional uncertain systems. In *Proceedings of the Multiconference on Systems and Control*, pages 644 – 649.

Ruhe, A. (1994). Rational Krylov algorithms for nonsymmetric eigenvalue problems. II. matrix pairs. *Linear Algebra and its Applications*, 197-198(0):283 – 295.

Saad, Y. (2003). *Iterative methods for sparse linear systems*. SIAM, 2^{nd} edition.

Scherer, C., Gahinet, P., and Chilali, M. (1997). Multiobjective output-feedback control via LMI optimization. *IEEE Transactions on Automatic Control*, 42(7):896 – 911.

Shaker, H. and Tahavori, M. (2014). Frequency-interval model reduction of bilinear systems. *IEEE Transactions on Automatic Control*, 59(7):1948 – 1953.

Sideris, A. and Sanchez Pena, R. (1990). Robustness margin calculation with dynamic and real parametric uncertainty. *IEEE Transactions on Automatic Control*, 35(8):970 – 974.

Sootla, A. (2011). Nu-gap model reduction in the frequency domain. In *American Control Conference*, pages 5025 – 5030.

Sorber, L., Van Barel, M., and De Lathauwer, L. (2012). Unconstrained optimization of real functions in complex variables. *SIAM Journal on Optimization*, 22(3):879 – 898.

Sorber, L., Van Barel, M., and De Lathauwer, L. (2013). Complex optimization toolbox v1.0. `http://esat.kuleuven.be/stadius/cot/`.

Sorensen, D. (1997). Implicitly restarted Arnoldi/Lanczos methods for large scale eigenvalue calculations. In Keyes, D., Sameh, A., and Venkatakrishnan, V., editors, *Parallel Numerical Algorithms*, volume 4 of *ICASE/LaRC Interdisciplinary Series in Science and Engineering*, pages 119 – 165. Springer.

Sorensen, D. and Antoulas, A. (2002). The Sylvester equation and approximate balanced reduction. *Linear Algebra and its Applications*, 351–352(0):671 – 700. Fourth Special Issue on Linear Systems and Control.

Stykel, T. and Simoncini, V. (2012). Krylov subspace methods for projected Lyapunov equations. *Applied Numerical Mathematics*, 62(1):35 – 50.

Unbehauen, H. and Rao, G. (1990). Continuous-time approaches to system identification - a survey. *Automatica*, 26(1):23 – 35.

Van Den Bos, A. (1994). Complex gradient and hessian. volume 141, pages 380 – 383.

Van Dooren, P., Gallivan, K., and Absil, P. (2008a). $\mathcal{H}_2$ optimal approximation of MIMO linear dynamical systems. Available as arXiv:0807.4807.

Van Dooren, P., Gallivan, K., and Absil, P. (2008b). $\mathcal{H}_2$-optimal model reduction of MIMO systems. *Applied Mathematics Letters*, 21(12):1267 – 1273.

Van Dooren, P., Gallivan, K., and Absil, P. (2010). $\mathcal{H}_2$-optimal model reduction with higher-order poles. *SIAM Journal on Matrix Analysis and Applications*, 31(5):2738 – 2753.

Vayssettes, J., Mercère, G., Vacher, P., and De Callafon, R. (2014). Frequency-domain identification of aircraft structural modes from short-duration flight tests. *International Journal of Control*, 87(7):1352 – 1372.

Vidyasagar, M. (1984). The graph metric for unstable plants and robustness estimates for feedback stability. *IEEE Transactions on Automatic Control*, 29(5):403 – 418.

Vinnicombe, G. (1992). *Measuring robustness of feedback systems*. PhD thesis, University of Cambridge.

Vizer, D., Mercere, G., Prot, O., Laroche, E., and Lovera, M. (2013). Linear fractional LPV model identification from local experiments: An $\mathcal{H}_\infty$-based optimization technique. In *Proceedings of the Conference on Decision and Control*, pages 4559 – 4564.

Vuillemin, P., Demourant, F., Biannic, J.-M., and Poussot-Vassal, C. (2014a). Global stability validation of an uncertain large-scale aircraft model. In *Proceedings of the Multi-conference on Systems and Control*.

Vuillemin, P., Poussot-Vassal, C., and Alazard, D. (2012a). Réduction de modèle de grande dimension, application à un avion souple de transport civil. In *Proceedings of the 7ème Conférence Internationale Francophone d'Automatique*.

Vuillemin, P., Poussot-Vassal, C., and Alazard, D. (2012b). A spectral expression for the frequency-limited $\mathcal{H}_2$-norm. Available as arXiv:1211.1858.

Vuillemin, P., Poussot-Vassal, C., and Alazard, D. (2013a). A frequency-limited $\mathcal{H}_2$ model approximation method with application to a medium-scale flexible aircraft. In Chu, Q., Mulder, B., Choukroun, D., van Kampen, E., de Visser, C., and Looye, G., editors, *Advances in Aerospace Guidance, Navigation and Control*, pages 569 – 583. Springer.

Vuillemin, P., Poussot-Vassal, C., and Alazard, D. (2013b). $\mathcal{H}_2$ optimal and frequency limited approximation methods for large-scale LTI dynamical systems. In *Proceedings of the IFAC Symposium on Systems Structure and Control*, pages 719 – 724.

Vuillemin, P., Poussot-Vassal, C., and Alazard, D. (2014b). Poles residues descent algorithm for optimal frequency-limited $\mathcal{H}_2$ model approximation. In *Proceedings of the European Control Conference*, pages 1080 – 1085.

Vuillemin, P., Poussot-Vassal, C., and Alazard, D. (2014c). Spectral expression for the frequency-limited $\mathcal{H}_2$-norm of LTI dynamical systems with high order poles. In *Proceedings of the European Control Conference*, pages 55 – 60.

Vuillemin, P., Poussot-Vassal, C., and Alazard, D. (2014d). Two upper bounds on the $\mathcal{H}_\infty$-norm of LTI dynamical systems. In *Proceedings of the IFAC World Congress*, pages 5562 – 5567.

Wang, G., Sreeram, V., and Liu, W. Q. (1999). A new frequency-weighted balanced truncation method and an error bound. *IEEE Transactions on Automatic Control*, 44(9):1734 – 1737.

Wang, Y., Andrew, W., and Palacios, R. (2014). Nonlinear model reduction for aeroelastic control of flexible aircraft described by large finite-element models. In *Proceedings of the AIAA/ASMe/ASCE/AHS/SC Structures, Structural Dynamics, and Materials Conference*.

Wilson, D. (1974). Model reduction for multivariable systems. *International Journal of Control*, 20(1):57 – 64.

Wood, G., Goddard, P., and Glover, K. (1996). Approximation of linear parameter-varying systems. In *Proceedings of the Conference on Decision and Control*, pages 406 – 411.

Yan, W.-Y. and Lam, J. (1999). An approximate approach to $\mathcal{H}_2$ optimal model reduction. *IEEE Transactions on Automatic Control*, 44(7):1341 – 1358.

Yousuff, A., Wagie, D., and Skelton, R. (1985). Linear system approximation via covariance equivalent realizations. *Journal of Mathematical Analysis and Applications*, 106(1):91 – 115.

Zhou, K. (1995). Frequency-weighted $\mathcal{L}_\infty$ norm and optimal Hankel norm model reduction. *IEEE Transactions on Automatic Control*, 40(10):1687 – 1699.

Zhou, K., Doyle, J. C., and Glover, K. (1995). *Robust and optimal control*. Pentice Hall.

Frequency-limited model approximation of large-scale dynamical models

Abstract. Physical systems are represented by mathematical models in order to be simulated, analysed or controlled. Depending on the complexity of the physical system it is meant to represent and the way it has been built, a model can have a varying complexity. This complexity can become an issue in practice due to the limited computational power and memory of computers. One way to alleviate this issue consists in using *model approximation* which is aimed at finding a *simpler* model that still represents faithfully the physical system. In the case of Linear Time Invariant (LTI) dynamical models, complexity translates into a large dimension of the state vector and one talks about *large-scale models*. Model approximation is in this case also called *model reduction* and consists in finding a model with a smaller state vector such that the input-to-output behaviours of both models are close with respect to some measure. The $\mathcal{H}_2$-norm has been extensively used in the literature to evaluate the quality of a reduced-order model. Yet, due to the limited bandwidth of actuators, sensors and the fact that models are generally representative on a bounded frequency interval only, a reduced-order model that faithfully reproduces the behaviour of the large-scale one over a bounded frequency interval only, may be more relevant. That is why, in this thesis, *the frequency-limited* $\mathcal{H}_2$-norm, or $\mathcal{H}_{2,\Omega}$-norm, which is the restriction of the $\mathcal{H}_2$-norm over a frequency interval Ω, has been considered. In particular, the problem of finding a reduced-order model that minimises the $\mathcal{H}_{2,\Omega}$-norm of the approximation error with the large-scale model has been addressed here. For that purpose, two approaches have been developed. The first one is an empirical approach based on the modification of a sub-optimal $\mathcal{H}_2$ model approximation method. Its performances are interesting in practice and compete with some well-known frequency-limited approximation methods. The second one is an optimisation method relying on the poles-residues formulation of the $\mathcal{H}_{2,\Omega}$-norm. This formulation naturally extends the one existing for the $\mathcal{H}_2$-norm and can also be used to derive two upper bounds on the $\mathcal{H}_\infty$-norm of LTI dynamical models which is of particular interest in model reduction. The first-order optimality conditions of the optimal $\mathcal{H}_{2,\Omega}$ approximation problem are derived and used to built a complex-domain descent algorithm aimed at finding a local minimum of the problem. Together with the $\mathcal{H}_\infty$ bounds on the approximation error, this approach is used to perform control of large-scale models. From a practical point of view, the methods proposed in this study have been successfully applied in an industrial context as a part of the global process aimed at controlling a flexible civilian aircraft.

Keywords : linear time invariant models, large-scale models, model reduction, model approximation, frequency-limited model approximation

Approximation de modèles dynamiques de grande dimension sur intervalles de fréquences limités

Résumé. Les systèmes physiques sont représentés par des modèles mathématiques qui peuvent être utilisés pour simuler, analyser ou contrôler ces systèmes. Selon la complexité du système qu'il est censé représenter, un modèle peut être plus ou moins complexe. Une complexité trop grande peut s'avérer problématique en pratique du fait des limitations de puissance de calcul et de mémoire des ordinateurs. L'une des façons de contourner ce problème consiste à utiliser *l'approximation de modèles* qui vise à remplacer le modèle complexe par un modèle *simplifié* dont le comportement est toujours représentatif de celui du système physique. Dans le cas des modèles dynamiques Linéaires et Invariants dans le Temps (LTI), la complexité se traduit par une dimension importante du vecteur d'état et on parle alors de *modèles de grande dimension*. L'approximation de modèle, encore appelée *réduction de modèle* dans ce cas, a pour but de trouver un modèle dont le vecteur d'état est plus petit que celui du modèle de grande dimension tel que les comportements entrée-sortie des deux modèles soient proches selon une certaine norme. La norme $\mathcal{H}_2$ a été largement considérée dans la littérature pour mesurer la qualité d'un modèle réduit. Cependant, la bande passante limitée des capteurs et des actionneurs ainsi que le fait qu'un modèle est généralement représentatif d'un système physique dans une certaine bande fréquentielle seulement, laissent penser qu'un modèle réduit dont le comportement est fidèle au modèle de grande dimension dans un intervalle de fréquences donné, peut être plus pertinent. C'est pourquoi, dans cette étude, *la norme $\mathcal{H}_2$ limitée en fréquence*, ou norme $\mathcal{H}_{2,\Omega}$, qui est simplement la restriction de la norme $\mathcal{H}_2$ sur un intervalle de fréquences Ω, a été considérée. En particulier, le problème qui vise à trouver un modèle réduit minimisant la norme $\mathcal{H}_{2,\Omega}$ de l'erreur d'approximation avec le modèle de grande dimension a été traité. Deux approches ont été proposées dans cette optique. La première est une approche empirique basée sur la modification d'une méthode sous-optimale pour l'approximation $\mathcal{H}_2$. En pratique, ses performances s'avèrent intéressantes et rivalisent avec certaines méthodes connues pour l'approximation de modèles sur intervalles de fréquences limités. La seconde est une méthode d'optimisation basée sur la formulation pôles-résidus de la norme $\mathcal{H}_{2,\Omega}$. Cette formulation généralise naturellement celle existante pour la norme $\mathcal{H}_2$ et permet également d'exprimer deux bornes supérieures sur la norme $\mathcal{H}_\infty$ d'un modèle LTI, ce qui est particulièrement intéressant dans le cadre de la réduction de modèles. Les conditions d'optimalité du premier ordre pour le problème d'approximation optimale en norme $\mathcal{H}_{2,\Omega}$ ont été exprimées et utilisées pour créer un algorithme de descente visant à trouver un minimum local au problème d'approximation. Couplée aux bornes sur la norme $\mathcal{H}_\infty$ de l'erreur d'approximation, cette méthode est utilisée pour le contrôle de modèles de grande dimension. D'un point de vue plus pratique, l'ensemble des méthodes proposées dans cette étude ont été appliquées, avec succès, dans un cadre industriel comme élément d'un processus global visant à contrôler un avion civil flexible.

Mots-clés : modèles linéaires invariants dans le temps, modèles de grande dimension, réduction de modèles, approximation de modèles, approximation de modèles sur intervalles de fréquences limités