Large-scale and infinite dimensional dynamical model approximation

Igor Pontes Duff Pereira

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 *11/01/2017* **par :**

Igor Pontes Duff Pereira

Large-scale and infinite dimensional dynamical model approximation

---

JURY

M. Olivi	Chargée de recherche	INRIA	Rapporteur
W. Michiels	Professeur	KU Leuven	Rapporteur
A. C. Antoulas	Professeur	Rice University	Examinateur
S. Gugercin	Professeur	Virginia Tech	Examinateur
J. Daafouz	Professeur	Université de Lorraine	Examinateur
A. Seuret	Chargé de recherche	LAAS CNRS	Examinateur
C. Poussot-Vassal	Ingénieur de recherche	ONERA	Directeur de thèse
C. Seren	Ingénieur de recherche	ONERA	Co-directeur de thèse

École doctorale et spécialité :

EDSYS : Automatique 4200046

Unité de Recherche :

ONERA - The French Aerospace Lab

Centre de Toulouse

Département de Commande des Systèmes et Dynamique du Vol

Directeurs de Thèse :

C. Poussot-Vassal (directeur de thèse)

C. Seren (co-directeur de thèse)

Rapporteurs :

Martine Olivi et *Wim Michiels*

-This Page Intentionally Left Blank-

Aos meus pais,
Ana Maria (in memoriam) e Carlos Eduardo,
por tudo que fizeram por mim
ao longo de minha vida

Acknowledgements

I am very grateful for all of the support and encouragement I have received from many people in the course of completing this work. First of all, many thanks go to my advisers Dr. Charles Poussot-Vassal and Dr. Cedric Seren for giving me the opportunity to achieve this Ph.D. thesis. I am also thankful for their guidance, constructive criticism and advice which have enabled me to complete this project successfully.

I would also like to thank Dr. Martine Olivi and Prof. Wim Michiels for accepting to review this manuscript. Their constructive comments have given me keys to improving this work and lead to extend it. Likewise, I would like to thank Prof. Athanasios C. Antoulas, Prof. Serkan Gugercin, Prof. Jamal Daafouz and Dr. Alexandre Seuret for accepting to be part of my jury.

I would like to thank as well Prof. Christopher Beattie and once again Prof. Serkan Gugercin for having me as a guest in the Department of Mathematics at Virginia Tech. This stay was very enjoyable and crucial for my training as a researcher. Also, I would like to thanks the Doctoral School EDSYS for making this stay possible through the international mobility scholarship.

I would also like to thanks my fellow Ph.D. students of the CSDV, who have helped to make these three years a very enjoyable and agreeable period through the numerous coffee breaks and other activities. Special thanks to Alvaro Perez, Flávio Ribeiro, Raquel Stella da Silva de Aguiar and Leandro Lustosa for the enormous support, great conversations and pleasant companionships.

I wish to express my gratitude to my girlfriend, Joana Eichenberger, for all the love, support, and comprehension.

Finally, I could not forget to thank in a special way all my family, who always supported me unconditionally in every possible way during my life. For this reason, I would like to express my kind wishes to everyone, especially to my father Carlos Eduardo, my mother-in-law Claudia and my brother Pedro.

Abstract

In the engineering area (e.g. aerospace, automotive, biology, circuits), dynamical systems are the basic framework used for modeling, controlling and analyzing a large variety of systems and phenomena. Due to the increasing use of dedicated computer-based modeling design software, numerical simulation turns to be more and more used to simulate a complex system or phenomenon and shorten both development time and cost. However, the need of an enhanced model accuracy inevitably leads to an increasing number of variables and resources to manage at the price of a high numerical cost. This counterpart is the justification for model reduction.

For linear time-invariant systems, several model reduction approaches have been effectively developed since the 60's. Among these, interpolation-based methods stand out due to their flexibility and low computational cost, making them a predestined candidate in the reduction of truly large-scale systems. Recent advances demonstrate ways to find reduction parameters that locally minimize the $\mathcal{H}_2$ norm of the mismatch error.

In general, a reduced-order approximation is considered to be a finite dimensional model. This representation is quite general and a wide range of linear dynamical systems can be converted in this form, at least in principle. However, in some cases, it may be more relevant to find reduced-order models having some more complex structures. As an example, some transport phenomena systems have their Hankel singular values which decay very slowly and are not easily approximated by a finite dimensional model. In addition, for some applications, it is valuable to have a structured reduced-order model which reproduces the physical behaviors. That is why, in this thesis, reduced-order models having delay structures have been more specifically considered.

This work has focused, on the one hand, in developing new model reduction techniques for reduced order models having delay structures, and, on the other hand, in finding new applications of model approximation. The major contribution of this thesis covers approximation topics and includes several contributions to the area of model reduction. A special attention was given to the $\mathcal{H}_2$ optimal model approximation problem for delayed structured models. For this purpose, some new theoretical and methodological results were derived and successfully applied to both academic and industrial benchmarks.

In addition, the last part of this manuscript is dedicated to the analysis of time-delayed systems stability using interpolatory methods. Some theoretical statements as well as an heuristic are developed enabling to estimate in a fast and accurate way the stability charts of those systems.

Résumé

Dans le domaine de l'ingénierie (par exemple l'aéronautique, l'automobile, la biologie, les circuits), les systèmes dynamiques sont le cadre de base utilisé pour modéliser, contrôler et analyser une grande variété de systèmes et de phénomènes. En raison de l'utilisation croissante de logiciels dédiés de modélisation par ordinateur, la simulation numérique devient de plus en plus utilisée pour simuler un système ou un phénomène complexe et raccourcir le temps de développement et le coût. Cependant, le besoin d'une précision de modèle améliorée conduit inévitablement à un nombre croissant de variables et de ressources à gérer au prix d'un coût numérique élevé. Cette contrepartie justifie la réduction du modèle.

Pour les systèmes linéaires invariant dans le temps, plusieurs approches de réduction de modèle ont été effectivement développées depuis les années 60. Parmi celles-ci, les méthodes basées sur l'interpolation se distinguent par leur souplesse et leur faible coût de calcul, ce qui en fait un candidat prédestiné à la réduction de systèmes véritablement à grande échelle. Les progrès récents démontrent des façons de trouver des paramètres de réduction qui minimisent localement la norme $\mathcal{H}_2$ de l'erreur d'incompatibilité.

En général, une approximation d'ordre réduit est considérée comme un modèle de dimension finie. Cette représentation est assez générale et une large gamme de systèmes dynamiques linéaires peut être convertie sous cette forme, du moins en principe. Cependant, dans certains cas, il peut être plus pertinent de trouver des modèles à ordre réduit ayant des structures plus complexes. A titre d'exemple, certains systèmes de phénomènes de transport ont leurs valeurs singulières Hankel qui se décomposent très lentement et ne sont pas facilement approchées par un modèle de dimension finie. En outre, pour certaines applications, il est intéressant de disposer d'un modèle structuré d'ordre réduit qui reproduit les comportements physiques. C'est pourquoi, dans cette thèse, les modèles à ordre réduit ayant des structures de retard ont été plus précisément considérés.

Ce travail a consisté, d'une part, à développer de nouvelles techniques de réduction de modèle pour des modèles à ordre réduit avec des structures de retard et, d'autre part, à trouver de nouvelles applications d'approximation de modèle. La contribution majeure de cette thèse couvre les sujets d'approximation et inclut plusieurs contributions au domaine de la réduction de modèle. Une attention particulière a été accordée au problème de l'approximation du modèle optimale pour les modèles structurés retardés. À cette fin, de nouveaux résultats théoriques et méthodologiques ont été obtenus et appliqués avec succès aux repères académiques et industriels.

De plus, la dernière partie de ce manuscrit est consacrée à l'analyse de la stabilité des systèmes retardés par des méthodes interpolatoires. Certaines déclarations théoriques ainsi qu'une heuristique sont développées permettant d'estimer de manière rapide et précise les diagrammes de stabilité de ces systèmes.

Notation

$\mathbb{N}$	Set of natural numbers including zero
$\mathbb{N}^*$	$\mathbb{N}\backslash\{0\}$
$\mathbb{Z}$	Set of integers
$\mathbb{R}$	Set of real numbers
$\mathbb{R}_+$	Set of strictly positive numbers
$\mathbb{R}_-$	Set of strictly negative numbers
$\mathbb{R}^n$	Set of real vectors with dimension n
$\mathbb{R}^*$	Set of non-null real numbers
$\mathbb{C}$	Set of complex numbers
i	Imaginary unit, *i.e.*, $i = \sqrt{-1}$
$\mathrm{Re}(s)$ and $\mathrm{Imag}(s)$	Real and imaginary parts of $s \in \mathbb{C}$ respectively
$\overline{s}$	Complex conjugate of $s \in \mathbb{C}$
$\mathbb{C}^+$	Open right-half plane
$\mathbb{C}^-$	Open left-half plane
$\mathbb{R}^{n_y \times n_u}$ and $\mathbb{C}^{n_y \times n_u}$	Set of real and complex matrices with dimensions $n_y \times n_u$
$A^T \in \mathbb{C}^{n_u \times n_y}$	Transpose of matrix $A \in \mathbb{C}^{n_y \times n_u}$
$A^* \in \mathbb{C}^{n_u \times n_y}$	Hermitian transpose of matrix $A \in \mathbb{C}^{n_y \times n_u}$
dim	Dimension of vector space
$\|\mathbf{b}\|_2$ or $\|A\|_2$	Vector and matrix 2 norm
$\|A\|_F$	Frobenius matrix norm
$\lambda_{\max}(A)$ and $\sigma_{\max}(A)$	Largest eigenvalue and largest singular value of matrix A
$\oplus$	Direct sum between vector spaces
$\mathbf{G}$	Dynamical system/full order model
$\hat{\mathbf{H}}$	Reduced order model
s	Laplace transform variable
$\mathbf{G}(s)$	Transfer function of the dynamical system $\mathbf{G}$
$\mathbf{G}'(s)$	Derivative of the function $\mathbf{G}(s)$ with respect to s
$\mathbf{G}_d$ and $\hat{\mathbf{H}}_d$	Dynamical system/Reduced order model having a delay structure
$\mathrm{order}(\mathbf{G})$	Dimension of the state space of a dynamical system $\mathbf{G}$.
$\mathcal{H}_2$ (or $\mathcal{H}_2(\mathbb{C}^+)$) and $\mathcal{H}_\infty$	Hardy spaces of holomorphic functions on $\mathbb{C}^+$
$\mathcal{H}_2(\mathbb{C}^-)$	Hardy space of holomorphic functions on $\mathbb{C}^-$
$\mathcal{L}_2(i\mathbb{R})$	$\mathcal{L}_2(i\mathbb{R}) = \mathcal{H}_2(\mathbb{C}^+) \oplus \mathcal{H}_2(\mathbb{C}^-)$
$\|\mathbf{G}\|_{\mathcal{H}_2}$ and $\|\mathbf{G}\|_{\mathcal{H}_\infty}$	$\mathcal{H}_2$ and $\mathcal{H}_\infty$ norm of the system $\mathbf{G} \in \mathcal{H}_2 \cap \mathcal{H}_\infty$
$\|\mathbf{G}\|_{\mathcal{L}_2(i\mathbb{R})}$	$\mathcal{L}_2(i\mathbb{R})$ norm of the system $\mathbf{G} \in \mathcal{L}_2(i\mathbb{R})$
$\mathbf{G}^s$	Stable part of $\mathbf{G} \in \mathcal{L}_2(i\mathbb{R})$, *i.e.*, $\mathrm{proj}_{\mathcal{H}_2(\mathbb{C}^+)}(\mathbf{G}) = \mathbf{G}^s$
$\mathbf{G}^a$	Unstable part of $\mathbf{G} \in \mathcal{L}_2(i\mathbb{R})$, *i.e.*, $\mathrm{proj}_{\mathcal{H}_2(\mathbb{C}^-)}(\mathbf{G}) = \mathbf{G}^s$
$\langle \mathbf{G}, \mathbf{H} \rangle_{\mathcal{H}_2}$ and $\langle \mathbf{G}, \mathbf{H} \rangle_{\mathcal{L}_2(i\mathbb{R})}$	$\mathcal{H}_2$ and $\mathcal{L}_2(i\mathbb{R})$ inner products
I_n	Identity matrix of size n

Acronyms

LTI	Linear Time-Invariant
TDS	Time-Delay Systems
DDEs	Delay-Differential Equation
LTI TDS	Linear Time-Invariant Time-Delay Systems governed by a retarded DDEs
ROM	Reduced order model (in general denoted by $\hat{\mathbf{H}}$ or $\hat{\mathbf{H}}_d$ with the hat symbol)

Algorithms

IRKA	Iterative Rational Krylov Algorithm
TF-IRKA	**IRKA** based on transfer function evaluations
MIMO IO-dIRKA	**IRKA** version including input and output delays (Chapter 5)
dTF-IRKA	**TF-IRKA** version including a single state-delay (Chapter 6)

Contents

Part I

Introduction

Chapter 1

Context and motivations

1.1 Context and problem formulation

Modeling is an essential step to well understand and interact with physical dynamical phenomena. It, among other, permits to analyze, simulate, optimize and control dynamical processes. Examples range from mechanical systems such as the resonance of a suspension bridge, the dynamics of flexible beam and the vibration of a drum, to biological ones such as the enzyme kinetics, the gene expression and the circulatory system. Models are extensively used in different domains in order to predict the dynamical behavior of a system without the need of any experimentation. Thereby, modeling has several advantages for the design of a new product, *e.g.*, avoidance in building expensive prototypes, predicting behavior in extreme scenarios and accelerating the design process.

The interest of a model lies in its ability to describe the reality as accurately as possible. Increasing demands on the accuracy typically bring about higher complexity of the model. In general, dynamical models are described by equations and their complexity is somehow linked to its number of equations and variables. Depending on the complexity of the physical system to be modeled, the means used to build the mathematical model and the desired accuracy of the model, this model can be more or less complex and representative. Although complex models have a high degree of representativeness with respect to reality, in practice, due to numerical limitations, they are problematic to manipulate. Actually, complex models are difficult to analyze and to control due to limited computational capabilities, storage constraints and finite machine precision. Therefore, a good model have to reach a trade-off between its accuracy and complexity.

However, some models are naturally designed to be very complex. This generally happens in two scenario : (i) the phenomena is governed by partial differential equations requiring a discretization (*e.g.*, transport phenomena, Black-Scholes equations, quantum mechanics, etc) (ii) the coupling of a large set of components (*e.g.*, electronic microchip, the international space station, satellite, etc). Numerical modeling tools (such as identification methods, finite elements methods, etc.) do not necessarily enable to restrain the complexity without loosing too much information leading to *large-scale models* (or high-fidelity models). In this context, an a posteriori method is generally preferred to reduce the complexity of the original, high-fidelity representation.

Model approximation (or model reduction) techniques serve to this purpose. Their general goal is to replace an existing high fidelity model by another one which is just as well suited for the engineering task but of lower complexity. Thereby, efficiency can be dramatically increased, as comparable results can be produced in far less time. The difficulty is to identify and extract

the components of the large-scale model that are pertinent while discarding the unnecessary part of the model. As mathematical models may take very different forms, depending on the kind of system they describe, every type of model requires customized techniques for its simplification.

In this thesis, continuous dynamical systems are considered. They can be represented, for example, by ordinary differential equations, differential algebraic equations or partial differential equations which can be linear, non-linear, time-invariant or time-variant. Among the possible representations, Linear Time Invariant (LTI) models are widely used, both in industry and research. Indeed, for many physical systems, they are sufficiently representative around an equilibrium point and numerous tools exist in order to analyze 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 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 by [Moore, 1981] and [Glover, 1984]. Then, the extensive use of numerical modeling 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-scale systems due to their inherent numerical complexity. Hence, original techniques that are numerically cheaper have been developed.

In general, a reduced-order approximation is considered to be a finite dimensional model $\hat{\mathbf{H}}$ represented by a differential equation as follows[1] :

$$\hat{\mathbf{H}} := \left\{ \begin{array}{rcl} \hat{E}\dot{\mathbf{x}}(t) & = & \hat{A}\mathbf{x}(t) + \hat{B}\mathbf{u}(t) \\ \mathbf{y}(t) & = & \hat{C}\mathbf{x}(t) + \hat{D}\mathbf{u}(t) \end{array} \right. \tag{1.1}$$

This representation is quite general and a wide range of linear dynamical systems can be converted to this form, at least in principle. However, in some cases, it may be more relevant to find Reduced-Order Model (ROM) having some more complex structures. Indeed, (i) some transport phenomena systems have their Hankel singular values which decay very slowly and are not easily approximated by a ROM as (1.1), (ii) for some applications, it is valuable to have structures in the reduced-order model which represents the real behavior, *e.g.*, input and output delays or second-order behavior. In addition, the problem of model approximation for a family of reduced-order models more general than (1.1) is an improvement. That is why we have chosen in this thesis, to address the problem of approximating systems by time-delay structured models.

This thesis tries to bring together the methodology used in optimal $\mathcal{H}_2$ model approximation [Gugercin et al., 2008] and the Loewner framework [Mayo and Antoulas, 2007], in order to find model approximations presenting time-delay structures. Let us now introduce some examples of system that will be used in this work.

1.2 Motivating examples

In this section, we present three examples illustrating the model approximation need and used along this thesis as numerical applications. The first one represents a large-scale system from [Leibfritz and Lipinski, 2003], the second one corresponds to a system having an input-delay behaviors from [Beattie and Gugercin, 2011] and the third one [Dalmas et al., 2016] is an infinite dimensional model represented by its irrational transfer function.

[1] The notation will be formally introduced in chapter 2.

Clamped beam model

This example arises from the theory of elasticity. A great amount of flexible structures are modeled by the fundamental mechanics laws. In particular, beam structures are usually modelled by the Euler–Bernoulli beam equation. This is a partial differential equation and is used for engineering purposes like simulation, optimization, analysis and control, for which a discretization appears necessary.

The model considered here represents one of this discretized (finite dimensional) beam model obtained by spatial discretization of an appropriate partial differential equation. The input represents the force applied to the structure at the free end, and the output is the resulting displacement. This model has 348 states whose poles are given in Figure 1.1. Its frequency behavior is given by its Bode plot in Figure 1.2.

We have chosen this particular example to illustrate possible difficulties that can arise throughout the reduction process. As it is shown, the original transfer function exhibits many peaks. Interpreting model approximation as an interpolation problem, coming up with a good reduced interpolant is very difficult since the function is not very smooth and, hence, a large number of interpolation points is needed for a successful reproduction. Moreover, a central question for the performance of a reduced-order model is the location of the interpolation points and throughout this thesis we provide several (known) statements about optimality with respect to a specific accuracy metric.

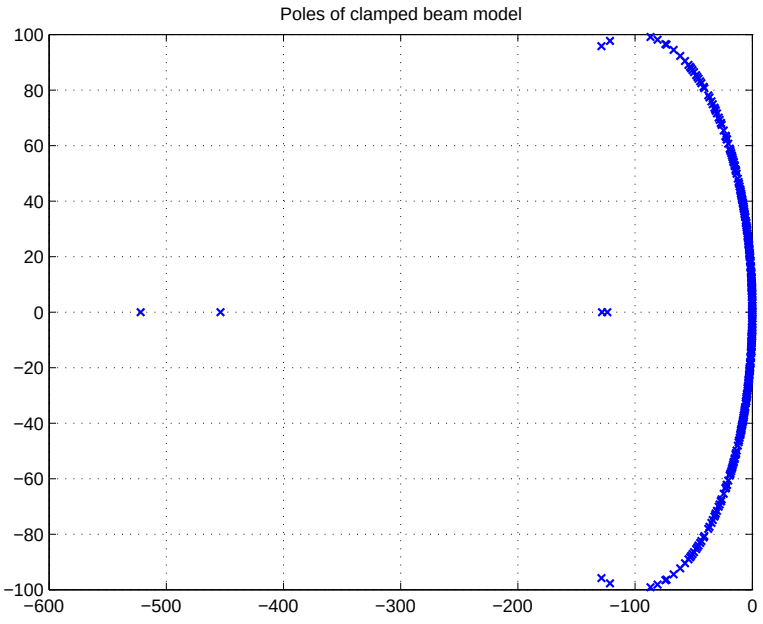

Figure 1.1: Poles of the discretized clamped beam model of order 348.

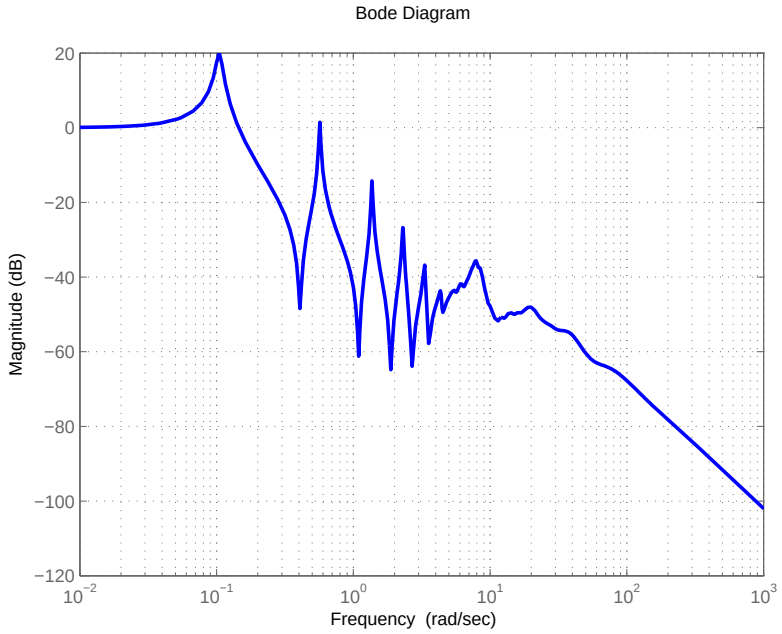

Figure 1.2: Bode diagram of the discretized clamped beam model of order 348.

Ladder network system

The following example is the Ladder network system proposed in [Beattie and Gugercin, 2011] and [Gugercin et al., 2012] in the context of Port-Hamiltonian systems. This model corresponds to a linear ladder network circuit as it is shown on Figure 1.3. In the article, the system is modeled by a finite dimensional state-space representation, *i.e.*, it is described by :

$$\mathbf{G}_{Ladder} := \left\{ \begin{array}{rcl} E\dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = & C\mathbf{x}(t) \end{array} \right. \tag{1.2}$$

and it has 100 states. Its impulse response is represented on Figure 1.4. Even if the system is represented by a finite dimensional system, it has an intrinsic input-delay behavior (with delay around 19s). This example leads to consider and look for an approximation having an input-delay structure.

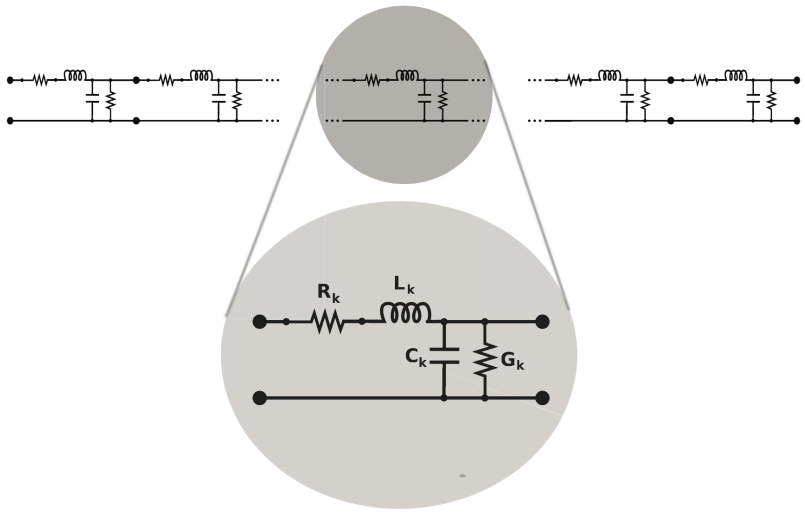

Figure 1.3: Ladder network circuit topology. Figure from [Beattie and Gugercin, 2011].

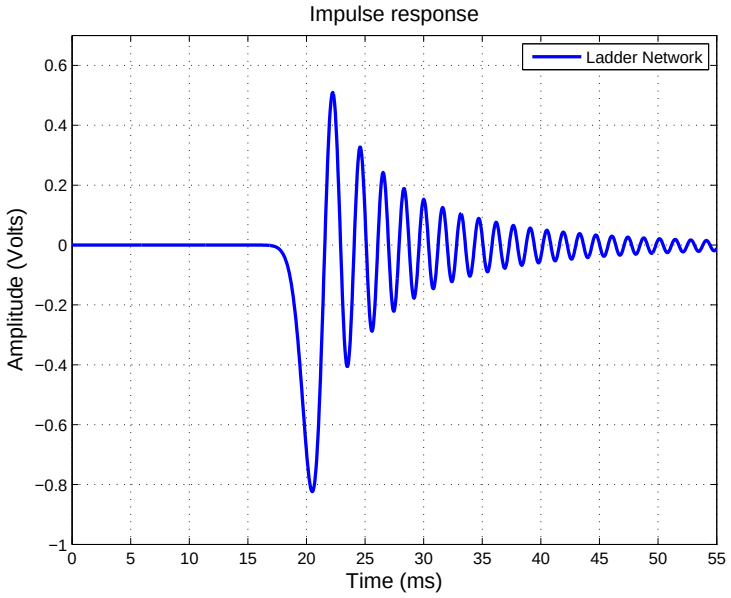

Figure 1.4: Impulse response of Ladder systems with intrinsic delay behavior.

Open channel flow for hydroelectricity

Run-of-the-river power plants rely on open-channel hydraulic systems. These large distributed systems are characterized by non-linearities and operating point dependent dynamic behaviors. The following example, taken from [Dalmas et al., 2016], model an open-channel hydraulic sys-

tems. The physical equations on which this work is done are Saint-Venant equations applied to a non-rectangular cross section channel, *i.e.*,

$$\begin{aligned} \frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} &= 0 \\ \frac{\partial Q}{\partial t} + \frac{\partial (Q^2/S)}{\partial x} + gS\frac{\partial H}{\partial x} &= gS(I-J), \end{aligned} \tag{1.3}$$

where $x \in [0; L]$ is the spatial variable, $H(x,t)$ the water depth, $S(x,t)$ the wetted area, and $Q(x,t)$ the discharge. These later are two coupled non-linear hyperbolic partial differential equations which are linearized (see [Dalmas et al., 2016]) and converted, by means of the Laplace transformation, into the following parameter dependent irrational transfer function :

$$\mathbf{H}_{flow}(s, Q_0) = \begin{bmatrix} \mathbf{G_e}(s, Q_0) & -\mathbf{G_s}(s, Q_0) \end{bmatrix} \begin{bmatrix} \mathbf{q_e}(s) \\ \mathbf{q_s}(s) \end{bmatrix} \tag{1.4}$$

with

$$\begin{aligned} \mathbf{G_e}(s, Q_0) &= \frac{\lambda_1(s)e^{\lambda_2(s)L+\lambda_1(s)x} - \lambda_2(s)e^{\lambda_1(s)L+\lambda_2(s)x}}{B_0 s(e^{\lambda_1(s)L} - e^{\lambda_2(s)L})} \\ \mathbf{G_s}(s, Q_0) &= \frac{\lambda_1(s)e^{\lambda_1(s)x} - \lambda_2(s)e^{\lambda_2(s)x}}{B_0 s(e^{\lambda_1(s)L} - e^{\lambda_2(s)L})} \end{aligned} \tag{1.5}$$

and

$$\lambda_{1,2}(s) = \frac{V_0 s + \varphi_0 \pm \sqrt{c_0^2 s^2 + \Phi_0 s + \varphi_0^2}}{\delta_0}, \tag{1.6}$$

where $\mathbf{q_e}$ and $\mathbf{q_s}$, representing the inflow and outflow receptively, are the inputs, and $\mathbf{H}_{flow}$, representing the measured water depth, is the output. The parameters x, B_0, L, V_0, ϕ_0, Φ_0, c_0, φ_0 and δ_0 are given by the physics of the problem and some nominal values can be found in [Dalmas et al., 2016]. In addition, $\mathbf{H}_{flow}$ depends on the nominal flow Q_0. The transfer function (1.4) is clearly irrational. For a frozen Q_0, Figures 1.5 and 1.6 shows its frequency behavior.

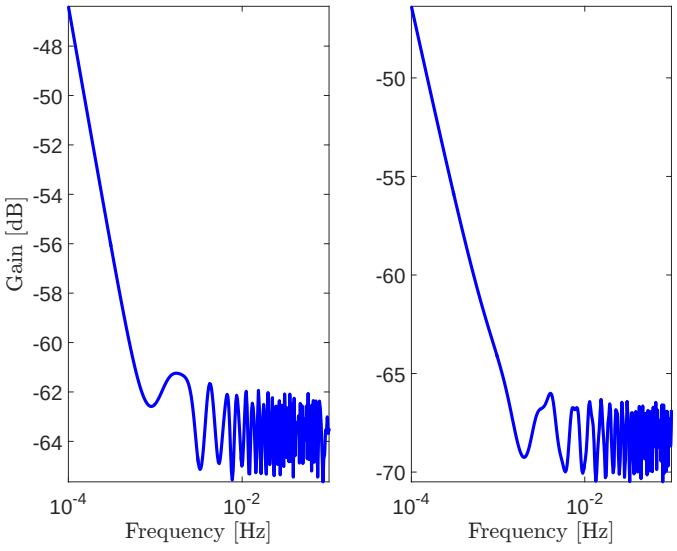

Figure 1.5: Bode magnitude diagram of the irrational transfer function of (1.4).

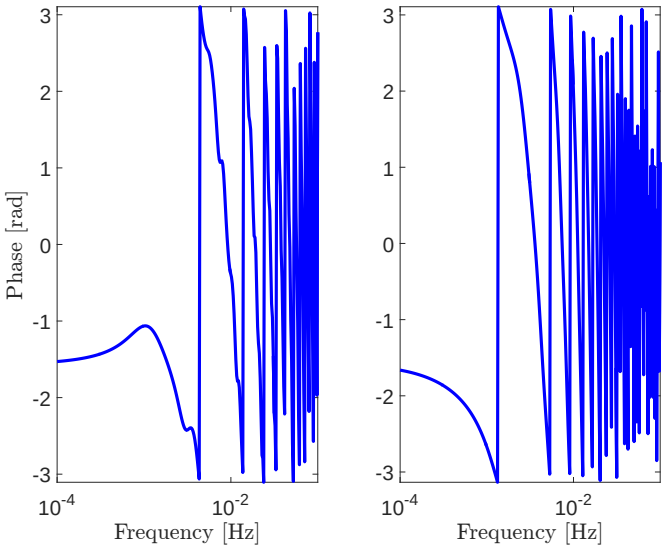

Figure 1.6: Bode phase diagram of the irrational transfer function of (1.4).

Experimentation and a great number of publications show that, even if this model does not present an explicit input/output delay, it behaves as having an intrinsic input-delay. Hence, it may be relevant to find reduced-order models having some input/output delays.

1.3 Problem formulation

Classical model approximation problem

The classical model approximation problem can be stated as follows :

Problem 1.1 (**Classical model approximation problem**). *Given a continuous LTI model* $\mathbf{G}$[a] *of order* N *represented by:*

$$\mathbf{G} := \left\{ \begin{array}{rcl} E\dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = & C\mathbf{x}(t) + D\mathbf{u}(t) \end{array} \right. \tag{1.7}$$

where $E, A \in \mathbb{R}^{N\times N}$, $B \in \mathbb{R}^{N\times n_u}$, $C \in \mathbb{R}^{n_y\times N}$, $D \in \mathbb{R}^{n_y\times n_u}$, *the model approximation problem consists in finding* $\hat{\mathbf{H}}$ *of order* $n \ll N$ *given by the following realization:*

$$\hat{\mathbf{H}} := \left\{ \begin{array}{rcl} \hat{E}\dot{\hat{\mathbf{x}}}(t) & = & \hat{A}\hat{\mathbf{x}}(t) + \hat{B}\hat{\mathbf{u}}(t) \\ \hat{\mathbf{y}}(t) & = & \hat{C}\hat{\mathbf{x}}(t) + \hat{D}\hat{\mathbf{u}}(t) \end{array} \right. \tag{1.8}$$

where $\hat{E}, \hat{A} \in \mathbb{R}^{n\times n}$, $\hat{B} \in \mathbb{R}^{n\times n_u}$, $\hat{C} \in \mathbb{R}^{n_y\times n}$, $\hat{D} \in \mathbb{R}^{n_y\times n_u}$, *which accurately reproduces the behavior of the full-order model* $\mathbf{G}$[b].

[a] Those notions will be formally defined in Chapter 2

[b] At this point, to reproduce the behavior seems a very vague sentence. This notion will be quantified with the introduction of system norms. That will be presented in Chapter 2.

If finding the model approximation is more computationally expensive than to work directly with the full-order model $\mathbf{G}$, then it is not worth approximating it. That is why a procedure intending to construct a model approximation $\hat{\mathbf{H}}$ should satisfy the following properties.

Remark 1.2 (Properties of a model approximation technique)**.** *The procedure to find the model approximation* $\hat{\mathbf{H}}$ *should satisfy the following properties:*

1. *The approximation error* $\|\mathbf{y}-\hat{\mathbf{y}}\|$[a] *is* small*. In addition, the existence of a* global *error bound is not required but recommended;*
2. *Systems properties (like* stability*) should be preserved;*
3. *The procedure have to be computationally efficient.*

[a]This notion will be formally defined in Chapter 2.

The reader should note that Problem 1.1 does not take into account systems $\mathbf{G}$ which cannot be represented by a finite dimensional realization as (1.7), *e.g.*, time-delay systems and irrational transfer functions. In general, the family of infinite dimensional models cannot be represented by (1.7). This family includes time-delay systems, transport equations, and more generally, systems governed by a partial differential equation.

In this thesis, we considered that the full-order model $\mathbf{G}$ that will be approximated can be represented by one of the three representations:

1. **Finite dimensional realization**: In this case, system $\mathbf{G}$ is represented by its time-domain realization of the form (1.7). In this case, its transfer function[2] is given by :

$$\mathbf{G}(s) = C(sE - A)^{-1}B + D \tag{1.9}$$

   and it is a matrix rational function.

2. **Irrational transfer function**: In this case, system $\mathbf{G}$ is represented by its frequency behavior, *i.e.*, by a matrix complex function :

$$\begin{array}{rccc} \mathbf{G} : \mathbb{C} & \to & \mathbb{C}^{n_y \times n_u} \\ s & \mapsto & \mathbf{G}(s) \end{array}, \tag{1.10}$$

   where n_u is the number of inputs and n_y is the number of outputs. Notice that if $\mathbf{G}(s)$ is a rational transfer function, then it is possible to find a finite dimensional realization (E, A, B, C) such that $C(sE - A)^{-1}B = \mathbf{G}(s)$. Otherwise, $\mathbf{G}(s)$ is said to be an irrational transfer function.

3. **Frequency data**: In this case, $\mathbf{G}$ is considered to be represented by a transfer function $\mathbf{G}(s)$ for which only a partial knowledge is considered to be known. As an example, for a set of different complex numbers $\{\sigma_k\}_{k=1}^n$, the only knowledge available from $\mathbf{G}$ is the values of

$$\mathbf{G}(\sigma_k) = \mathbf{w}_k \in \mathbb{C}^{n_y \times n_u}, \quad \text{for} \quad k = 1, \ldots, n.^{3} \tag{1.11}$$

Additionally, we are interested in finding a reduced-order model $\hat{\mathbf{H}}$ in a family of models that is bigger than the one modeled by (1.8). We have chosen to work with the family of models having

[2]defined formally in Chapter 2.

[3]In Chapter 3, the notion of tangential interpolation data will be introduced which is weaker than point-wise interpolation of (1.11).

a time-delay structure, *e.g.*, input/output delays and state-delays. For this family of models, a great variety of simulation, analysis, optimization and control techniques were already developed [Richard, 2003; Briat, 2015; Michiels and Niculescu, 2014] and a great amount of systems can be represented by Time-Delay Systems (TDS).

In this work we consider a model approximation problem which generalizes Problem 1.1

Problem 1.3 (Extended model approximation problem). *Let* $\mathbf{G}$ *be a continuous LTI system represented by one of the three possibilities below (finite dimensional realization, irrational tranfer function or frequency data). The model approximation problem consists in finding* $\hat{\mathbf{H}}$ *from a subset of the TDS family*[a] *which accurately reproduces the behavior of the full-order model* $\mathbf{G}$*.*

[a] This will be detailed later in this monograph. Reduced order models having input and output delay structures are considered in Chapter 5. State-delay structures are considered in Chapters 6 and 7.

Evidently, the procedure to construct $\hat{\mathbf{H}}$ for Problem 1.3 should satisfy the properties from Remark 1.2.

1.4 Overview of the contributions

In this thesis, we were particularly interested in infinite dimensional systems approximation, and more specifically, time-delay and irrational ones. Some examples of this class of systems can be found in the context of networked systems, chemical reactions, traffic jam and heating systems. Since for this class of systems, standard analysis and control methods are not directly applicable, it might be very appealing to approximate them by finite order models. Even if many dedicated approaches have been derived to handle infinite dimensional problems, most of them are limited to systems with low order state space vector and associated methods are not scalable when the order of the model increase.

The aim of this thesis is to investigate the approximation of large-scale and infinite dimensional systems. We are particularly interested in tackling the problem using **data-driven** (**Loewner**) and **interpolatory** methods. As made clearer in what follows, this work has focused, on one hand, in finding new applications of those model approximation techniques, *e.g.*, estimation of stability charts for a time-delay system, on the other hand, to develop new model reduction techniques for reduced order models based on more complex structures (integrating delays for example). We believe that the originality of this work comes from the combination of results from two different fields: time-delay systems and model approximation. The following list summarizes the main contributions and publications.

$\mathcal{H}_2$ model approximation for input/output delay reduced order model

Given $\mathbf{G}$, a finite dimensional stable model, we have studied the $\mathcal{H}_2$ optimal model approximation problem when the ROM includes input/output delays, *i.e.*,

$$\hat{\mathbf{H}}_\mathbf{d} = \Delta_o(s)\hat{\mathbf{C}}(\hat{\mathbf{E}}s - \hat{\mathbf{A}})^{-1}\hat{\mathbf{B}}\Delta_i(s).$$

The $\Delta_{i/o}(s)$ blocks represent here multiple input and output delays having the form $\Delta_{i/o}(s) = \mathbf{diag}(e^{-\tau_1 s}, \ldots, e^{-\tau_M s})$ with $\{\tau_1, \ldots, \tau_M\} \in \mathbb{R}_+^*$. We derived the $\mathcal{H}_2$ optimality conditions for this problem based on the pole/residue decomposition. Finally, a two stage algorithm, in order to practically obtain such an approximation, has been proposed. Those results were submitted to [Pontes Duff et al., 2016b] and presented in Chapter 5.

Data-based model interpolation for single time-delay systems

We have developed an extension of the Loewner framework, initially settled for delay-free reduced models, to the single state-delay case, *i.e.*, when the reduced order model has the following delay structure :

$$\hat{\mathbf{H}}_{\mathbf{d}}(s) = \hat{\mathbf{C}}(\hat{\mathbf{E}}s - \hat{\mathbf{A}}e^{-s\tau})^{-1}\hat{\mathbf{B}}.$$

Secondly, using the Lambert function, a new algorithm called **dTF-IRKA** enabling the construction of a reduced order delay-model satisfying some interpolation conditions has been developed. Theoretical results and several numerical examples are presented in [Pontes Duff et al., 2015a] and in Chapter 6.

More general $\mathcal{H}_2$-optimality conditions :

The $\mathcal{H}_2$-optimal model reduction problem was revisited when the reduced system has a more general pole-residue structure and new $\mathcal{H}_2$-optimality conditions were obtained. These conditions are no longer simple interpolation conditions as in the finite-dimensional case but rather interpolation of series which depends on the spectral structure of the reduced order model. These results are presented in [Pontes Duff et al., 2016a] and in Chapter 7.

Model reduction and stability charts for time-delay systems:

The problem of determining approximate stability regions for large-scale time-delay systems is treated using model approximation techniques. To achieve this, an $\mathcal{H}_2$-oriented approximation algorithm (see [Beattie and Gugercin, 2012]) is used in order to estimate the stability of a given TDS. We show how model reduction can be efficiently used to approximate time-delay systems with multiple delays and estimate their stability regions with respect to those delays. Theoretical results and several numerical examples are presented in [Pontes Duff et al., 2015b] and in Chapters 8 and 9.

1.5 Manuscript outline

This manuscript is divided into four parts and ten chapters. Part I is dedicated to the introduction, notation and state of the art of model approximation. Part II is dedicated to the main theoretical developments of this thesis, *i.e.*, model approximation by structured time-delay reduced order models. Part III presents a model approximation based framework to estimate stability of a TDS. Finally, Part IV concludes the manuscript with a global analysis and provides some perspectives. The chapter are organized as follows:

Part I: Introduction

Chapter 2: Fundamentals of LTI system theory

This chapter aims at recalling some general elements about linear continuous systems theory and introduces the general notations used along this thesis. In particular, two elements that form the basis of this thesis are recalled : the partial fraction decomposition of a finite dimensional system and the $\mathcal{H}_2$ norm of LTI models. In addition, a very brief introduction on time-delay systems are provided. Experts in control theory might skip this chapter.

Chapter 3: State of the art of model approximation

In this chapter, some well-known model approximation techniques based either on state-space truncation or tangential interpolation are recalled. In particular, the modal and balanced truncation are recalled since they are the most popular approximation methods. Model approximation by tangential interpolation is also presented. Finally the Loewner framework, which enables to construct reduced-order models via data interpolation, is described. This framework can be applied in a larger class of system representations, including irrational transfer functions. Readers who are familiar with the model approximation literature might skip this chapter.

Part II: Model approximation by structured time-delay reduced order models

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

The optimal $\mathcal{H}_2$ approximation problem is introduced together with some methods that address it. Firstly, a state of the art on $\mathcal{H}_2$ approximation is presented. Then, the $\mathcal{H}_2$ necessary optimality conditions are derived in two different ways. Finally, two algorithms achieving $\mathcal{H}_2$ model approximations are presented: (i) **IRKA**, which is based on the interpolation of the large-scale model through projection on some specific Krylov subspaces and (ii) **TF-IRKA**, which is based on the interpolation of the transfer functions through the Loewner framework. This chapter plays a pivotal role in this manuscript, since it links recent results on $\mathcal{H}_2$ model approximation while introducing the necessary elements required in the developments of the main contributions reported in this part, detailed in Chapters 5, 6 and 7.

Chapter 5: Optimal $\mathcal{H}_2$ model approximation by input/output-delay structure reduced models

In this chapter, the $\mathcal{H}_2$ optimal approximation problem by a finite dimensional model including input/output delays, is addressed. Firstly, the $\mathcal{H}_2$ inner product formulas are revisited in the case where the models have input/output delays. Secondly, the approximation error is formulated as a function of the pole/residue decomposition. Then, by taking the gradient of the error, the $\mathcal{H}_2$-optimality conditions of the approximation problem are obtained as an extension of the tangential interpolatory conditions in the delay-free case. It is also demonstrated that for fixed delay values, this problem can be recast as a delay-free one. The approach followed in this first part is similar to what has been done in the delay-free case in Chapter 4. The similarities and the key differences from the delay-free case are highlighted. The results are compared with some simpler interpolation conditions. Finally, an iterative algorithm, entitled **IO-dIRKA:**, is sketched out and numerical results assess the theoretical contributions.

Chapter 6: Data-driven model approximation by single state-delay structure reduced order models

In this chapter, the Loewner framework is revisited in the case where the interpolating model might be a single state-delay dependent model. To this aim, the Loewner framework, initially settled for delay-free realization, is firstly generalized to the single-delay case. Secondly, the finite dimensional inspired interpolation conditions are established through the use of the Lambert functions. Finally, a iterative scheme, named **dTF-IRKA**, similar to the **TF-IRKA**, is proposed to reach a part of the aforementioned optimality conditions. The proposed method validity and interest are assessed on different numerical examples.

Chapter 7: $\mathcal{H}_2$ optimality conditions derivation for state-delay reduced models

In this chapter, we first attempt to generalize the $\mathcal{H}_2$ optimal interpolation conditions for more general reduced order models. To this aim, we first expose the necessary optimality conditions in the case where the reduced system is of dimension one and have a single state delay structure. This can be viewed as a first step towards the $\mathcal{H}_2$ optimal model approximation where the reduced system corresponds to an infinite dimensional one. Finally, we illustrate the results with an academic example.

Part III: Stability charts of time-delay systems and model approximation

In this part of the thesis, corresponding to Chapters 8 and 9, we apply model approximation in order to analyze the stability of time-delay systems. As they represent infinite dimensional systems, the stability analysis of any time-delay system is a complex problem and is not simply established solving a matrix eigenvalue problem. Our philosophy here is to address the stability problem using interpolatory model approximation. Some theoretical results and heuristics are developed and were partially presented in [Pontes Duff et al., 2015b].

Chapter 8: Model approximation framework for evaluating time-delay systems' stability

In this chapter, a model approximation framework is developed in order to evaluate the stability of a time-delay system. Henceforth, systems are considered to be elements of $\mathcal{L}_2(i\mathbb{R})$ (instead of $\mathcal{H}_2$). The set of stable and unstable systems are characterized as subsets of $\mathcal{L}_2(i\mathbb{R})$, leading to some topological and approximation results. Thereby, a certificate of instability is derived based on the numerical estimation of the $\mathcal{L}_2(i\mathbb{R})$ norm. Equipped with those results, one can estimate the stability of any time-delay systems by finding a model approximation with is good enough in the sense of $\mathcal{L}_2(i\mathbb{R})$. Finally, interpolation-based techniques are used to build such approximations and some numerical results are presented.

Chapter 9: Stability chart approximation from interpolatory methods

This chapter is dedicated to the use of model approximation techniques to determine the stability charts of time-delay systems. We give a major attention to the algorithm **TF-IRKA** (see [Beattie and Gugercin, 2012]), which enables to achieve the $\mathcal{H}_2$ optimal model approximation. This algorithm has been shown to be well suited for the approximation of infinite-dimensional systems into finite-dimensional ones. Some theoretical arguments are, then, provided to justify the stability-property preservation in the reduced order model. Discussions regarding the adaptation of existing algorithms to the considered problem are also provided. Several numerical examples illustrate the efficiency and the accuracy of the approach, including in the large-scale setting.

Part IV: Conclusion

Chapter 10: Discussion and perspectives

This chapter aims at recalling the contributions of this thesis as well as their limitations. Additionally, some short-term extensions are presented and some long-term outlook concerning the extension of the methods and tools developed in this thesis to other type of models.

Chapter 2

Fundamentals of LTI system theory

In this chapter, some central concepts about continuous Linear Time Invariant (LTI) systems theory and the associated notations are recalled. We also introduce several objects associated with LTI systems that will be used throughout this thesis. The material is standard and is covered in many monographs such as [Antoulas, 2005; Partington, 2004, 1997]. Additional references are mentioned through the text when required.

This chapter is decomposed in three different sections. In the first one, the notion of LTI model is defined by the convolution integral. Then, the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ spaces are defined and some useful results are recalled. In the second section, some theoretical results related to finite dimensional LTI models are given, including the partial fraction decomposition and the notion of gramians. Finally, the third section recalls some facts about time-delay models, pointing out two structures: input/output delays and state-delay models.

Contents

2.1 Signals, systems and norms

In Section 2.1, some elements about the general theory of LTI models are recalled. The subjects that are highlighted here are: (i) the representation of an LTI model as a convolution with the impulse response, (ii) the Hardy spaces $\mathcal{H}_2$ and $\mathcal{L}_2(i\mathbb{R})$ and (iii) the inequalities involving systems and signals.

2.1.1 Signals and norms

Let us define a *signal* as a Lebesgue measurable function $\mathbf{f}$ that maps the real numbers $\mathbb{R}$ to $\mathbb{C}^n$. The set of signals is defined as :

$$\mathcal{S}_n = \{\mathbf{f} : \mathbb{R} \to \mathbb{C}^n : \mathbf{f} \quad \text{measurable}\}. \tag{2.1}$$

A particular $\mathcal{S}_n$ subspace is the signals with finite energy denoted by:

$$L_2^n(-\infty, \infty) = \{\mathbf{f} \in \mathcal{S}_n : \int_{-\infty}^{\infty} \|\mathbf{f}(t)\|_2^2 dt < \infty\}. \tag{2.2}$$

The space $L_2^n(-\infty, \infty)$ is a Hilbert space equipped with the following *inner product* :

$$\langle \mathbf{f}, \mathbf{g} \rangle_{L_2} = \int_{-\infty}^{+\infty} \mathbf{g}^*(t)\mathbf{f}(t)dt, \text{ for } \mathbf{f}, \mathbf{g} \in L_2^n(-\infty, \infty) \tag{2.3}$$

whose induced norm is given by $\|\mathbf{f}\|_{L_2} = \langle \mathbf{f}, \mathbf{f} \rangle_{L_2}^{\frac{1}{2}} = \Big(\int_{-\infty}^{\infty} \|\mathbf{f}(t)\|_2^2 dt \Big)^{\frac{1}{2}}$. We often call the L_2 norm of a signal its *energy*.

From a practical point of view, it is convenient to define the two following subspaces :

$$\begin{array}{ll} L_2^n[0, \infty) & = \{\mathbf{f} \in L_2^n(-\infty, \infty) : f(t) = 0 \text{ for all } t < 0\} \text{ and} \\ L_2^n(-\infty, 0] & = \{\mathbf{f} \in L_2^n(-\infty, \infty) : f(t) = 0 \text{ for all } t > 0\}. \end{array} \tag{2.4}$$

The $L_2^n[0, \infty)$ and $L_2^n(-\infty, 0]$ spaces are also Hilbert spaces with respect to the inner product (2.3). Moreover,

$$\Big(L_2^n[0, \infty)\Big)^{\perp} := \{\mathbf{f} \in L_2^n(-\infty, \infty) : \langle \mathbf{f}, \mathbf{g} \rangle_{L_2} = 0, \forall \mathbf{g} \in L_2^n[0, \infty)\} = L_2^n(-\infty, 0]$$

and

$$L_2^n(-\infty, \infty) = L_2^n[0, \infty) \oplus L_2^n(-\infty, 0].$$

In what follows, we give the definition of an LTI system based on the convolution product between given signals.

2.1.2 Systems and norms

A *dynamical system* or *model* consists in the mathematical equations that represent a physical process that evolves in time. In this monograph, we are particularly interested in continuous LTI models. In addition, in this work, the words *system* and *model* will be considered as synonyms. For a more general axiomatic definition of system/model see [Sontag, 2013] and the references therein.

Time-domain representation of LTI models

A continuous LTI model $\mathbf{H}$ is an "input-output" map which associates to an input signal $\mathbf{u} \in \mathcal{D}(\mathcal{S}_{n_u})$, where $\mathcal{D}(\mathcal{S}_{n_u})$ is a sub-domain of $\mathcal{S}_{n_u}$, an output one $\mathbf{y} \in \mathcal{S}_{n_y}$ by means of convolution operation

$$\begin{array}{rlll} \mathbf{H}: & \mathcal{D}(\mathcal{S}_{n_u}) & \mapsto \mathcal{S}_{n_y} & \\ & \mathbf{u}(t) & \mapsto \mathbf{y}(t) = \displaystyle\int_{-\infty}^{\infty} \mathbf{h}(t-\tau)\mathbf{u}(\tau)d\tau = \mathbf{h} * \mathbf{u}(t), \end{array}$$

where $\mathbf{h}(t) \in \mathbb{R}^{n_y \times n_u}$ is called the impulse response of the system. Models with $n_u > 1$ or $n_y > 1$ are referred to as Multi-Input Multi-Output (MIMO) models; the special case $n_u = n_y = 1$ is called Single-Input Single-Output (SISO) models. A LTI model is causal if and only if $\mathbf{h}(t) = 0, t < 0$ and, in this case, the convolution product with an input $\mathbf{u}$ ($\mathbf{u}(t) = 0, t < 0$) is reduced to the integral

$$\mathbf{y}(t) = \int_0^t \mathbf{h}(t-\tau)\mathbf{u}(\tau)d\tau = \int_0^t \mathbf{h}(\tau)\mathbf{u}(t-\tau)d\tau \tag{2.5}$$

For now on, we suppose that all systems are causal.

Frequency-domain representation of LTI models

The *unilateral Laplace transform* $\mathscr{L}$ of the impulse response $\mathbf{h}$ of an LTI model is defined as :

$$\mathbf{H}(s) = \mathscr{L}(\mathbf{h}) = \int_0^\infty \mathbf{h}(\tau)e^{-s\tau}d\tau \in \mathbb{C}^{n_y \times n_u}, \tag{2.6}$$

and taking the Laplace transformations of (2.5), one obtains

$$\mathbf{Y}(s) = \mathbf{H}(s)\mathbf{U}(s), \tag{2.7}$$

where $\mathbf{Y}(s)$ and $\mathbf{U}(s)$ are the Laplace transform of $\mathbf{y}(t)$ and $\mathbf{u}(t)$. The $n_y \times n_u$ matrix function $\mathbf{H}(s)$ is called the *transfer function* of the LTI model. A model is said to be *real* if its impulse response matrix $\mathbf{h}(t)$ is a real valued matrix function. As a consequence, the transfer function $\mathbf{H}(s) = \mathscr{L}(\mathbf{h})$ satisfies for $s \in \mathbb{C}$

$$\overline{\mathbf{H}(s)} = \mathbf{H}(\overline{s}). \tag{2.8}$$

In this work, all models are assumed to be real. A transfer function $\mathbf{H}(s)$ is said to represent a *proper model* if for sufficiently large ρ (see [Curtain and Morris, 2009]) :

$$\sup_{Re(s)\geq 0, |s|\geq\rho} \|\mathbf{H}(s)\|_2 < \infty.$$

It is said to be a *strictly proper model* if the limit of $\mathbf{H}(s)$ at infinity is 0.

In this work, an LTI system $\mathbf{H}$ is said to be *stable* if and only if its transfer function $\mathbf{H}(s)$ is bounded and analytic on $\mathbb{C}^+$, *i.e.*, it has no poles in the closed right half-plane. Causal stable models are represented by transfer functions that are analytic in the right half-plane. In addition, the family of stable models can be regarded as a functional space of analytic functions on the right half-plane and is therefore a Hardy space. These spaces are detailed in the following subsection.

2.1.3 Hardy spaces and norms

Hardy spaces are linear spaces of functions which are bounded and analytic in a certain region of the complex plane. These spaces are equipped with a *norm* $\|\ .\ \|_{\mathcal{H}}$ which allows to quantify the distance $d(\mathbf{a}, \mathbf{b})$ between two different elements $\mathbf{a}$ and $\mathbf{b}$. Hence, one can use this norm to measure the distance between two LTI models $\mathbf{H}_1$ and $\mathbf{H}_2$ by considering the norm of the difference of their transfer function, *i.e.*, $d(\mathbf{H}_1, \mathbf{H}_2) = \|\mathbf{H}_1 - \mathbf{H}_2\|_{\mathcal{H}}$. This notion is particularly interesting in the domain of model approximation since it provides a way to quantify the quality of a model approximation (see Part II).

In this section, we recall the definition and some properties of some Hardy spaces that are particularly important in this work. The theoretical elements related to Hardy spaces can also be found in the bibliographical references [Partington, 1997; Hoffman, 1962].

The $\mathcal{H}_2$ and $\mathcal{L}_2$ spaces

Firstly, let us define the $\mathcal{L}_2(i\mathbb{R})$ space.

Definition 2.1 ($\mathcal{L}_2(i\mathbb{R})$ **space).** *The $\mathcal{L}_2(i\mathbb{R})$ space is the vector-space of matrix-valued functions $\mathbf{F} : \mathbb{C} \to \mathbb{C}^{n_y \times n_u}$ which satisfy*

$$\int_{\mathbb{R}} \text{trace}[\overline{\mathbf{F}(i\omega)}\mathbf{F}(i\omega)^T]d\omega = \int_{\mathbb{R}} \|\mathbf{F}(i\omega)\|_F^2 d\omega < \infty.$$

This space is an Hilbert space equipped with the inner product defined as

$$\langle \boldsymbol{H}, \boldsymbol{G} \rangle_{\mathcal{L}_2(i\mathbb{R})} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{trace}\Big(\overline{\mathbf{H}(i\omega)}\mathbf{G}(i\omega)^T\Big)d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \langle \mathbf{H}(i\omega), \mathbf{G}(i\omega)\rangle_F d\omega, \qquad (2.9)$$

with corresponding induced-norm $\|\mathbf{H}\|_{\mathcal{L}_2(i\mathbb{R})} = \langle \mathbf{H}, \mathbf{H}\rangle^{\frac{1}{2}}_{\mathcal{L}_2(i\mathbb{R})}$.

Two subspaces of $\mathcal{L}_2(i\mathbb{R})$ will be important in this work. We now introduce the $\mathcal{H}_2$ space, which plays a very important role for model reduction. Its definition follows as :

Definition 2.2 ($\mathcal{H}_2$ **space).** *The Hardy space $\mathcal{H}_2(\mathbb{C}^+)$ is the space of the matrix-valued functions $\mathbf{F}(s) \in \mathbb{C}^{n_y \times n_u}$ such that*

(a) $\mathbf{F}(s)$ is analytic in the open right half-plane $\mathbb{C}^+$,

(b) for almost every real number ω,

$$\lim_{\sigma \to 0^+} \mathbf{F}(\sigma + i\omega) = \mathbf{F}(i\omega),$$

(c) $\sup_{\sigma \geq 0} \int_{-\infty}^{+\infty} \text{trace}[\overline{\mathbf{F}((\sigma + i\omega))}\mathbf{F}((\sigma + i\omega))^T]d\omega < \infty.$

This space is an Hilbert space equipped with the inner product defined as

$$\langle \boldsymbol{H}, \boldsymbol{G} \rangle_{\mathcal{H}_2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{trace}\Big(\overline{\mathbf{H}(i\omega)}\mathbf{G}(i\omega)^T\Big)d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \langle \mathbf{H}(i\omega), \mathbf{G}(i\omega)\rangle_F {}^{a} d\omega. \qquad (2.10)$$

for $\mathbf{H}, \mathbf{G} \in \mathcal{H}_2(\mathbb{C}^+)$ and its induced norm is defined as $\|\mathbf{H}\|_{\mathcal{H}_2} = \langle \mathbf{H}, \mathbf{H}\rangle^{\frac{1}{2}}_{\mathcal{H}_2}$.

[a] If $(a_{i,j}) = A \in \mathbb{C}^{n_y \times n_u}$ is a complex matrix, its Frobenious norm is defined as $\|A\|_F^2 = \text{trace}[\overline{A}A^T] = \sum_{i,j} |a_{i,j}|^2$. The Frobenious norm is induced by the inner product $\langle A, B\rangle_F = \text{trace}[\overline{A}B^T]$.

The $\mathcal{H}_2$ space is the vector-space of the transfer functions whose impulse responses $\mathbf{h}(t)$ are stable[1] and have finite energy.[2]

Let us denote by $\mathcal{H}_2(\mathbb{C}^-)$ the left half-plane analog to $\mathcal{H}_2(\mathbb{C}^+)$, *e.g.*, $\mathbf{G}(s) \in \mathcal{H}_2(\mathbb{C}^-)$ if and only if $\mathbf{G}(-s) \in \mathcal{H}_2(\mathbb{C}^+)$. This space is the vector-space of the LTI models $\mathbf{H}(s)$ whose all

[1] It depends, of course, on the definition of stability. Here, in general, we consider that if a system is $\mathcal{H}_2$, we consider it to be stable.

[2] Moreover, if $\mathbf{h}(t) \in L_2^n[0,\infty) \cap \mathcal{L}_1^n[0,\infty)$, the Riemann-Lebesgue lemma guaranties that the associated transfer function $\mathbf{H}(s)$ satisfies $\lim_{\omega \to \pm\infty} \mathbf{H}(i\omega) = 0$ and thus, the model is strictly proper.

the poles are in $\mathbb{C}^-$, *i.e.*, are all unstable. The space $\mathcal{H}_2(\mathbb{C}^-)$ corresponds to the space of the *anti-stable models*.

The space $\mathcal{L}_2(i\mathbb{R})$ is the vector-space of models whose transfer functions are square-integrable over the imaginary axis. In addition, $\mathcal{L}_2(i\mathbb{R})$ can be understood as the Laplace transform image of the signal space $L_2^n(-\infty,\infty)$, *i.e.*, $\mathscr{L}(L_2^n(-\infty,\infty)) = \mathcal{L}_2(i\mathbb{R})$. Moreover, the following proposition holds:

Proposition 2.3 ($\mathcal{L}_2(i\mathbb{R})$ **decomposition).** *The Laplace transform gives the following bijections between the signal spaces (defined in Section 2.1) and the Hardy spaces:*

$$\mathscr{L} : L_2^n[0,\infty) \to \mathcal{H}_2(\mathbb{C}^+),$$

and

$$\mathscr{L} : L_2^n(-\infty,0] \to \mathcal{H}_2(\mathbb{C}^-).$$

Moreover, $\mathcal{H}_2(\mathbb{C}^-)$ and $\mathcal{H}_2(\mathbb{C}^+)$ are closed subspaces of $\mathcal{L}_2(i\mathbb{R})$ and

$$\mathcal{L}_2(i\mathbb{R}) = \mathcal{H}_2(\mathbb{C}^-) \bigoplus \mathcal{H}_2(\mathbb{C}^+),$$

i.e., *$\mathcal{L}_2(i\mathbb{R})$ is the direct sum of $\mathcal{H}_2(\mathbb{C}^-)$ and $\mathcal{H}_2(\mathbb{C}^+)$.*

Proof. See *e.g.* [Partington, 1997, Chapter 1]. □

Proposition 2.3 shows that every element $\mathbf{H} \in \mathcal{H}_2(\mathbb{C}^+)$ (respectively $\mathbf{G} \in \mathcal{H}_2(\mathbb{C}^-)$) can be uniquely associated to an element $\mathbf{h} \in L_2^n[0,\infty)$ (respectively $\mathbf{g} \in L_2^n(-\infty,0]$). In addition, the following functional analysis' theorem shows that the Laplace transform preserves the inner product, and hence, the orthogonality property.

Theorem 2.4 (Plancherel's theorem). *For $\mathbf{h_1}, \mathbf{h_2} \in L_2^n(-\infty,\infty)$, one has*

$$\langle \mathbf{H_1}, \mathbf{H_2} \rangle_{\mathcal{L}_2(i\mathbb{R})} = \langle \mathscr{L}(\mathbf{h_1}), \mathscr{L}(\mathbf{h_2}) \rangle_{\mathcal{L}_2(i\mathbb{R})} = \langle \mathbf{h_1}, \mathbf{h_2} \rangle_{L_2}.$$

Moreover, $\mathcal{H}_2(\mathbb{C}^-)$ is orthogonal to $\mathcal{H}_2(\mathbb{C}^+)$ with respect to the $\mathcal{L}_2(i\mathbb{R})$ inner product, i.e., *if $\mathbf{H}_s \in \mathcal{H}_2(\mathbb{C}^+)$ and $\mathbf{H}_a = \mathcal{H}_2(\mathbb{C}^-)$, then*

$$\langle \mathbf{H}_s, \mathbf{H}_a \rangle_{\mathcal{L}_2(i\mathbb{R})} = 0.$$

Proof. See [Partington, 1988, chapter 1]. □

In other words, the above Proposition 2.3 and Theorem 2.4 state that given a model $\mathbf{H} \in \mathcal{L}_2(i\mathbb{R})$, there is a stable model $\mathbf{H}_s \in \mathcal{H}_2(\mathbb{C}^+)$ and an anti-stable model $\mathbf{H}_a = \mathcal{H}_2(\mathbb{C}^-)$ such that $\mathbf{H} = \mathbf{H}_s + \mathbf{H}_a$ and $\langle \mathbf{H}_s, \mathbf{H}_a \rangle_{\mathcal{L}_2(i\mathbb{R})} = 0$. These orthogonality results will play an important role in the stability estimation in the Chapters 8 and 9.

The $\mathcal{H}_\infty$ space

The definition of the $\mathcal{H}_\infty$ space is given as follows :

Definition 2.5 ($\mathcal{H}_\infty$ **space).** *The $\mathcal{H}_\infty(i\mathbb{R})$ space is the vector-space of meromorphic matrix-valued functions $\mathbf{F} : \mathbb{C} \to \mathbb{C}^{n_y \times n_u}$ analytic on $\mathbb{C}^+$ for which $\sup_{Re(s)>0} \|\mathbf{F}(s)\|_2 < \infty$. This*

space is equipped with the norm

$$\|\mathbf{F}\|_{\mathcal{H}_\infty} := \max_{\omega\in\mathbb{R}} \sigma_{max}(\mathbf{F}(i\omega)).$$

This norm allows to measure the worst case scenario and it is essential for robust control methods. In addition, one should note that the computation of the $\mathcal{H}_\infty$ norm can be a very difficult task and it is usually achieved by an iterative bisection algorithm (see *e.g.* [Bruinsma and Steinbuch, 1990] for $\mathcal{H}_\infty$ norm computation of finite dimensional model). This is the reason this norm is not widely used in the large-scale setting.

Norm inequalities and stability

Norms of systems can be used to measure how "large" an output $\mathbf{y}(t)$ will be if an LTI model is subjected to an input $\mathbf{u}(t)$. In others words, they can be used to estimate of an upper bound of the output $\mathbf{y}(t)$, given an input $\mathbf{u}(t)$. Different norms lead to different inequalities. In this section we will present two of these bounds that are useful to motivate the model reduction problem.

Proposition 2.6 (Inequality with $\mathcal{H}_2$ norm). *Let $\mathbf{H} \in \mathcal{H}_2$ and $\mathbf{u} \in L_2$. Then*

$$\|\mathbf{y}\|_{L_\infty} := \sup_{t\geq 0} \|\mathbf{y}(t)\|_2 \leq \|\mathbf{H}\|_{\mathcal{H}_2} \|\mathbf{u}\|_{L_2}. \tag{2.11}$$

Proof. Firstly we denote $\mathbf{U}(s)$ and $\mathbf{Y}(s)$ the Laplace transforms of the input $\mathbf{u}$ and output $\mathbf{y}$, respectively. In addition, Equation (2.7) gives the frequency relationship between $\mathbf{U}(s)$, $\mathbf{Y}(s)$ and $\mathbf{H}(s)$. Hence,

$$\begin{aligned}
\|\mathbf{y}\|_{L_\infty} &:= \sup_{t\geq 0} \|\mathbf{y}(t)\|_2 = \sup_{t\geq 0} \left\| \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{Y}(i\omega) e^{i\omega t} d\omega \right\|_2 \\
&\leq \sup_{t\geq 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} \|\mathbf{Y}(i\omega)\|_2 \underbrace{|e^{i\omega t}|}_{=1} d\omega \\
&\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \|\mathbf{Y}(i\omega)\|_2 d\omega \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \|\mathbf{H}(i\omega)\mathbf{U}(i\omega)\|_2 d\omega \\
&\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \|\mathbf{H}(i\omega)\|_2 \|\mathbf{U}(i\omega)\|_2 d\omega \\
&\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \|\mathbf{H}(i\omega)\|_F \|\mathbf{U}(i\omega)\|_2 d\omega \\
&\underbrace{\leq}_{\text{Cauchy-Schwartz Inequality}} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \|\mathbf{H}(i\omega)\|_F^2 d\omega\right)^{1/2} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \|\mathbf{U}(i\omega)\|_2^2 d\omega\right)^{1/2} \\
&\underbrace{=}_{\text{Parseval's identity}} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \|\mathbf{H}(i\omega)\|_F^2 d\omega\right)^{1/2} \left(\int_0^{+\infty} \|\mathbf{u}(t)\|_2^2 dt\right)^{1/2} \\
&= \|\mathbf{H}\|_{\mathcal{H}_2} \|\mathbf{u}\|_{L_2}.
\end{aligned}$$

□

Hence, a model $\mathbf{H}$ possessing a "small" $\mathcal{H}_2$ norm will produces output signals whose peak amplitude is also "small". Proposition 2.6 is largely applied in the context of model reduction, as illustrated in Chapter 5.

In a similar way, the $\mathcal{H}_\infty$ norm provides the following inequality

Proposition 2.7 (Inequality with $\mathcal{H}_\infty$ norm). *Let $\mathbf{H} \in \mathcal{H}_\infty$ and $\mathbf{u} \in L_2$. Then*

$$\|\mathbf{y}\|_{L_2} \leq \|\mathbf{H}\|_{\mathcal{H}_\infty} \|\mathbf{u}\|_{L_2}. \tag{2.12}$$

Moreover, the $\mathcal{H}_\infty$ norm is the induced norm by the signal L_2 norm, i.e.*,*

$$\|\mathbf{H}\|_{\mathcal{H}_\infty} = \sup_{\mathbf{u} \in L_2} \frac{\|\mathbf{y}\|_{L_2}}{\|\mathbf{u}\|_{L_2}},$$

and it is also called the L_2-L_2 induced norm.

Hence, a model $\mathbf{H}$ possessing "small" $\mathcal{H}_\infty$ norm will produces output signals with "small" energy. A system is considered to be stable a stable if small inputs lead to responses that do not diverge. Different input/output norms can lead to different notions of stability. In this work we considered the following two notions

- $\mathcal{H}_2$ **stability.** A system is $\mathcal{H}_2$ stable if it transfer function $\mathbf{H}(s)$ lies in $\mathcal{H}_2$. In this case the inequality (2.11) holds, and L_2 bounded inputs produces L_∞ bounded outputs.
- $\mathcal{H}_\infty$ **stability.** A system is $\mathcal{H}_\infty$ stable if it transfer function $\mathbf{H}(s)$ lies in $\mathcal{H}_\infty$. In this case, the inequality (2.12) holds, and L_2 bounded inputs produces L_2 bounded outputs.

In this thesis, we are interested in methods based on the $\mathcal{H}_2$ norm and that is why the notion of $\mathcal{H}_2$ stability will be used henceforth. In this work, a system refereed to be stable when it is $\mathcal{H}_2$ stable. It should be noticed that a transfer function $\mathcal{H}_2$ stable, analytic over the imaginary axis and strictly proper is automatically $\mathcal{H}_\infty$ stable.

Now, let us recall the definition and some representation properties of finite dimensional models.

2.2 Finite dimensional models

In this section, some elements about the representation of LTI finite dimensional models are recalled. The subjects that are highlighted are: (i) the partial fraction representation and (ii) the infinite gramians.

2.2.1 Descriptor realisation

Time-domain representation

A LTI dynamical model $\mathbf{H}$ is considered to be *finite dimensional model* if it can be represented in matrix terms by the following equations :

$$\mathbf{H} := \left\{ \begin{array}{rcl} E\dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = & C\mathbf{x}(t) + D\mathbf{u}(t) \end{array} \right. , \tag{2.13}$$

where $E, A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times n_u}$, $C \in \mathbb{R}^{n \times n_y}$ and $D \in \mathbb{R}^{n_y \times n_u}$ are the matrices defining the model. Also, $\mathbf{x}(t) \in \mathbb{R}^n$, $\mathbf{u}(t) \in \mathbb{R}^{n_u}$ and $\mathbf{y}(t) \in \mathbb{R}^{n_y}$ denote the state, input and output vectors,

respectively. The dimension n is the order of the model $\mathbf{H}$ and we denote order$(\mathbf{H}) = n$. One another common short notation for $\mathbf{H}$, given by (2.13), is

$$\mathbf{H} := (E, A, B, C, D) \quad \text{or} \quad \mathbf{H} := \left(\begin{array}{c|c} E, A & B \\ \hline C & D \end{array}\right). \tag{2.14}$$

For convenience, we shall suppose *zero initial conditions*, *i.e.*, that $x(0) = 0$. See [Dai, 1989; Kunkel and Mehrmann, 2006; Benner and Sokolov, 2006] for complete explanations about differential algebraic equations.

If $E = I_n$, identity matrix of order n, we call (2.13) a *standard state space realization*, otherwise a *generalized state space model*. Models with $\det(E) = 0$ are referred to as Differential-Algebraic Equations (DAE), but play a minor role in this thesis. Unless specified otherwise, we will assume E to be non-singular. In the case where E is assumed to be non-singular, the model $\mathbf{H}$ can be written in a standard state space realization. Even though this would always allow to get rid of the E matrix by multiplying the state equation of (2.13) with its inverse from the left, this should usually be avoided in a large scale-setting due to numerical issues and the (2.13) structure should be kept.

The impulse response of the LTI model is the Green function associated to the model, *i.e.*, the output response when the model is subjected by a Dirac input. In the case where the model $\mathbf{H}$ is governed by (2.13), the impulse response is given by

$$\mathbf{h}(t) = Ce^{E^{-1}At}E^{-1}B + D\boldsymbol{\delta}(t) \quad \text{for } t \geq 0, \tag{2.15}$$

and the model (2.13) can be represented by the convolution operator, *i.e.*,

$$y(t) = \int_0^t \mathbf{h}(t-\tau)\mathbf{u}(\tau)d\tau$$

as in (2.5).

Transfer function of a finite dimensional model

The transfer function $\mathbf{H}(s) = \mathcal{L}(\mathbf{h}(t))$ associated with $\mathbf{H} := (E, A, B, C, D)$ is given by

$$\mathbf{H}(s) = C(sE - A)^{-1}B + D \in \mathbb{C}^{n_y \times n_u}, \tag{2.16}$$

and represents the model in the frequency-domain. Since $\det(sE - A)$ is a polynomial in s, the transfer function of a finite dimensional model is a matrix-valued function with rational entries. The poles of the transfer function $\mathbf{H}(s)$ are the generalized eigenvalues of the pencil (E, A), denoted as $\Lambda(E, A)$ and defined as follows:

Definition 2.8 (Generalized eigenvalues). *The complex number $\lambda \in \mathbb{C}$ is an* eigenvalue *of the matrix pencil (E, A) if there exists a non-trivial vector $\mathbf{x} \in \mathbb{C}^n$ such that*

$$A\mathbf{x} = \lambda E\mathbf{x}.$$

In addition, the vector $\mathbf{x}$ is the right eigenvector *associated with the eigenvalue λ. Moreover if there exists a basis $X = \begin{bmatrix}\mathbf{x}_1 & \dots & \mathbf{x}_n\end{bmatrix}$ of eigenvectors of (E, A), we say that the pencil (E, A) is diagonalizable and*

$$AX = EX\Sigma,$$

where $\Sigma = \boldsymbol{diag}(\lambda_1, \dots, \lambda_n)$ and, for $k = 1, 2, \dots, n$, λ_k is the eigenvalue associated with the eigenvector $\mathbf{x}_k$. We also denote $\Lambda(E, A) = \{\lambda_1, \dots, \lambda_n\}$ as the generalized eigenvalues

of the pencil (E, A).

In addition, for a real model, if λ is a complex pole of $\mathbf{H}(s)$, $\overline{\lambda}$ is also a pole of $\mathbf{H}(s)$. Moreover, if Φ_k denotes the kth residue of $\mathbf{H}(s)$ with respect to λ_k, we have :

$$\Phi_k = \lim_{s \to \lambda_k} (s - \lambda_k)\mathbf{H}(s),$$

then Φ_k^* is the residue of $\mathbf{H}(s)$ with respect to $\overline{\lambda}$.

If E is non-singular, the poles are also the eigenvalues of the matrix $E^{-1}A$. Depending on the multiplicity of the poles, $\mathbf{H}(s)$ can be represented as the following poles-residues decomposition (see [Vuillemin, 2014; Vuillemin et al., 2014b]). The following proposition provides the partial fraction decomposition when the pencil (E, A) is diagonalizable.

Proposition 2.9 (**Partial fraction decomposition, semi-simple poles**). *Let* $\mathbf{H} = (E, A, B, C, D)$*, with* E *non-singular. In addition, suppose that the pencil* (E, A) *is* diagonalizable*, i.e., there is non-singular matrix* $X \in \mathbb{C}^{n \times n}$ *such that* $AX = EX\Sigma$*,* $\Sigma = \boldsymbol{diag}(\lambda_1, \ldots, \lambda_n)$*. Then*

$$\mathbf{H}(s) = \sum_{k=1}^{n} \frac{\Phi_k}{s - \lambda_k} + D, \tag{2.17}$$

where

$$\Phi_k = CX\mathbf{e}_k\mathbf{e}_k^T(EX)^{-1}B,$$

where $\mathbf{e}_k \in \mathbb{R}^n$ *is the vector whose all entries are zero, with exception the* k*-th entry which is equal 1.*

Reader should note that if λ_k is a simple pole, its associated residue Φ_k can be obtained by

$$\Phi_k = \lim_{s \to \lambda_k} (s - \lambda_k)H(s).$$

The pencil (E, A) is said to be *defective* if it is not diagonalisible. In this case, one would need superior order partial fraction decomposition, as stated in the following proposition:

Proposition 2.10 (**Partial fraction decomposition, higher multiplicity pole**). *Let* $\mathbf{H} = (E, A, B, C, D)$*, with* E *non-singular. In addition, let us suppose that the pencil* (E, A) *has only one single eigenvalue* $\lambda \in \mathbb{C}$ *and that there exists a non-singular transformation* X *such that* $AX = EXJ$*, where* J *is the following Jordan block :*

$$J = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \ddots \\ & & \ddots & \\ & & & \lambda \end{bmatrix}, \quad \lambda \in \mathbb{C}.$$

Then

$$\mathbf{H}(s) = \sum_{k=1}^{n} \frac{\Phi_k}{(s - \lambda)^k} + D,$$

where the residues $\Phi_k \in \mathbb{C}^{n_y \times n_u}$ *can be computed by*

$$\Phi_k = CF_kE^{-1}B$$

where

$$F_1 = I_n, \quad F_2 = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}, \ldots, \quad F_n = \begin{bmatrix} 0 & 0 & & 1 \\ & 0 & \ddots & \\ & & \ddots & 0 \\ & & & 0 \end{bmatrix}.$$

Proposition 2.10 assumes that the pencil (E, A) has only one eigenvalue. In the general case, the pencil (E, A) can be written in its Jordan's form J, *i.e.*, there is a non-singular matrix $V \in \mathbb{C}^{n\times n}$ such that

$$AV = EVJ, \quad \text{where} \quad J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_{n_b} \end{bmatrix} \in \mathbb{C}^{n\times n},$$

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

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

Then, the transfer function $\mathbf{H}(s) = C(Es - A)^{-1}B + D$ can be written as

$$\mathbf{H}(s) = \left(\sum_{i=1}^{n_u} \mathbf{H}_{\lambda_i}(s) \right) + D,$$

where

$$\mathbf{H}_{\lambda_i}(s) = C_i(sI - J_i)^{-1}E^{-1}B_i,$$

and $C_i \in \mathbb{C}^{n_y\times n_i}$ denotes (respectively $B_i \in \mathbb{C}^{n_i\times n_u}$) the columns of CV (respectively the lines) associated to the block J_i. Hence, to obtain the partial fraction decomposition of $\mathbf{H}(s)$ one should apply Proposition 2.10 to each term $\mathbf{H}_{\lambda_i}(s)$. Let us illustrate this results by the following example:

Example 2.11 (Example of partial fraction decomposition). *Let us consider the SISO model* $\mathbf{H}$ *whose realization is given by the following matrices :*

$$E = I_3, \quad A = \begin{bmatrix} -5 & 3 & 1 \\ 2 & -10 & -2 \\ -11 & 5 & -1 \end{bmatrix}, B = C^T = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, D = 0.$$

First, A can be expressed in its Jordan's form as follows:

$$AV = VJ, \quad \text{where} \quad J = \left[\begin{array}{c|cc} 8 & 0 & 0 \\ \hline 0 & 4 & 1 \\ 0 & 0 & 4 \end{array}\right] \quad \text{and} \quad V = \begin{bmatrix} -1/4 & 2 & 5/4 \\ 1/2 & -4 & -1/2 \\ -3/4 & 14 & 3/4 \end{bmatrix}.$$

Hence,

$$\begin{aligned}\mathbf{H}(s) &= C(Es-A)^{-1}B = CV(sI_3-J)^{-1}V^{-1}B \\ &= C_1(s-8)^{-1}B_1 + C_2(sI_2 - \begin{bmatrix}4 & 1\\ 0 & 4\end{bmatrix})^{-1}B_2 \\ &= \tfrac{\Phi}{s-8} + \tfrac{\Psi_1}{s-4} + \tfrac{\Psi_2}{(s-4)^2}\end{aligned}$$

where

$$\Phi = C_1B_1 = -3, \quad \Psi_1 = C_2I_2B_2 = 6, \text{ and } \quad C_2\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}B_2 = 18.$$

and

$$C_1 = -\frac{1}{2}, B_1 = 6, B_2 = \begin{bmatrix}5/16\\ 3/2\end{bmatrix} \quad \text{and} \quad C_2 = \begin{bmatrix}12 & 3/2\end{bmatrix}.$$

These partial fraction decomposition will plays a very important role in the spectral $\mathcal{H}_2$ inner product computation in Chapter 4. We remark that if the matrix E is singular, then the partial fraction representation of $\mathbf{H}(s)$ may have a polynomial part, implying that the model is no longer in $\mathcal{H}_2$ (see [Benner and Sokolov, 2006] for some applications in model approximation). That is why we want to avoid this case and we suppose that $\det(E) \neq 0$. Note that in the Loewner framework (see Chapter 3) this particular case is handled. Now, let us move on towards the gramians and $\mathcal{H}_2$ norm computation.

2.2.2 Gramians and $\mathcal{H}_2$ norm computation

Controllabiliy and observability gramians

The controllability and observability gramians $\mathcal{P}$ and $\mathcal{Q}$ are symmetric semi-positive matrices that are used to determine whether or not a stable and finite dimensional LTI model is controllable or observable. In general, a finite dimensional LTI model is controllable and observable and only if and only if the matrices $\mathcal{P}$ and $\mathcal{Q}$ are non-singular. In addition, their associated bilinear form is related with the degree of controllability and observability of a given state $\overline{\mathbf{x}} \in \mathbb{R}^n$. For a descriptor model $\mathbf{H}$ as in (2.13), the gramians can be defined as follows :

Definition 2.12. *Given an asymptotically stable model $\mathbf{H} = (E, A, B, C, D)$, the infinite controllability and observability gramians associated with $\mathbf{H}$, denoted by $\mathcal{P}$ and $\mathcal{Q}$ respectively, are defined as*

$$\mathcal{P} = \int_0^\infty e^{E^{-1}At}E^{-1}BB^TE^{-T}e^{A^TE^{-T}t}dt \tag{2.18}$$

$$\mathcal{Q} = \int_0^\infty e^{A^TE^{-T}t}C^TCe^{E^{-1}At}dt \tag{2.19}$$

in the time domain, and as

$$\mathcal{P} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(\omega iE - A)^{-1}BB^T(-\omega iE^T - A^T)^{-1}d\omega \tag{2.20}$$

$$\mathcal{Q} = \frac{1}{2\pi}E^T\int_{-\infty}^{\infty}(-\omega iE^T - A^T)^{-1}C^TC(\omega iE - A)^{-1}d\omega E \tag{2.21}$$

in the frequency domain.

The controllabillity and observability gramians can be also computed as the solution of the following Lyapunov equations :

$$\begin{aligned} A\mathcal{P}E^T + E\mathcal{P}A^T + BB^T &= 0 \quad (2.22)\\ A^T\mathcal{Q}E + E^T\mathcal{Q}A + C^TC &= 0. \quad (2.23)\end{aligned}$$

Now let us exploit the grammians to compute the $\mathcal{H}_2$ norm of a dynamical system.

Gramian-based $\mathcal{H}_2$ norm computation

The $\mathcal{H}_2$ norm of a finite dimensional model can be expressed by the infinite gramians of the model as presented in Theorem 2.13.

Theorem 2.13 (Gramian formulation of the $\mathcal{H}_2$ norm). *Let $\mathbf{H} = (E, A, B, C)$ be a finite dimensional model, $\mathbf{H} \in \mathcal{H}_2$, with infinite observability and controllability gramians are $\mathcal{Q}$ and $\mathcal{P}$, respectively. The $\mathcal{H}_2$ norm of $\mathbf{H}$, is given by*

$$\|\mathbf{H}\|_{\mathcal{H}_2} = \operatorname{trace}(C\mathcal{P}C^T) = \operatorname{trace}(B^T\mathcal{Q}B). \quad (2.24)$$

Proof. See *e.g.* [Gugercin et al., 2008]. □

Theorem 2.13 provides a way to compute the $\mathcal{H}_2$ norm of a finite dimensional model. Hence, given $\mathbf{H} = (E, A, B, C, D)$, one should solve the Lyapunov equations (2.22) and apply the formula (2.24) to obtain the $\mathcal{H}_2$ norm of $\mathbf{H}$.

$\mathcal{H}_2$ inner product computation based on Sylvester equations

Analogously, $\mathcal{H}_2$ inner product between two finite dimensional models can be expressed by the Sylvester equations as presented in Theorem 2.14.

Theorem 2.14 (Cross Gramians formulation of the $\mathcal{H}_2$ inner product). *Let $\mathbf{H} = (E, A, B, C)$ and $\mathbf{G} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ be two finite dimensional models, and assume $\mathbf{H}, \mathbf{G} \in \mathcal{H}_2$. Then, the $\mathcal{H}_2$ inner product is given by*

$$\langle \mathbf{H}, \mathbf{G}\rangle_{\mathcal{H}_2} = \operatorname{trace}(B^T\mathcal{Y}\hat{B}) = \operatorname{trace}(C\mathcal{X}\hat{C}^T) = \operatorname{trace}(C^T\mathcal{Z}\hat{B}), \quad (2.25)$$

where the matrices $\mathcal{X}, \mathcal{Y}$ and $\mathcal{Z}$ satisfy the Sylvester equations

$$\begin{aligned} A\mathcal{X}\hat{E}^T + E\mathcal{X}\hat{A}^T + B\hat{B}^T &= 0\\ A^T\mathcal{Y}\hat{E} + E^T\mathcal{Y}\hat{A} + C^T\hat{C} &= 0 \quad . \quad (2.26)\\ A\mathcal{Z}\hat{E}^T + E^T\mathcal{Y}\hat{A} + B\hat{C}^T &= 0\end{aligned}$$

Proof. See [Mehrmann and Stykel, 2005] □

Theorems 2.13 and 2.14 characterize the $\mathcal{H}_2$ inner product and norm by Sylvester and Lyapunov equations in the finite dimensional case. For more details about the numerical procedures to efficiently compute them in the large-scale setting, reader is invited to refer to [Antoulas, 2005; Sorensen and Antoulas, 2002; Benner et al., 2009, 2006]. In Chapter 4 an alternative characterization based on the spectral decomposition of a dynamical system will be provided.

2.3 A glimpse of LTI time-delay systems

Hereafter, some elements about the theory of linear time-delay systems are recalled. The subjects that are focused on: (i) input and output delays and (ii) state-delay systems.

Delay-differential equations

A great amount of real systems present delays in their inner structure, due to transport, propagation or communication phenomena. In practice, most of the time they are ignored for the sake of simplicity. On the other hand, those delays can be the cause of bad performance or even instability, and therefore, in order to analyze properly and design controllers for those systems, it is mandatory to take their effects into account. Another important source of delay is the feedback control itself, with this delay induced by the sensors, actuators and, in more modern digital controllers, the time of calculation.

Time-delays systems are an important class of infinite dimensional models extensively studied (see [Richard, 2003; Niculescu, 2001a; Gu et al., 2003]). Such time-delay representations occur often in systems in engineering, biology, chemistry, physics, and ecology [Niculescu, 2001b]. Time-delay systems can be represented by Delay Differential Equations (DDEs), which belong to the class of functional differential equations, and have been extensively studied over the past decades. See [Bellman and Cooke, 1963; Gu et al., 2003; Briat, 2015; Richard, 2003] for a bibliographical references on DDEs. In this work, the following two kinds of time-delay structure are being considered :

2.3.1 Input and output time-delay models

A dynamical system presents an input delays when the action of the control input takes a certain amount of time before it affects systems dynamics. It presents an output delay when the measurement process takes a certain amount of time to produce a measurement. Those kinds of behavior appear frequently in industrial processes as well as in economical and biological systems. These can be the delays in the process itself and delays caused by processing sensed signals, transport phenomena, among others. See [Balluchi et al., 2006] for an example of modeling and control of a rail injection system using an input delay model. For this kind of systems, a great variety of control design techniques can be found in the literature, *e.g.*, the well-know Smith predictor (see [Smith, 1957; Alevisakis and Seborg, 1973; Watanabe and Ito, 1981]).

In the case where the system presents only a single input delay, it is also referred in the literature of control systems as a *dead-time systems*. Reader might refer to [Dym et al., 1995; Meinsma and Zwart, 2000; Foias et al., 1996; Tadmor, 1997] for this kind of systems. The input and output delay models present in this work can be defined as follows.

Time-domain representation of an input and output delay model

First, let us consider the *shift time input operator*, denoted by $\mathbf{\Delta_i}$, and the *shift time output operator*, denoted by $\mathbf{\Delta_o}$, defined respectively as follows:

$$\begin{aligned} \mathbf{\Delta_i}: \quad \mathcal{S}_{n_u} &\mapsto \mathcal{S}_{n_u} \\ \Big(\mathbf{u}(t)\Big)_i &\mapsto \Big(\mathbf{\Delta_i}(\mathbf{u})(t)\Big)_i = \mathbf{u}_i(t-\tau_i), \quad \text{for} \quad i = 1, \dots, n_u. \end{aligned} \tag{2.27}$$

$$\begin{aligned} \mathbf{\Delta_o}: \quad \mathcal{S}_{n_y} &\mapsto \mathcal{S}_{n_y} \\ \Big(\mathbf{y}(t)\Big)_j &\mapsto \Big(\mathbf{\Delta_o}(\mathbf{y})(t)\Big)_j = \mathbf{y}_j(t-\gamma_j), \quad \text{for} \quad j = 1, \dots, n_y. \end{aligned} \tag{2.28}$$

for some *input delays* $\{\tau_1, \ldots, \tau_{n_u}\} \in \mathbb{R}_+$, and *output delays* $\{\gamma_1, \ldots, \gamma_{n_u}\} \in \mathbb{R}_+$.

Then, we say that $\mathbf{H}_d$ is a *multiple-input/output delays MIMO system* if it is represented by:

$$\mathbf{H}_d : \left\{ \begin{array}{rcl} E\dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + B\boldsymbol{\Delta}_i(\mathbf{u}(t)) \\ \mathbf{y}(t) & = & \boldsymbol{\Delta}_o(C\mathbf{x}(t)) \end{array} \right. , \tag{2.29}$$

where $E, A \in \mathbb{R}^{n\times n}$ (with state dimension $n \in \mathbb{N}^*$), $B \in \mathbb{R}^{n\times n_u}$, $C \in \mathbb{R}^{n_y\times n}$.

Let $\mathbf{H}$ be the delay-free version of $\mathbf{H}_d$, *i.e.*, the model $\mathbf{H}$ is represented by

$$\mathbf{H} : \left\{ \begin{array}{rcl} E\dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = & C\mathbf{x}(t) \end{array} \right. \tag{2.30}$$

In addition, suppose that $\mathbf{h}(t) \in \mathbb{R}^{n_y\times n_u}$ is the impulse response of the delay-free model $\mathbf{H}$. Then, the impulse response of the input and output delay model $\mathbf{H}_d$, denoted here by $\mathbf{h}_d(t) \in \mathbb{R}^{n_y\times n_u}$ is given by:

$$\mathbf{h}_d(t) = \begin{bmatrix} \mathbf{h}_{1,1}(t-\gamma_1-\tau_1) & \ldots & \mathbf{h}_{1,n_u}(t-\gamma_1-\tau_{n_u}) \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{n_y,1}(t-\gamma_{n_y}-\tau_1) & \ldots & \mathbf{h}_{n_y,n_u}(t-\gamma_{n_y}-\tau_{n_u}) \end{bmatrix} \in \mathbb{R}^{n_y\times n_y}. \tag{2.31}$$

Transfer function of an input and output delay model

Recall that the Laplace transform satisfies the following shift property: if $\mathcal{L}(\mathbf{h}_{ij}(t)) = \mathbf{H}_{ij}(s)$, then $\mathcal{L}(\mathbf{h}(t-\gamma_i-\tau_j)) = e^{-s\gamma_i}\mathbf{H}_{ij}(s)e^{-s\tau_i}$. Hence, the transfer function associated to the input and output delay model $\mathbf{H}_d$ is given by

$$\begin{array}{rcl} \mathbf{H}_d(s) & = & \mathcal{L}(\mathbf{h}_d(t)) \\ & = & \boldsymbol{\Delta}_o(s)\mathcal{L}(\mathbf{h}(t))\boldsymbol{\Delta}_i(s) \\ & = & \boldsymbol{\Delta}_o(s)\mathbf{H}(s)\boldsymbol{\Delta}_i(s) \end{array}$$

where

$$\left\{ \begin{array}{rcl} \boldsymbol{\Delta}_i(s) & = & \mathbf{diag}(e^{-s\tau_1} \ \ldots \ e^{-s\tau_{n_u}}) \in \mathcal{H}_\infty^{n_u\times n_u} \\ \boldsymbol{\Delta}_o(s) & = & \mathbf{diag}(e^{-s\gamma_1} \ \ldots \ e^{-s\gamma_{n_y}}) \in \mathcal{H}_\infty^{n_y\times n_y}. \end{array} \right. \tag{2.32}$$

are the matrix transfer functions which $\boldsymbol{\Delta}_i(s)$ and $\boldsymbol{\Delta}_o(s)$ represent the frequency behavior of the delays operators $\boldsymbol{\Delta}_i$ and $\boldsymbol{\Delta}_o$, respectively, and

$$\mathbf{H}(s) = C(sE - A)^{-1}B \in \mathbb{C}^{n_y\times n_u} \tag{2.33}$$

is the transfer function of the delay-free version $\mathbf{H}$. The order of $\mathbf{H}_d$ is defined as the order of the underlying delay-free model $\mathbf{H}$. In addition, if $\mathbf{H}$ lies in $\mathcal{H}_2$, then $\mathbf{H}_d$ also lies in $\mathcal{H}_2$.

Reader should note that a *multiple-input/output delays MIMO system* $\mathbf{H}_d$ as in (2.30) is an infinite dimensional model even if it has a finite number of poles, since its transfer function is not rational. However, because they have a simple transfer function, they are well adapted for analysis and controller design. Input and output reduced order models will be considered in Chapter 5.

2.3.2 State-delay systems

The models considered here present one or many state-delays. They are represented by functional differential equations of retarded type, *i.e.*, by a retarded delay differential equation of the form

$$\dot{x} = F(x(t), x(t-\tau_1), \ldots, x(t-\tau_{max})), \tag{2.34}$$

where $\tau_1 < \tau_2 < \dots \tau_{max} \in \mathbb{R}_+$, equipped with an initial condition (also called *history function*)

$$x(t) = \phi(t), \quad \text{for} \quad t \in (-\tau_{max}, 0).$$

The reader should refer to [Bellman and Cooke, 1963] for more details. There exist as well the class of neutral time-delay systems, where the function F in (2.34) depends also on $\dot{x}(t-\tau_k)$, and the class of advanced time-delay systems, where the function F in (2.34) depends only on $\dot{x}(t-\tau_k)$. Those systems are not considered in this thesis and the reader should refer to [Bellman and Cooke, 1963; Michiels and Niculescu, 2014] for more details. In addition, there exists the class of distributed retarded time-delay systems where the delays are not discrete. An example of this is given by the following distributed delay differential equation

$$\dot{x} = \int_{-\tau}^{0} x(t+\theta) f(\theta) d\theta.$$

This class of systems is not consider in this work and reader should refer to [Richard, 2003]. In what follows, we present the class of linear discrete state-delays models.

Time-domain representation of a linear discrete state-delay models

The following general class of *LTI discrete retarded time-delay systems* will also be considered in this work:

$$\mathbf{H}_d = \left\{ \begin{array}{rl} E\dot{\mathbf{x}}(t) & = A_0\mathbf{x}(t) + \sum_{k=1}^{n_d} A_k\mathbf{x}(t-\tau_k) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = C\mathbf{x}(t) \end{array} \right. , \tag{2.35}$$

where $E, A_k \in \mathbb{R}^{n\times n}$, for $k = 1, \dots, n_d$, $B \in \mathbb{R}^{n\times n_u}$, $C \in \mathbb{R}^{n_y \times n}$, $\tau_1 < \tau_2 < \dots < \tau_d \in \mathbb{R}_+^*$, $\mathbf{x}(t) \in \mathbb{R}^n$, $\mathbf{u}(t) \in \mathbb{R}^{n_u}$, $\mathbf{y}(t) \in \mathbb{R}^{n_y}$ are the delays, the state of the system, the input and the output, respectively. In this case, we say that the system $\mathbf{H}_d$ has order n, because it is defined by n DDEs. We suppose also that E is a non-singular matrix, otherwise this family might include neutral systems.

Transfer function of a discrete state-delay model

Considering $\phi(t) = 0$, for $-\tau_d < t < 0$ taking the Laplace transformation of equation (2.35), one is able to find that the transfer function of $\mathbf{H}_d$ is given by

$$\mathbf{H}_d(s) = C\left(Es - A_0 - \sum_{k=1}^{d} A_k e^{-s\tau_k}\right)^{-1} B. \tag{2.36}$$

The transfer function of $\mathbf{H}_d$ is a matrix-valued function with meromorphic $(\mathbf{H}_d)_{ij}(s)$ entries expressed as

$$(\mathbf{H}_d(s))_{ij} = \frac{n_{ij}(s, e^{-s\tau_1}, \dots, e^{-s\tau_d})}{d_{ij}(s, e^{-s\tau_1}, \dots, e^{-s\tau_d})},$$

where $n(s, e^{-s\tau_1}, \dots, e^{-s\tau_d})$ and $d(s, e^{-s\tau_1}, \dots, e^{-s\tau_d})$ are polynomials on the variables s, $e^{-s\tau_1}$, $\dots, e^{-s\tau_d}$. This is reflected in the characteristic equation, *i.e.*,

$$p(s) = \det(Es - A_0 - \sum_{k=1}^{d} A_k e^{-s\tau_k}),$$

which is a quasi-polynomial whose zeros are the poles of $\mathbf{H}_d$. Since E is supposed to be , $p(s)$ does not depend on $s^n e^{-s\tau_k}$, otherwise the model would be neutral. Those systems present nice properties and some of them are listed hereafter (see [Michiels and Niculescu, 2014]) :

- ➤ the poles of $\mathbf{H}_d$ are roots of an analytic function.
- ➤ there are only finitely many roots in any right-half plane.
- ➤ the system $\mathbf{H}_d$ is strictly proper. Hence, it is ($\mathcal{H}_2$ and $\mathcal{H}_\infty$) stable if and if all its poles lie in the left-half plane.
- ➤ the spectral abcissa is continuous with respect to the parameters (see Chapter 8, Property 8.1 for more details).

These properties will be developed and used in Chapter 8. In addition, the transfer function of $\mathbf{H}_d$ is clearly irrational and that is why $\mathbf{H}_d$ is considered to be an infinite dimensional model. A special case of retarded time-delay system, namely single state-delay model, is considered in Chapters 6 and 7. Stability of time-delay systems is considered in Chapter 8 and 9.

Conclusions

In this chapter, some important results related to the LTI system theory were recalled in order to unify the necessary background for the development of this work. Firstly, Section 2.1 defines the notion of LTI models by convolution integral and by transfer functions living in a Hardy space. Secondly, Section 2.2 has recalled the definition of a finite dimensional LTI models pointing out the notion of partial fraction decomposition and gramians. Finally, Section 2.3 recalls some facts about time-delay models, highlighting two structures : input/output delays and state-delay models.

Chapter 3

State of the art of model approximation

This chapter aims at providing a (non-exhaustive) state-of-the-art on model approximation techniques. The main goal of this chapter is, on one hand, to provide to the non-familiar reader a brief introduction on model approximation, and, on the other hand, to present some methods that have inspired the results from this thesis. Moreover, some of the theoretical results and some mathematical proofs are provided for a better understanding.

The chapter is organized as follows. Firstly, a general overview and the projection-based model approximation framework are presented in Section 3.1. Secondly, Section 3.2 briefly recalls the notion of model approximation by truncation, especially the well-known *modal truncation* and *balanced truncation* techniques. Then, in Section 3.3, *model approximation by tangential interpolation* is presented. Finally, Section 3.4 presents the *Loewner framework*. These two last frameworks are the main basis of the work done in this thesis. Optimal $\mathcal{H}_2$ model approximation will be developed in Chapter 4.

Contents

3.1 Introduction

3.1.1 General overview

A great amount of physical phenomena can be modeled by LTI systems, which are usually derived from a linearisation at the operating point of interest. In this work, only LTI models are considered (see Chapter 2). In this chapter, some model approximation techniques are presented. Firstly, the projection-based model approximation framework is presented. A great variety of model approximation methods can be described in this framework. In the sequel, two projection-based model approximation methods are described: (i) model approximation by truncation in 3.2 and (ii) model approximation by tangential interpolation in 3.3. These techniques can be successfully applied to a large number of applications. The major drawback is that they require that both the full-order model $\mathbf{G}$ and the reduced order model $\hat{\mathbf{H}}$ to be represented by finite-dimensional realizations. Hence, they are not directly applicable to infinite dimensional systems or irrational transfer functions and they do not allow the reduced order model to have any particular structure. However, in Section 3.3, a preserving structure projection-based method is presented, which already enables model approximation for a larger family of models. Finally, Section 3.4 presents the Loewner framework. This framework performs data-based tangential interpolation and, in the context of model approximation, can be used to construct reduced order models directly from the evaluation of the original transfer function or for data provided by any simulator.

In what follows, let us consider $\mathbf{G}$ to be a large-scale finite dimensional model whose realization is given by

$$\mathbf{G} := \begin{cases} E\dot{\mathbf{x}}(t) &= A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) &= C\mathbf{x}(t) + D\mathbf{u}(t) \end{cases} \tag{3.1}$$

where $\mathbf{x}(t) \in \mathbb{R}^N$, $\mathbf{u}(t) \in \mathbb{R}^{n_u}$ and $\mathbf{y}(t) \in \mathbb{R}^{n_y}$ denote the state, input and output, respectively. The order N of $\mathbf{G}$ is considered to be large, *e.g*, $N = 10^2, \ldots, 10^7$ and the necessity to find a low-order approximation will depend on the application. The matrix E is assumed to be non-singular, which means that the model does not contain any algebraic constraint.

Model approximation seeks for a simpler model $\hat{\mathbf{H}}$ which has a similar behavior as $\mathbf{G}$. Simpler, in this context, means to be described by less equations, or, in other words, to have fewer state variables. Hence, if $\hat{\mathbf{H}}$ has a finite dimensional realization, it takes the form

$$\hat{\mathbf{H}} := \begin{cases} \hat{E}\dot{\hat{\mathbf{x}}}(t) &= \hat{A}\hat{\mathbf{x}}(t) + \hat{B}\mathbf{u}(t) \\ \hat{\mathbf{y}}(t) &= \hat{C}\hat{\mathbf{x}}(t) + \hat{D}\mathbf{u}(t) \end{cases} \tag{3.2}$$

where $\hat{E}, \hat{A} \in \mathbb{R}^{n\times n}$, $\hat{B} \in \mathbb{R}^{n\times n_u}$, $\hat{C} \in \mathbb{R}^{n_y\times n}$ and $\hat{D} \in \mathbb{R}^{n_y\times n_u}$, and the state variable $\hat{\mathbf{x}}(t) \in \mathbb{R}^n$ is a lower dimension vector, *i.e.*, $n \ll N$, the input $\mathbf{u}(t) \in \mathbb{R}^{n_u}$ and the output $\hat{\mathbf{y}}(t) \in \mathbb{R}^{n_y}$ have the same dimensions that $\mathbf{u}(t)$ and $\mathbf{y}(t)$, respectively. We recall that the transfer functions of $\mathbf{G}$ and $\hat{\mathbf{H}}$ are, respectively,

$$\mathbf{G}(s) = C(sE - A)^{-1}B + D \quad \text{and} \quad \hat{\mathbf{H}}(s) = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{C} + \hat{D}.$$

As model approximation is concerned to match the transfer behavior, a ROM usually features an equal feed-through term, *i.e.*, $D = \hat{D} \in \mathbb{R}^{n_y\times n_u}$. One should note that if $D \neq \hat{D}$, then $(\mathbf{G}-\hat{\mathbf{H}})(i\infty) = D-\hat{D} \neq 0$ and, hence, $\|\mathbf{G}-\hat{\mathbf{H}}\|_{\mathcal{H}_2} = \infty$. So, if one requires a good approximation in the sense of the norm $\mathcal{H}_2$, then we must have $D = \hat{D}$. However, the feed-through $\hat{D}$ term

may still be used as an additional degree of freedom in order to improve the approximation. For example, the work of [Flagg et al., 2013] presents a framework which uses the feed-through term $\hat{D}$ in order to improve the approximation in the $\mathcal{H}_\infty$ norm sense. In addition, the weighted (see [Breiten et al., 2015] and [Anić et al., 2013]) and the frequency limited (see [Vuillemin, 2014] and [Petersson and Löfberg, 2014]) model approximation might provide reduced order models having the feed-through $\hat{D}$ different than zero.

In this work, since we study $\mathcal{H}_2$-oriented approaches, we will consider $D = \hat{D} = 0$. In addition, we denote the systems $\mathbf{G}$ and $\hat{\mathbf{H}}$ having, respectively, the realizations from equations (3.1) and (3.2) by $\mathbf{G} = (E, A, B, C)$ and $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$.

In the next section, we present the *projection-based model approximation framework* (see [Antoulas, 2005, Chapter 1]). This framework has led to various model approximation methods for finite dimensional LTI models, *e.g.*, the balanced truncation and the model approximation by tangential interpolation methods. Both of these methods are reviewed in Sections 3.2 and 3.3.

3.1.2 Projection-based model approximation framework

In this section, the projection-based model approximation framework is introduced. Let us consider a large-scale model $\mathbf{G} = (E, A, B, C)$ described by (3.1). Let us assume that there exists a n-dimensional subspace, with $n \ll N$, that contains the most dominant dynamics. Suppose that this subspace is generated by the basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, where $\mathbf{v}_k \in \mathbb{C}^N$ and let $V = \begin{bmatrix}\mathbf{v}_1 \ldots \mathbf{v}_n\end{bmatrix} \in \mathbb{C}^{N\times n}$. Then one can find the approximation

$$\mathbf{x}(t) = V\hat{\mathbf{x}}(t) + \mathbf{e}(t) \tag{3.3}$$

to be admissible. In (3.3), $\mathbf{e}(t) \in \mathbb{C}^N$ is the resulting error which comes from the approximation of $\mathbf{x}(t)$. Then, by substituting (3.3) in the state equation of the original model, we get

$$EV\dot{\hat{\mathbf{x}}}(t) = AV\hat{\mathbf{x}}(t) + B\mathbf{u}(t) + \epsilon(t), \tag{3.4}$$

where the residual variable $\epsilon(t) \in \mathbb{C}^N$ contains the errors and satisfies $\epsilon(t) = A\mathbf{e}(t) - E\dot{\mathbf{e}}(t)$. It is worth to note that the differential equation (3.4) has N equations for n unknowns $\hat{\mathbf{x}}(t)$. To handle this problem, one can enforce the *Petrov-Galerkin conditions* by finding a subspace $\mathcal{W}$ which is orthogonal to $\epsilon(t)$. By denoting $W \in \mathbb{C}^{N\times n}$ be the matrix whose $span(W) = \mathcal{W}$, by multiplying (3.4) by W^T on the left, one obtains the reduced order model $\hat{\mathbf{H}}$, as follows:

$$\hat{\mathbf{H}} = \begin{cases} \underbrace{W^TEV}_{\hat{E}}\dot{\hat{\mathbf{x}}}(t) &= \underbrace{W^TAV}_{\hat{A}}\hat{\mathbf{x}}(t) + \underbrace{W^TB}_{\hat{B}}\mathbf{u}(t) + \underbrace{W^T\epsilon(t)}_{=0} \\ \hat{\mathbf{y}}(t) &= \underbrace{CV}_{\hat{C}}\hat{\mathbf{x}}(t) \end{cases} \tag{3.5}$$

Hence, the reduced order model $\hat{\mathbf{H}}$ obtained by the projection over the space $\mathcal{V}$ along the directions $\mathcal{W}^\perp$ is given by $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$, where

$$\hat{E} = W^TEV, \quad \hat{A} = W^TAV, \quad \hat{B} = W^TB \text{ and } \hat{C} = CV. \tag{3.6}$$

The following remark characterizes the implicit projector operator used in this procedure.

Remark 3.1 (Projector Π). *Assume that W^TEV is invertible, so that the projection is regular. Let $\Pi = EV(W^TEV)^{-1}W^T \in \mathbb{C}^{N\times N}$. Then*

$$\Pi = \Pi^2 \in \mathbb{C}^{N\times N}$$

is an oblique projector onto the n-th subspace spanned by the columns of span(EV) along the kernel of W^* (see [Saad, 2003]). In addition, model $\hat{\mathbf{H}}$ can be represented by the following realization

$$\hat{\mathbf{H}} = \begin{cases} \Pi EV\dot{\hat{\mathbf{x}}}(t) &= \Pi AV\hat{\mathbf{x}}(t) + \Pi B\mathbf{u}(t) \\ \hat{\mathbf{y}}(t) &= CV\hat{\mathbf{x}}(t) \end{cases}.$$

In this regard, matrices V and W are called projection matrices *because they are the fundamental blocks to generate the projector. In addition, if the subspaces $\mathcal{V}$ and $\mathcal{W}$ are equal, then one talks about* orthogonal projection*. Otherwise, one talks about oblique projection.*

To sum up, given a large-scale model $\mathbf{G} = (E, A, B, C)$ represented by equation (3.1), a reduced order model $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ can be constructed by a pair of projection matrices V and W and the relations from (3.6). As a consequence, in this framework finding a good approximation $\hat{\mathbf{H}}$ is equivalent to a pair of projection matrices V, W such that the reduced order model constructed $\hat{\mathbf{H}}$ by projection has a similar behavior as $\mathbf{G}$.

In the reasoning, only the subspaces $span(V)$ and $span(W)$ play an important role in the projection framework and the projection matrices are not important in this regard. Indeed, this result can be properly stated mathematically as follows :

Proposition 3.2 ([Gallivan et al., 2004b])**.** *Let T_1, T_2 be two non-singular matrices. Then, the projected transfer function $\hat{\mathbf{H}}(s) = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B}$ obtained by the projection matrices V, W is unchanged if we replace the projectors by $\tilde{V} = VT_1$ and $\tilde{W} = WT_2$.*

Proof. Note that the matrix V (respect. W) span the same subspace as the matrix $\tilde{V}$ (respect. $\tilde{W}$). Hence,

$$\begin{aligned} \hat{\mathbf{H}}(s) &= \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B} \\ &= CV(sW^TEV - W^TAV)^{-1}W^TB \\ &= CVT_1(sT_2^TW^TEVT_1 - T_2^TW^TAVT_1)^{-1}T_2^TW^TB \\ &= C\tilde{V}(s\tilde{W}^TE\tilde{V} - \tilde{W}^TA\tilde{V})^{-1}\tilde{W}^TB \end{aligned}$$

which concludes the proof. □

Proposition 3.2 states that only the projection spaces $span(V)$ and $span(W)$ are important for the construction of the reduced order model $\hat{\mathbf{H}}$. In other words, similar projection matrices lead to the same reduced order model. As a consequence, in this framework, finding a reduced order model consists in finding projection spaces encoded by W and V which determine the approximation. However, different projection matrices can lead to the same reduced order model, as long as they span the same subspace.

A projection framework for model approximation was introduced by [Villemagne and Skelton, 1987]. A great amount of model approximation techniques can be stated using the projection-based framework. This chapter does not intend to provide an extensive review of all of them. Here there is a list of model approximation techniques that are not discussed in what follows:

- **Hankel norm approximation:** this procedure is presented in [Glover, 1984]. The theory enables to construct reduced order models which are optimal in the sense of the Hankel-norm (norm induced by the Hankel operator linked to the system). This approach provides a bound of the $\mathcal{H}_\infty$ mismatch error norm which is related to the Hankel singular values that were neglected. This bound is the same as the one presented in the balanced truncation (see Section 3.2). The article [Glover et al., 1988] extended the Hankel-norm approximation for infinite dimensional models. The book [Sasane, 2002] gives a good overview of the Hankel approximation in the larger context.

- **Weighted and frequency limited model approximation :** weighted model approximation enables to construct a reduced order model $\hat{\mathbf{H}}$ which is a good approximation of a system $\mathbf{G}$ after the application of a filter $\mathbf{W}$. In order words, the goal is to find $\hat{\mathbf{H}}$, such that the weighted error $\|(\mathbf{G}-\hat{\mathbf{H}})\mathbf{W}\|$ is small. See [Halevi, 1992], [Anić et al., 2013] and [Breiten et al., 2015] for some references. Similarly, frequency limited model approximation seeks a reduced order model $\hat{\mathbf{H}}$ which is a good approximation of $\mathbf{G}$ over a given frequency bound $\Omega = [\omega_1, \omega_2]$, $0 \leq \omega_1 < \omega_2$. In this case the $\mathcal{H}_{2,\Omega}$ norm (see Chapter 2 for definition) is used to measure the quality of the approximation. See [Vuillemin, 2014], [Vuillemin et al., 2014a] for a pole/residue-based approach, and [Petersson, 2013], [Petersson and Löfberg, 2014] for a limited gramian-based approach.

- **Proper orthogonal decomposition :** the philosophy is to snapshot the time-domain simulations of a possibly non-linear model. Generally, empirical gramians are built to project the initial model which leads to the so-called Proper Orthogonal Decomposition. See [Antoulas, 2005, Chapter 9] and references therein for further information on this method.

- **Dominant poles :** dedicated algorithms have been developed to efficiently compute iteratively the dominant poles of a large-scale model. In [Martins et al., 1996], the so-called dominant pole algorithm is presented. This algorithm uses the Newton's method to compute a dominant pole of a SISO model. Later, this algorithm is improved and extended to a robust and efficient method for MIMO systems in [Rommes and Martins, 2006] and for second order models in [Rommes and Martins, 2008]. Then, [Rommes and Sleijpen, 2008] presents some convergent properties, theoretical results and some comparisons with the Rayleigh quotient iteration.

For a general overview, the interested reader can refer to the books [Antoulas, 2005] and [Schilders et al., 2008], the classical survey papers [Gugercin and Antoulas, 2004], [Antoulas et al., 2001] and the recent survey papers [Beattie and Gugercin, 2015], [Baur et al., 2014]. In addition, for model approximation of parametric dynamical systems, topic which is not addressed in this thesis, see the recent survey of [Benner et al., 2015]. For differential algebraic systems see also [Benner and Stykel, 2015].

In what follows, we present two well-known projection-based methods, named the modal/balanced truncation and model approximation by moment matching.

3.2 Model approximation by truncation

In this section, *model approximation by truncation* is first defined. Then, two classical techniques are presented : *modal* and *balanced truncations*. Firstly, let us start by defining model approximation by truncation. Let us consider a system $\mathbf{G} = (E, A, B, C)$, whose realization is given by

$$\mathbf{G} = \left\{ \begin{array}{rcl} E\dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = & C\mathbf{x}(t) \end{array} \right. \tag{3.7}$$

The state vector $\mathbf{x}(t)$ can be decomposed between the states which must be retained $\mathbf{x}_1(t) \in \mathbb{R}^n$ and those which will be neglected $\mathbf{x}_2(t) \in \mathbb{R}^{N-n}$. Hence,

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix}$$

and the state space representation (3.7) can be structured as follows,

$$E = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}, \quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2. \end{bmatrix} \tag{3.8}$$

The truncated state-space realization $\hat{\mathbf{H}}$ of $\mathbf{G}$ only keeps the sub-matrices associated with $\mathbf{x}_1(t)$. The order of $\hat{\mathbf{H}}$ is then given by the number of states which have been retained. As a consequence, $\hat{\mathbf{H}} = (E_{11}, A_{11}, B_1, C_1) = (W^T EV, W^T AV, BV, W^T C)$, where the projection spaces are encoded by the matrices $W, V \in \mathbb{R}^{N\times n}$,

$$V = \begin{bmatrix} I_n \\ 0 \end{bmatrix} \in \mathbb{R}^{N\times n}, \quad \text{and} \quad W = V.$$

The procedure of truncation on the canonical basis can be considered as an orthogonal projection since $V = W$. In general, the truncated model $\hat{\mathbf{H}}$ does not have any particular property from the original model $\mathbf{G}$. In particular, any property such as stability, controlabillity, etc. might be lost. This procedure is often interesting when there is a physical meaning behind. For example, if some of the states are faster than others, one might consider to use a simple truncation. One advantage is that the state space variables of $\hat{\mathbf{H}}$ keep a physical meaning.

Another way to proceed is to find a particular linear transformation leading the system to a similar state-space representation. If in this new state-space the states are organized following their degree of importance, then the reduced order model obtained by truncation will be a better approximation. In what follows, two approaches based on this idea are presented: *modal* and *balanced truncations*.

3.2.1 Modal truncation

Model approximation by truncation is a technique widely used in many engineer applications due to its simplicity and physical meaning. The idea behind is to find the modal (pole/residue) decomposition of the full order model $\mathbf{G}$ and then construct a reduced order model preserving the more relevant dynamics. In general, the high frequency modes are neglected, since they represent fast dynamics playing a secondary role in the transfer, at least for control engineers. One advantage of modal truncation is that the reduced order model preserves some properties from the full order model, *e.g.* stability, controllability and observability. In addition, the preserved modes still have a physical meaning. The major drawback is that one needs to compute the eigenvalue decomposition of the pencil (E, A) and this task can be very expensive for large-scale systems.

Suppose that the full order model $\mathbf{G}$ is described by (3.7) and that the pencil (E, A) is diagonalizable, *i.e.*, there is a non-singular matrix $X \in \mathbb{C}^n$ such that $AX = EX\Sigma$, $\Sigma = \mathbf{diag}(\lambda_1, \ldots, \lambda_N)$. Then, the transformation

$$\mathbf{G} = (EX, AX, B, CX^{-1}) = (I_N, \underbrace{\Sigma}_{\tilde{A}}, \underbrace{(EX)^{-1}B}_{\tilde{B}}, \underbrace{CX}_{\tilde{C}})$$

leads the modal state space realization

$$\tilde{A} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} \mathbf{b}_1^T \\ \vdots \\ \mathbf{b}_N^T \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} \mathbf{c}_1 & \ldots & \mathbf{c}_N \end{bmatrix} \tag{3.9}$$

whose transfer function can be written as

$$\mathbf{G}(s) = \sum_{k=1}^{N} \frac{\mathbf{c}_i \mathbf{b}_i^T}{s - \lambda_i}.$$

Then, the reduced order model $\hat{\mathbf{H}}$ is obtained by applying a truncation over the modal realization (3.10). Hence, $\hat{\mathbf{H}} = (I_n, \hat{A}, \hat{B}, \hat{C})$, where

$$\hat{A} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \quad \hat{B} = \begin{bmatrix} \mathbf{b}_1^T \\ \vdots \\ \mathbf{b}_n^T \end{bmatrix} \quad \hat{C} = \begin{bmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_n \end{bmatrix}. \tag{3.10}$$

In the projection-based framework, let $\mathbf{G} = (E, A, B, C)$ and suppose X be the left eigenvectors of the pencil (E, A). Then, the reduced order model obtained by truncation reads $\hat{\mathbf{H}} = (W^T EV, W^T AV, W^T B, CV)$ where the projection matrices are :

$$V = X \begin{bmatrix} I_n \\ 0 \end{bmatrix} \in \mathbb{C}^{N \times n} \quad W = (EX)^{-T} \begin{bmatrix} I_n \\ 0 \end{bmatrix} \in \mathbb{C}^{N \times n} \tag{3.11}$$

Moreover, the approximation error

$$\mathbf{G}(s) - \hat{\mathbf{H}}(s) = \sum_{i=n+1}^{N} \frac{\mathbf{c}_i \mathbf{b}_i^T}{s - \lambda_i},$$

and from this expression a $\mathcal{H}_\infty$ bound can be derived as follows:

$$\begin{aligned} \|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{H}_\infty} &= \max_{\omega \in \mathbb{R}} \left\| \sum_{i=n+1}^{N} \frac{\mathbf{c}_i \mathbf{b}_i^T}{i\omega - \lambda_i} \right\|_2 \\ &\leq \max_{\omega \in \mathbb{R}} \sum_{i=n+1}^{N} \left\| \frac{\mathbf{c}_i \mathbf{b}_i^T}{i\omega - \lambda_i} \right\|_2 \\ &\leq \sum_{i=n+1}^{N} \frac{\|\mathbf{c}_i\|_2 \|\mathbf{b}_i\|_2}{|Re(\lambda_i)|}. \end{aligned}$$

In practice, this $\mathcal{H}_\infty$ approximation bound is very conservative (see [Vuillemin, 2014, Chapter 3]). This comes from the commutation between the sum and the max operator. In general, modal truncation is less effective compared to other model approximation methods when the task is to reduce the $\mathcal{H}_2$ or $\mathcal{H}_\infty$ approximation errors.

To sum up, modal approximation is a simple approximation technique that has the advantage of preserving some modes of the full order model. This is particularly interesting when some modes with a particular physical meaning must be retained. Yet, modal truncation is often less efficient than other methods in the sense of the $\mathcal{H}_2$ or $\mathcal{H}_\infty$ norms. See [Vuillemin, 2014] and the examples therein for more details.

3.2.2 Balanced truncation

One of the most widely used model approximation methods is the so-called *balanced truncation*. It was first introduced in [Mullis and Roberts, 1976] and then later by [Moore, 1981] in systems and control literature. This method has the nice physical interpretation: the states that are removed are those that are difficult to reach and to observe simultaneously, and consequently, which least contribute to the energy transfer from the input to the output. Therefore, the starting point of balanced truncation is to suitably characterize "observability" and "controllability" of a system, which are quantified by the gramians (see Chapter 2). The method can be summarized in finding the balanced realization (balancing step) and, subsequently, eliminating the states that are associated to the least Hankel singular values (truncation step).

Balanced realization step: the balanced realization is the state-space representation where the states are ordered by their importance in the input-output energy transfer. Let us recall its definition.

Definition 3.3 (Balanced realization). *A model* $\mathbf{H}_b = (E_b, A_b, B_b, C_b)$ *is represented in the balanced state-space realization if the associated gramians* $\mathcal{P}$ *and* $\mathcal{Q}$ *(defined in Chapter 2, equations (2.18) and (2.19)) satisfy*

$$\mathcal{P} = E_b^T \mathcal{Q} E_b = \Sigma = \boldsymbol{diag}(\sigma_1, \ldots, \sigma_N), \tag{3.12}$$

where $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_N \geq 0$ *are the Hankel singular values of* $\mathbf{H}_b$, i.e., $\sigma_i = \sqrt{\lambda_i(\mathcal{P} E_b^T \mathcal{Q} E_b)}$, *for* $i = 1, \ldots, N$.

It is worth mentioning that the gramians $\mathcal{P}$ and $\mathcal{Q}$ both depend on the realization (E, A, B, C). However, the Hankel singular values are invariants, *i.e.*, do not depend on the realization. In order to find the transformation leading the system to the balanced realization, let us apply on the realization a change in the variables of the form

$$\mathbf{z} = T\mathbf{x}, \text{ with } T \in \mathbb{R}^{N \times N} \text{ non-singular.}$$

This implies that the Gramians in this new realization are

$$\tilde{\mathcal{P}} = T \mathcal{P} T^T, \ \tilde{\mathcal{Q}} = T^{-T} \mathcal{Q} T^{-T}.$$

Hence, finding the balanced realization is similar as finding a transformation T which diagonalizes both $\mathcal{P}$ and $E^T \mathcal{Q} E$. This consists in the simultaneous diagonalization of two positive definite matrices. This is always possible and there are different ways to construct this transformation (see [Gugercin and Antoulas, 2004]). The following lemma recalls the so-called *square root algorithm* (see [Laub et al., 1987]).

Lemma 3.4. *Let* $\mathcal{P} = UU^T$ *and* $\mathcal{Q} = LL^T$ *be the Cholesky decompositions of the gramians* $\mathcal{P}$ *and* $\mathcal{Q}$. *In addition, let* $L^T E U = M \Sigma N$ *be the singular value decomposition of* $L^T E U$, *where* $M^T = M^{-1}$ *and* $N^T = N^{-1}$ *are orthogonal matrices, and* $\Sigma = \boldsymbol{diag}(\sigma_1, \ldots, \sigma_N)$ *are the Hankel singular values of the system. Then, the state transformation* $\mathbf{z} = T\mathbf{x}$ *given by*

$$T = \Sigma^{-\frac{1}{2}} M^T L^T E \quad \text{and} \quad T^{-1} = U N \Sigma^{-\frac{1}{2}},$$

allows to transform the system in its balanced realization.

Hence, Lemma 3.4 gives a constructive way to find the balanced realization of a system. This generalized version takes into consideration the matrix E and it is due to [Hsu et al., 1983]. There exist other algorithms to find this transformation and they will briefly discussed latter.

Truncation step: once the system is in its balanced realization $\mathbf{H} = (E_b, A_b, B_b, C_b)$, the balanced truncation is obtained by eliminating the states corresponding **to the least Hankel singular values**. In this way, let us write

$$\Sigma = \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix}, \ E_b = \begin{bmatrix} E_1 & E_2 \\ E_3 & E_4 \end{bmatrix}, \ A_b = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}, \ B_b = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \ C_b = \begin{bmatrix} C_1 & C_2 \end{bmatrix},$$

where $\Sigma_1 = \mathbf{diag}(\sigma_1, \ldots, n)$ and $\Sigma_2 = \mathbf{diag}(\sigma_{n+1}, \ldots, \sigma_N)$. Therefore, the reduced-order model obtained by balanced truncation is

$$\hat{\mathbf{H}} = \left(\begin{array}{c|c} E_1, A_1 & B_1 \\ \hline C_1 & 0 \end{array}\right) \tag{3.13}$$

The following theorem provides some properties of the balanced truncation.

Theorem 3.5 (Balanced truncation properties [Antoulas, 2005]). *Let $\mathbf{G} = (E, A, B, C)$ be a stable, controllable and observable realization. Then the reduced-order model $\hat{\mathbf{H}} = (E_1, A_1, B_1, C_1)$ of order n obtained by balanced truncation as in (3.13), with $\sigma_n > \sigma_{n+1}$, has the following properties: it is stable, balanced, minimal and it satisfies*

$$\sigma_{n+1} \leq \|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{H}_\infty} \leq 2 \sum_{i=n+1}^{N} \sigma_i.$$

The full proof of Theorem 3.5 can be found in [Antoulas, 2005, Chap.7]. The proof of stability of the reduced-order model is based on inertia results (see [Antoulas, 2005, Chap.6]). The theorem provides an important reassuring insight into the potential of balanced truncation: since the singular numbers of exponentially stable LTI systems generally decay exponentially, the upper bound of Theorem 3.5 is not expected to be much larger than the lower bound.

In order to obtain the reduced-order model by projection, let us consider the matrices L, U, M, N and Σ from Lemma 3.4. One can considered that the matrices Σ, M and N can be partitioned as,

$$\Sigma = \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix}, \quad M = \begin{bmatrix} M_1 & M_2 \end{bmatrix}, \quad N = \begin{bmatrix} N_1 & N_2 \end{bmatrix},$$

where $M_1 \in \mathbb{R}^{N \times n}$, $N_1 \in \mathbb{R}^{N \times n}$ and $\Sigma_1 \in \mathbb{R}^{n \times n}$. Therefore, the balanced truncation can be simply stated using the projection framework as follows :

Theorem 3.6 (Balanced truncation - Projection framework). *Using the notation predefined, let*

$$W^T = \Sigma_1^{-\frac{1}{2}} M_1^T L^T \in \mathbb{R}^{n \times N} \quad \text{and} \quad V = U N_1 \Sigma_1^{-\frac{1}{2}} \in \mathbb{R}^{N \times n},$$

be projection matrices. These matrices lead to the construction of the reduced-order model $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ by balanced truncation as follows

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

Notice that by construction, the matrix $\hat{E}$ is equal to identity. To sum up, balanced truncation is a model approximation method that generally yields a good approximation, preserves stability and provides an *a priori* error bound. The drawback is that it requires to find the solution of Lyapunov equations which can be a difficult task for large-scale systems. In addition, the provided error bound is conservative in practice (see [Vuillemin, 2014, Chapter 3]).

A great amount of extensions of the balanced truncation and numerical improvements are available. Here is a non-exhaustive list of examples:

Extensions :

- **Differential algebraic equations** : see [Bender, 1987; Stykel, 2004] for extension of the theory. See [Stykel, 2006] for application to the semidiscretized Stokes equation. In addition, the paper [Mehrmann and Stykel, 2005] gives a very good survey of balanced truncation for differential algebraic systems.

- **Second-order systems** : both papers [Meyer and Srinivasan, 1996] and [Chahlaoui et al., 2006] develop different structures preserving model reduction methods for second-order systems.

- **Unstable systems** : it was first developed in [Therapos, 1989] for unstable non-minimal systems and then generalized in [Zhou et al., 1999].

- **Inhomogeneous initial conditions** : [Heinkenschloss et al., 2011] presents a new method allowing model approximation for systems with inhomogeneous initial conditions, by adding auxiliary inputs derived from the initial conditions.

- **Infinite dimensional systems** : [Glover et al., 1988] extends balanced truncation to the class of infinite-dimensional continuous-time systems. More recently, papers [Reis and Selig, 2014; Guiver and Opmeer, 2014] generalize those results.

- **Time-varying systems** : one extension of balanced truncation for time-varying systems is developed in [Sandberg and Rantzer, 2004]. Applications using a projection method are presented in [Sandberg, 2006].

- **Nonlinear systems** : Balanced truncation for nonlinear systems is introduced in [Scherpen, 1993]. In addition, balanced truncation based on the empirical gramians to project the initial model leads to the so-called Proper Orthogonal Decomposition. See [Antoulas, 2005] and the references therein for further information.

The survey [Gugercin and Antoulas, 2004] presents some other methods that were not mentioned here, *e.g.*, positive real and frequency weighted balancing. In addition, model approximation by balanced truncation requires the solution of Lyapunov equations which can be a hard task in the large-scale setting. Some well know methods, *e.g.*, the Bartels-Stewart algorithm, Hammarling's method, sparse Alternating Direction Implicit (ADI) Krylov subspaces, can be found in the survey paper [Benner and Saak, 2013].

In the next section, other projection-based method are presented. The philosophy proposed no longer to eliminate some states, but rather to find a reduced order model satisfying some interpolation conditions.

3.3 Model approximation by interpolation

Generally, interpolation consists in finding a simple function, *e.g*, polynomials, splines, trigonometric functions, passing through some given data (x_i, y_i) for $i = 1, \ldots, n$. The same idea can be employed to build finite dimensional reduced-order models, which are essentially rational functions. That motivates the introduction of rational interpolation methods. Given a transfer function $\mathbf{G}(s) \in \mathbb{C}^{n_y \times n_u}$, one could consider the problem of finding a rational transfer function $\hat{\mathbf{H}} = (E, A, B, C)$ which satisfies some point-wise interpolation conditions as

$$\mathbf{G}(\sigma_k) = \hat{\mathbf{H}}(\sigma_k) \in \mathbb{C}, \text{ at each interpolation point } \quad \sigma_k \in \mathbb{C}. \tag{3.14}$$

However, those point-wise interpolation is to strong since, for each interpolation point σ_k, it requires $n_y \times n_u$ constraints to be satisfied. Hence, instead of point wise conditions like (3.14), we require only that the interpolating matrix function match the original only along certain directions, known as *right* and *left tangential directions*. This new interpolation problem is called *tangential interpolation problem* and can be stated as follows.

Problem 3.7 (**Tangential interpolation problem**). *Given a set of n right interpolation points $\{\sigma_i\}_{i=1}^n \in \mathbb{C}$, n right tangential directions $\{\mathbf{r}_i\}_{i=1}^n \in \mathbb{C}^{n_u}$, n left interpolation points $\{\mu_i\}_{i=1}^n$ and n left tangential directions $\{\mathbf{l}_i\}_{i=1}^n \in \mathbb{C}^{n_y}$ find a reduced order model $\hat{\mathbf{H}}$ satisfying*

$$\left.\begin{array}{rcl} \mathbf{G}(\sigma_i)\mathbf{r}_i & = & \hat{\mathbf{H}}(\sigma_i)\mathbf{r}_i \\ \mathbf{l}_i^T\mathbf{G}(\mu_i) & = & \mathbf{l}_i^T\hat{\mathbf{H}}(\mu_i) \end{array}\right\} \quad \text{for } i = 1, \ldots, n. \tag{3.15}$$

In addition, we say that $\hat{\mathbf{H}}$ satisfies a bitangential Hermite condition with respect to $\mathbf{G}(s)$ at $\sigma \in \mathbb{C}$ along the right tangent direction $\mathbf{r}_i \in \mathbb{C}^{n_u}$ and the left tangent direction $\mathbf{l}_i \in \mathbb{C}^{n_y}$, if

$$\mathbf{l}_i^T\mathbf{G}'(\sigma_i)\mathbf{r}_i = \mathbf{l}_i^T\hat{\mathbf{H}}'(\sigma_i)\mathbf{r}_i, \tag{3.16}$$

where $\mathbf{G}'(\sigma_i)$ denotes the differentiation of $\mathbf{G}(s)$ with respect to s.

Notice that in the SISO case, the right and left tangential directions are scalars and can be choosen as one, *e.g.*, $\mathbf{r}_i, \mathbf{l}_i = 1$, for $i = 1, \ldots, n$. It is worth to note that tangential interpolation generalizes the conditions from (3.14). Indeed, if for the same interpolation point σ_k, the canonical vectors $\mathbf{e}_i$ are chosen as right tangential directions, then we obtain (3.14). The tangential interpolation problem was proposed and solved in [Gallivan et al., 2004a]. It will be shown in Chapter 4 that the tangential interpolation considerations are adequate to characterize necessary optimality conditions for the $\mathcal{H}_2$ optimal model approximation problem.

In the following, we briefly present the notion of moment matching and the historical developments behind the solution of Problem 3.7.

3.3.1 Model approximation by moment matching

Historically, the first approach tackling Problem 3.7 was the so-called *model approximation by moment matching*. The moments of a transfer function are presented in Definition 3.8 and the problem of model approximation by moment matching is introduced in Problem 3.9.

Definition 3.8 (**Moments of a transfer function**). *Let us consider a model $\mathbf{G} = (E, A, B, C)$. Its transfer function $\mathbf{G}(s) = C(sE - A)^{-1}B \in \mathbb{C}^{n_y \times n_u}$ can be expanded using the Taylor series in the neighborhood of a given shift point $s_0 \in \mathbb{C}$ if s_0 is not a pole of $\mathbf{G}(s)$. Then,*

$$\mathbf{G}(s) = \sum_{k=0}^{\infty} \eta_i(s_0)\frac{(s - s_0)^k}{i!},$$

and

$$\eta_i(s_0) = (-1)^k \frac{d^i}{ds^i}\mathbf{G}(s)\Big|_{s=s_0} \in \mathbb{C}^{n_y \times n_u}$$

is the i-th moment[a] of $\mathbf{G}(s)$ at s_0.

[a] Notice that this definition is related to the notion of moment of a measure. Given an impulse response of a model $\mathbf{h}(t) \in \mathbb{R}^{n_y \times n_u}$, the k-th moment of $\mathbf{h}(t)$ at $s = s_0$ is given by $\eta_k(s_0) = \int_0^\infty t^k \mathbf{h}(t) e^{-s_0 t} dt$. This

can be seen as the i-th moment of the measure $d\mu = \mathbf{h}(t)e^{-s_0 t}dt$.

Notice that moments are invariants of the system. In order to simplify the notation, let us denote $A_\sigma = (A - \sigma E)$. Let $\mathbf{G}(s) = C(sE - A)^{-1}B$ and $s_0 \in \mathbb{C}$ be given, then the i-th moment at s_0 is given by

$$\eta_i(s_0) = -C(A_{s_0}^{-1}E)^i A_{s_0}^{-1}B, \text{ for } i = 1, 2, \ldots$$

In addition, if $s_0 = \infty$, the moments are defined as follows

$$\eta_i(\infty) = C(E^{-1}A)^{i-1}E^{-1}B. \text{ for } i = 1, 2, \ldots.$$

In this case, the moments are called *Markov parameters* and are related to the *partial realization problem* (see [Antoulas, 2005, Chapter 4]).

Now, let us define the problem of model approximation by moment matching.

Problem 3.9 (Moment matching problem). *Given a N-th oder model $\mathbf{G}$ and r shift points $\{s_1, \ldots, s_r\} \in \mathbb{C}$, whose i-th moment is denoted by $\eta_i(s_k)$, the moment matching problem consists in finding a reduced-order model $\hat{\mathbf{H}}$, whose i-th moment, is denoted by $\hat{\eta}_i(s_k)$, satisfies*

$$\eta_i(s_k) = \hat{\eta}_i(s_k) \quad \text{for} \quad i = 1, \ldots, n \text{ and } k = 1, \ldots, r.$$

Notice that Problem 3.9 for moments of order 0 and 1 is equivalent to Hermite interpolation problem 3.7 in the SISO case. In the MIMO case, moment matching corresponds to point-wise interpolation and the tangential interpolation problem is a generalization of it. The simplest approach to solve Problem 3.9 is to compute explicitly the moments of $\mathbf{G}$ and to find a reduced-order model $\hat{\mathbf{H}}$ which matches those moments (see [Villemagne and Skelton, 1987] and [Grimme, 1997]). However, computing the moments is a numerically expensive and ill-conditioned problem.

The method described in this section is numerically efficient and relies on the projection framework. The following theorem enables to construct the reduced-order model implicitly without explicitly computing the moments. It was first presented in [Grimme, 1997] in the framework of matching moments and then generalized to the tangential interpolation framework by [Gallivan et al., 2004a].

Theorem 3.10 (Tangential interpolation by projection). *Let $\mathbf{G} = (E, A, B, C)$ be a system whose transfer function is $\mathbf{G}(s) = C(sE - A)^{-1}B \in \mathbb{C}^{n_y \times n_u}$ and $\hat{\mathbf{H}}$ denotes a reduced-order model obtained by projection as in (3.6) using the matrices $\mathbf{V}$ and $\mathbf{W}$. Let $\sigma, \mu \in \mathbb{C}$ be interpolation points which are not poles of $\mathbf{G}$. Then, if*

$$(\sigma E - A)^{-1}B\mathbf{r} \in span(V) \tag{3.17}$$

then the following right tangential interpolation condition

$$\mathbf{G}(\sigma)\mathbf{r} = \hat{\mathbf{H}}(\sigma)\mathbf{r}, \tag{3.18}$$

is satisfied. If

$$\left(\mathbf{l}^T C(\sigma E - A)^{-1}\right)^T \in span(W), \tag{3.19}$$

then the following left tangential interpolation condition

$$\mathbf{l}^T\mathbf{G}(\mu) = \mathbf{l}^T\hat{\mathbf{H}}(\mu), \tag{3.20}$$

is satisfied. Moreover, if $\sigma = \mu$ and both (3.27) and (3.29) hold, then the bitangential Hermite condition

$$\mathbf{l}^T \mathbf{G}'(\sigma)\mathbf{r} = \mathbf{l}^T \hat{\mathbf{H}}'(\sigma)\mathbf{r}, \tag{3.21}$$

is satisfied.

Proof. We only check the right tangential interpolation condition. The other conditions follow similarly and the development can be found in [Gallivan et al., 2004a] and [Gugercin et al., 2008]. Let us write

$$\begin{aligned}
\hat{\mathbf{H}}(\sigma)\mathbf{r} &= \hat{C}(\sigma\hat{E} - \hat{A})^{-1}\hat{B}\mathbf{r} \\
&= CV\left(\sigma W^T EV - W^T AV\right)^{-1} W^T B\mathbf{r} \\
&= CV\left(\sigma W^T EV - W^T AV\right)^{-1} W^T \underbrace{(\sigma E - A)(\sigma E - A)^{-1}}_{=I_N} B\mathbf{r}.
\end{aligned}$$

Notice that $(\sigma E - A)^{-1}B\mathbf{r} \in span(V)$. Hence there exists $\mathbf{x} \in \mathbb{C}^n$ such that

$$(\sigma E - A)^{-1}B\mathbf{r} = V\mathbf{x}.$$

Therefore,

$$\begin{aligned}
\hat{\mathbf{H}}(\sigma)\mathbf{r} &= CV\left(\sigma W^T EV - W^T AV\right)^{-1} W^T(\sigma E - A)V\mathbf{x} && (3.22)\\
&= CV(\sigma W^T EV - W^T AV)^{-1}(\sigma W^T EV - W^T AV)\mathbf{x} && (3.23)\\
&= CV\mathbf{x} = C(\sigma E - A)^{-1}B\mathbf{r} = \mathbf{G}(\sigma)\mathbf{r}. && (3.24)
\end{aligned}$$

□

Theorem 3.10 enables to construct the projection matrices V and W such that the reduced-order model obtained by projection satisfies left and right tangential interpolation conditions. Furthermore, if some left and right tangential interpolation condition are imposed for the same shift σ, the bitangential Hermite condition (3.31) is satisfied for free. In addition, a similar theorem enforcing higher order derivatives interpolation can be stated (see for example [Grimme, 1997; Gallivan et al., 2004a] and [Antoulas et al., 2010, Theorem 2]).

Projection matrices V and W can be constructed to match moments at the same interpolation point. This can be done in the robust numerical framework of rational Krylov methods (see [Grimme, 1997; Antoulas, 2005]), in particular using Lanczos and Arnoldi methods (see [Saad, 2003, Chapter 6] and [Antoulas, 2005, Chapter 11]). They enable the construction of the projectors V, W without explicitly computing the moments and the procedure can be implemented iteratively. The following papers [Gallivan et al., 1994; Feldmann and Freund, 1995; Freund, 2003; Antoulas et al., 2010; Bai, 2002] and the references therein develop related work to model approximation using rational Krylov methods and interpolation. In addition, the projectors V, W can be encoded as a solution of Sylvester equations. Indeed, there is a connection between Krylov subspaces and Sylvester equations which was reveled by the PhD thesis [Vandendorpe, 2004] and the paper [Gallivan et al., 2004b].

3.3.2 Generalized coprime framework

Until now, the full and the reduced-order models were assumed to be finite dimensional and described by a realization $\mathbf{G} = C(sE - A)^{-1}B$ and $\hat{\mathbf{H}} = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B}$. As it was discussed in Chapter 1, even if this representation is quite general and a wide range of linear dynamical

systems can be converted to this form, some more complex structures cannot be represented by those equations. Problem formulations often lead to somewhat different structures that reflect the underlying physics or other important system features. In the following, we introduce the results from [Beattie and Gugercin, 2009a]. It presents interpolatory methods that preserve relevant system structure in reduced models.

In what follows, we consider a transfer function having the following generalized coprime representation:

$$\mathbf{G}(s) = \mathbf{C}(s)\mathbf{K}^{-1}(s)\mathbf{B}(s). \tag{3.25}$$

where both $\mathbf{C}(s) \in \mathbb{C}^{n_y \times N}$, $\mathbf{B}(s) \in \mathbb{C}^{N \times n_u}$ are analytic in the right half plane, and $\mathbf{K}(s) \in \mathbb{C}^{N \times N}$ is analytic and full rank throughout the right half plane. Examples of systems represented by the coprime representation are given in what follows :

- **Second-order systems:** let $\mathbf{G}$ be a second order system whose transfer function is given by

  $$\mathbf{G}(s) = C(s^2M + sK + D)^{-1}B,$$

  where $M, K, D \in \mathbb{R}^{N \times N}$, $B \in \mathbb{R}^{N \times n_u}$ and $C \in \mathbb{R}^{n_y \times N}$. Its coprime representation is given by $\mathbf{C}(s) = C$, $\mathbf{B}(s) = B$ and $\mathbf{K}(s) = s^2M + sK + D$. The structured preserving model approximation problem is to find $\hat{\mathbf{H}}$ a reduced-order model whose transfer function is given by

  $$\hat{\mathbf{H}}(s) = \hat{C}(s^2\hat{M} + s\hat{K} + \hat{D})^{-1}\hat{B},$$

  where $\hat{M}, \hat{K}, \hat{D} \in \mathbb{R}^{n \times n}$, $\hat{B} \in \mathbb{R}^{n \times n_u}$, $\hat{C} \in \mathbb{R}^{n_y \times n}$ and $n \in \mathbb{N}, n \ll N$.

- **Time-delay systems:** let $\mathbf{G}$ be a time-delay system whose the transfer function reads

  $$\mathbf{G}(s) = C\left(sE - \sum_{k=1}^{n_\tau} A_k e^{-s\tau_k}\right)^{-1} B,$$

  where $A_k \in \mathbb{R}^{N \times N}$, for $k = 1, \ldots, n_\tau$, $B \in \mathbb{R}^{N \times n_u}$ and $C \in \mathbb{R}^{n_y \times N}$. Its coprime representation is given by $\mathbf{C}(s) = C$, $\mathbf{B}(s) = B$ and $\mathbf{K}(s) = sE - \sum_{k=1}^{n_\tau} A_k e^{-s\tau_k}$. Similarly, the structured preserving model approximation problem is to find $\hat{\mathbf{H}}$, a reduced-order model, whose transfer function is given by

  $$\hat{\mathbf{H}}(s) = \hat{C}\left(s\hat{E} - \sum_{k=1}^{n_\tau} \hat{A}_k e^{-s\tau_k}\right)^{-1} \hat{B},$$

  where $\hat{A}_k \in \mathbb{R}^{n \times n}$, for $k = 1, \ldots, n_\tau$, $\hat{B} \in \mathbb{R}^{n \times n_u}$, $\hat{C} \in \mathbb{R}^{n_y \times N}$ and $n \ll N$.

More generally, the goal is to construct a reduced-order model $\hat{\mathbf{H}}$ having the same structure as $\mathbf{G}$. This is possible by theorem from [Beattie and Gugercin, 2009a], which enables to generalize tangential interpolation for generalized coprime representation.

Theorem 3.11 (Tangential interpolation for generalized coprime factors). *Let* $\mathbf{G}$ *be a system whose transfer function is given by*

$$\mathbf{G}(s) = \mathbf{C}(s)\mathbf{K}^{-1}(s)\mathbf{B}(s),$$

and $\hat{\mathbf{H}}(s) = \hat{\mathbf{C}}(s)\hat{\mathbf{K}}^{-1}(s)\hat{\mathbf{B}}(s)$ *be the reduced-order model transfer function obtained by the projection matrices as follows*

$$\hat{\mathbf{K}}(s) = W^T\mathbf{K}(s)V, \quad \hat{\mathbf{B}}(s) = W^T\mathbf{B}(s), \quad \text{and} \quad \hat{\mathbf{C}}(s) = \mathbf{C}(s)V. \tag{3.26}$$

Assume that σ and μ, the right and left interpolation points, are not poles of $\mathbf{C}(s)$, $\mathbf{K}(s)$ and $\mathbf{B}(s)$, and the matrices $\mathbf{K}(\sigma)$ and $\mathbf{K}(\mu)$ have full rank. In addition, let $\mathbf{r} \in \mathbb{C}^{n_u}$ and $\mathbf{l} \in \mathbb{C}^{n_y}$. Then, if

$$\mathbf{K}^{-1}(\sigma)\mathbf{B}(\sigma)\mathbf{r} \in span(V), \tag{3.27}$$

then the following right tangential interpolation condition

$$\mathbf{G}(\sigma)\mathbf{r} = \hat{\mathbf{H}}(\sigma)\mathbf{r}, \tag{3.28}$$

is satisfied. If

$$\left(\mathbf{l}^T\mathbf{C}(\sigma)\mathbf{K}^{-1}(\sigma)\right)^T \in span(W), \tag{3.29}$$

then the following left tangential interpolation condition

$$\mathbf{l}^T\mathbf{G}(\mu) = \mathbf{l}^T\hat{\mathbf{H}}(\mu), \tag{3.30}$$

is satisfied. Moreover, if $\sigma = \mu$ and both (3.27) and (3.29) hold, then the bitangential Hermite condition

$$\mathbf{l}^T\mathbf{G}'(\sigma)\mathbf{r} = \mathbf{l}^T\hat{\mathbf{H}}'(\sigma)\mathbf{r}, \tag{3.31}$$

is satisfied.

Theorem 3.11 enables to perform tangential interpolation for a larger class of models. It generalizes the results from Theorem 3.10 with similar conditions on the projection matrices V, W. Some extensions were developed for other classes of models. Here is a non-exhaustive list of extensions and relative works:

- **Non-linear systems:** model approximation by moment matching for non-linear systems was developed in [Astolfi, 2010]. It generalizes the notion of moment for non-linear models. For bilinear systems using Volterra series, see [Bai and Skoogh, 2006] and the generalizations in [Breiten and Damm, 2010] and [Benner and Breiten, 2012].
- **Time-delay systems:** a similar approach to Theorem 3.11 was given in [Michiels et al., 2011] which proposes a Krylov-based model approximation for time-delay systems. Moment matching were also generalized for time-delay systems in [Scarciotti and Astolfi, 2016] and [Schulze and Unger, 2015].
- **Port-Hamiltonian systems:** model approximation by interpolation for port-Hamiltonian systems was studied in [Gugercin et al., 2012] and [Ionescu and Astolfi, 2013] in the linear context and in [Beattie and Gugercin, 2011] and [Chaturantabut et al., 2016] in the non-linear case.
- **Parametric systems**: the paper [Benner et al., 2015] is a recent survey on model approximation by projection applied to parametric systems.

In what follows, we present a framework enabling to derive reduced-order models based on interpolation of given frequency data, instead of the system realization.

3.4 Data-driven model reduction

So far, projection-based model approximation has been presented in such a way that it aims at finding a finite dimensional approximation $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ for a full-order model $\mathbf{G} = (E, A, B, C)$. In addition, Theorem 3.11 enables preserving-structure model approximation.

In this section, a new framework due to [Mayo and Antoulas, 2007] is presented. This framework permits to build a finite dimensional approximation $\hat{\mathbf{H}}$ using the evaluations of the transfer function of $\mathbf{G}$ only. Hence, with this framework one can construct an approximation $\mathbf{H} = (E, A, B, C)$ of $\mathbf{G}$, even if $\mathbf{G}$ is given by an irrational transfer function or simply if its realization is not available. The main advantage of this procedure is to avoid any discretization, which is usually necessary when dealing with an infinite-dimensional system.

3.4.1 Tangential interpolation data problem

We introduce the data-based framework for model approximation from [Mayo and Antoulas, 2007]. Instead of assuming known the frequency behavior of $\mathbf{G}(s)$ for all complex number $s \in \mathbb{C}$, we suppose that some frequency data of $\mathbf{G}$ are available only. Inspired by the tangential interpolation, given a set of n right interpolation points $\{\sigma_i\}_{i=1}^n \in \mathbb{C}$, n right tangential directions $\{\mathbf{r}_i\}_{i=1}^n \in \mathbb{C}^{n_u}$, n left interpolation points $\{\mu_i\}_{i=1}^n$ and n left tangential directions $\{\mathbf{l}_i\}_{i=1}^n \in \mathbb{C}^{n_y}$, the objective consists in finding a reduced order model $\hat{\mathbf{H}}$ satisfying,

$$\left\{ \begin{array}{rcl} \mathbf{G}(\lambda_i)\mathbf{r}_i & = & \mathbf{w}_i \\ \mathbf{l}_i^T\mathbf{G}(\mu_i) & = & \mathbf{v}_i \end{array} \right. . \tag{3.32}$$

We assume that only the right and left tangential evaluations of $\mathbf{G}$ are available. The goal is to find a reduced order model $\hat{\mathbf{H}}$ matching this frequency data set. This is summarized by the following problem.

Problem 3.12 (Tangential interpolation data problem). *Let*

$$\begin{array}{ll} \{(\lambda_i, \mathbf{r}_i, \mathbf{w}_i) | & \lambda_i \in \mathbb{C}, \mathbf{r}_i \in \mathbb{C}^{n_u}, \mathbf{w}_i \in \mathbb{C}^{n_y}\} \\ \{(\mu_i, \mathbf{l}_i, \mathbf{v}_i) | & \mu_i \in \mathbb{C}, \mathbf{l}_i \in \mathbb{C}^{n_y}, \mathbf{v}_i \in \mathbb{C}^{n_u}\} \end{array} \quad \text{for } i = 1, \ldots, n \tag{3.33}$$

be the tangential interpolation data set. Find a (minimal) reduced order model $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ satisfying the right and left interpolation conditions, respectively,

$$\left\{ \begin{array}{rcl} \hat{\mathbf{H}}(\lambda_i)\mathbf{r}_i & = & \mathbf{w}_i \\ \mathbf{l}_i^T\hat{\mathbf{H}}(\mu_i) & = & \mathbf{v}_i \end{array} \right. \text{. for } \quad i = 1, \ldots, n, \tag{3.34}$$

In the SISO case, Problem 3.12 is also known as *rational interpolation* (see [Anderson and Antoulas, 1990; Antoulas and Anderson, 1986; Ionita, 2013] and the references therein for some developments in the solution of this problem). Problem 3.12 supposes that the numbers of left and right interpolation conditions are the same. A more general tangential interpolation problem can be considered supposing that the number of left and right interpolation conditions can be different (see [Mayo and Antoulas, 2007]). For clarity, we present only the case where the number of right and left conditions are the same and readers should refer to [Mayo and Antoulas, 2007] for the general problem.

Let us introduce the following matrices associated with the interpolation problem

$$\left\{ \begin{array}{rcl} \Lambda & = & \mathbf{diag}(\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^{n \times n} \\ R & = & \begin{bmatrix} \mathbf{r}_1 & \ldots & \mathbf{r}_n \end{bmatrix} \in \mathbb{C}^{n_u \times n} \\ W & = & \begin{bmatrix} \mathbf{w}_1 & \ldots & \mathbf{w}_n \end{bmatrix} \in \mathbb{C}^{n_y \times n} \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{rcl} M & = & \mathbf{diag}(\mu_1, \ldots, \mu_n) \in \mathbb{C}^{n \times n} \\ L^T & = & \begin{bmatrix} \mathbf{l}_1 & \ldots & \mathbf{l}_n \end{bmatrix} \in \mathbb{C}^{n_y \times n} \\ V^T & = & \begin{bmatrix} \mathbf{v}_1 & \ldots & \mathbf{v}_n \end{bmatrix} \in \mathbb{C}^{n_u \times n} \end{array} \right. \tag{3.35}$$

These matrices represent the tangential interpolation Problem 3.12 and will be useful in later developments.

3.4.2 The Loewner framework

Now, we present the main tool of this approach: *the Loewner pencil*. The Loewner matrix $\mathbb{L}$ and the shifted Loewner matrix $\mathbb{L}_\sigma$ associated to the tangential interpolation data (3.33) are defined as follows:

Definition 3.13 (The Loewner pencil [Mayo and Antoulas, 2007]). *The* Loewner matrix $\mathbb{L}$ *associated to the data* (3.33) *is defined as*

$$\mathbb{L} = \begin{bmatrix} \dfrac{\mathbf{v}_1^T\mathbf{r}_1 - \mathbf{l}_1^T\mathbf{w}_1}{\mu_1 - \lambda_1} & \cdots & \dfrac{\mathbf{v}_1^T\mathbf{r}_n - \mathbf{l}_1^T\mathbf{w}_n}{\mu_1 - \lambda_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\mathbf{v}_n^T\mathbf{r}_1 - \mathbf{l}_n^T\mathbf{w}_1}{\mu_n - \lambda_1} & \cdots & \dfrac{\mathbf{v}_n^T\mathbf{r}_n - \mathbf{l}_n^T\mathbf{w}_n}{\mu_n - \lambda_n} \end{bmatrix} \in \mathbb{C}^{n\times n}. \tag{3.36}$$

The shifted Loewner matrix $\mathbb{L}_\sigma$ *associated to the data* (3.33) *is defined as*

$$\mathbb{L}_\sigma = \begin{bmatrix} \dfrac{\mu_1\mathbf{v}_1^T\mathbf{r}_1 - \lambda_1\mathbf{l}_1^T\mathbf{w}_1}{\mu_1 - \lambda_1} & \cdots & \dfrac{\mu_1\mathbf{v}_1^T\mathbf{r}_n - \lambda_n\mathbf{l}_1^T\mathbf{w}_n}{\mu_1 - \lambda_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\mu_n\mathbf{v}_n^T\mathbf{r}_1 - \lambda_1\mathbf{l}_n^T\mathbf{w}_1}{\mu_n - \lambda_1} & \cdots & \dfrac{\mu_n\mathbf{v}_n^T\mathbf{r}_n - \lambda_n\mathbf{l}_n^T\mathbf{w}_n}{\mu_n - \lambda_n} \end{bmatrix} \in \mathbb{C}^{n\times n}. \tag{3.37}$$

The pair $(\mathbb{L}, \mathbb{L}_\sigma)$ *is called the Loewner pencil associated to the tangential interpolation data* (3.33).

For instance, it is assumed that the interpolation points satisfy $\lambda_j \neq \mu_i$, for $i, j = 1, \ldots, n$, in order to ensure that the Loewner and shifted Loewner matrices are well-defined. This restriction will be later dropped in Theorem 3.16. These matrices are the key tool to solve the data tangential interpolation Problem 3.12.

First, notice that these two matrices satisfy the following Sylvester equations:

$$M\mathbb{L} - \mathbb{L}\Lambda = VR - LW, \tag{3.38}$$

and

$$M\mathbb{L}_\sigma - \mathbb{L}_\sigma\Lambda = MVR - LW\Lambda, \tag{3.39}$$

where M, Λ, V, W, R and L are defined as in (3.35).

Now, let us assume that the data from (3.33) are sampled from a system whose transfer function is

$$\hat{\mathbf{H}} = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B} \in \mathbb{C}^{n_y\times n_u},$$

where $\hat{A}, \hat{E} \in \mathbb{R}^{n\times n}$. For model $\hat{\mathbf{H}}$, we define the *generalized tangential controllability* and *observability matrices* associated to the tangential interpolation data (3.33) as:

$$\left.\begin{aligned} \mathcal{O}_n &= \begin{bmatrix} \mathbf{l}_1^T\hat{C}(\mu_1\hat{E} - \hat{A})^{-1} \\ \vdots \\ \mathbf{l}_n^T\hat{C}(\mu_n\hat{E} - \hat{A})^{-1} \end{bmatrix} \\ \mathcal{R}_n &= \begin{bmatrix} (\lambda_1\hat{E} - \hat{A})^{-1}\hat{B}\mathbf{r}_1 & \ldots & (\lambda_n\hat{E} - \hat{A})^{-1}\hat{B}\mathbf{r}_n \end{bmatrix} \end{aligned}\right\}. \tag{3.40}$$

Hence, the following matrix factorization result follows,

$$\begin{aligned}[\mathbb{L}]_{ji} &= \frac{\mathbf{v}_j^T\mathbf{r_i} - \mathbf{l}_j^T\mathbf{w}_i}{\mu_j - \lambda_i} \\ &= -\mathbf{l_j^T}\hat{C}(\mu_j\hat{E} - \hat{A})^{-1}\hat{E}(\lambda_i\hat{E} - \hat{A})^{-1}\hat{B}\mathbf{r_i},\end{aligned}$$

and, similarly,

$$\begin{aligned}[\mathbb{L}_\sigma]_{ji} &= \frac{\mu_j\mathbf{v}_j^T\mathbf{r_i} - \mathbf{l}_j^T\mathbf{w}_i\lambda_i}{\mu_j - \lambda_i} \\ &= -\mathbf{l_j}^T\hat{C}(\mu_j\hat{E} - \hat{A})^{-1}\hat{A}(\lambda_i\hat{E} - \hat{A})^{-1}\hat{B}\mathbf{r_i},\end{aligned}$$

leading to

$$\mathbb{L} = -\mathcal{O}_n\hat{E}\mathcal{R}_n \quad \text{and} \quad \mathbb{L}_\sigma = -\mathcal{O}_n\hat{A}\mathcal{R}_n. \tag{3.41}$$

Hence, the Loewner matrix and the shifted Loewner matrix can be factorized using the generalized controllability and observability grammians. This leads to the first important result about the Loewner pencil.

Lemma 3.14 (Complexity of interpolant [Mayo and Antoulas, 2007]). *Given tangential samples of a minimal descriptor system $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$, construct the matrices $\mathbb{L}$ and $\mathbb{L}_\sigma$ as in Definition 3.13. Assuming that we have enough samples, and the left and right directions $\mathbf{l}_i$ and $\mathbf{r}_i$ are chosen such that $\mathcal{O}_n$ and $\mathcal{R}_n$ have full rank, then the following statements hold:*

1. ***rank***$(\mathbb{L}) =$ ***rank***$(\hat{E}) =$ McMillan degree of the system $\hat{\mathbf{H}}$[a];
2. ***rank***$(\mathbb{L}_\sigma) =$ ***rank***$(\hat{A})$.

[a] if the system is strictly proper, its McMillan degree is the order of one of its minimal realizations.

Lemma 3.14 states that the Loewner matrix $\mathbb{L}$ encodes the McMillan degree of a system. Hence, the rank of the Loewner matrix somewhere encodes the number of states necessary to describe the underlying system.

Let us now state the solution for Problem 3.12 based on the Loewner pencil.

Theorem 3.15. *Let $\det(\sigma\mathbb{L} - \mathbb{L}_\sigma) \neq 0$ for all $\sigma \in \{\lambda_i\} \cup \{\mu_i\}$. Then, the n-th reduced-order model*

$$\hat{\mathbf{H}} : \begin{cases} -\mathbb{L}\dot{\hat{\mathbf{x}}}(t) &= -\mathbb{L}_\sigma\hat{\mathbf{x}}(t) + V\mathbf{u}(t) \\ \hat{\mathbf{y}}(t) &= W\hat{\mathbf{x}}(t) \end{cases}, \tag{3.42}$$

is a minimal realization of an interpolant of the data, i.e., *its transfer function*

$$\hat{\mathbf{H}}(s) = W(\mathbb{L}_\sigma - s\mathbb{L})^{-1}V,$$

satisfies the bitangential interpolation conditions from Problem 3.12.

Proof. Multiplying equation (3.38) by s and subtracting it from the equation (3.39) we obtain

$$(\mathbb{L}_\sigma - s\mathbb{L})\Lambda - M(\mathbb{L}_\sigma - s\mathbb{L}) = LW(\Lambda - sI_n) - (M - sI_n)VR.$$

Then, by multiplying this equation by $\mathbf{e}_i$ on the right and setting $s = \lambda_i$, we have

$$\begin{aligned}(\lambda_i I_n - M)(\mathbb{L}_\sigma - \lambda_i\mathbb{L})\mathbf{e}_i &= (\lambda_i I_n - M)V\mathbf{r}_i \\ (\mathbb{L}_\sigma - \lambda_i\mathbb{L})\mathbf{e}_i &= V\mathbf{r}_i.\end{aligned}$$

Since, by hypothesis, $(\mathbb{L}_\sigma - \lambda_i \mathbb{L})$ is non singular, we deduce

$$\mathbf{w}_i = W\mathbf{e}_i = W(\mathbb{L}_\sigma - \lambda_i \mathbb{L})^{-1} V \mathbf{r}_i = \hat{\mathbf{H}}(\lambda_i)\mathbf{r}_i.$$

The left tangential interpolation condition follows similarly. □

Theorem 3.15 gives a solution to the tangential interpolation data problem 3.12. In the context of model approximation, the advantage is that it *does not require* that the model $\mathbf{G}$ is given by a finite dimensional realization. Hence, given a transfer function $\mathbf{G}$, irrational or not, we are able to construct a model approximation $\mathbf{H} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ satisfying some tangential interpolation conditions.

This is not the case of the projection-based approach of Theorem 3.10. Indeed, the projection-based model approximation framework requires that $\mathbf{G}$ has a finite dimensional realization, *e.g.*,

$$\mathbf{G}(s) = C(sE - A)^{-1}B,$$

and irrational transfer function cannot be handled by this approach directly, requiring a discretization step of the underlying infinite dimensional system *a priori*. Therefore, the Loewner framework enables to perform model approximation of infinite dimensional models without discretization.

The following theorem enables to extend Theorem 3.15 to the case where the right and left interpolation points are identical, *i.e.*, $\lambda_i = \mu_i$, for $i = 1, \dots n$.

Theorem 3.16. *Let $\mathbf{G}(s) \in \mathbb{C}^{n_u \times n_y}$ be a transfer functions. If*

$$(\mathbb{L})_{i,j} := \begin{cases} \dfrac{\mathbf{l}_i^T(\mathbf{G}(\sigma_i) - \mathbf{G}(\sigma_j)\mathbf{r}_j}{\sigma_i - \sigma_j} & \text{if} \quad i \neq j \\ \mathbf{l}_i^T \mathbf{G}'(\sigma_i)\mathbf{r}_i & \text{if} \quad i = j \end{cases}$$

$$(\mathbb{L}_\sigma)_{i,j} := \begin{cases} \dfrac{\mathbf{l}_i^T(\sigma_i\mathbf{G}(\sigma_i) - \sigma_j\mathbf{G}(\sigma_j)\mathbf{r}_j}{\sigma_i - \sigma_j} & \text{if} \quad i \neq j \\ \mathbf{l}_i^T [s\mathbf{G}'(s)]'|_{s=\sigma_i}\mathbf{r}_i & \text{if} \quad i = j \end{cases}$$

and

$$W := \begin{bmatrix} \mathbf{G}(\sigma_1)\mathbf{r}_1 & \dots & \mathbf{G}(\sigma_n)\mathbf{r}_n \end{bmatrix}, \quad V := \begin{bmatrix} \mathbf{l}_1^T\mathbf{G}(\sigma_1) \\ \dots \\ \mathbf{l}_n^T\mathbf{G}(\sigma_n) \end{bmatrix}, \tag{3.43}$$

then, $\hat{\mathbf{H}}(s) = W(\mathbb{L}_\sigma - s\mathbb{L})V$ satisfies the bitangential Hermite interpolation conditions as follows:

$$\begin{cases} \hat{\mathbf{H}}(\sigma_i)\mathbf{r}_i &= \mathbf{G}(\sigma_i)\mathbf{r}_i \\ \mathbf{l}_i^T\hat{\mathbf{H}}(\sigma_i) &= \mathbf{l}_i^T\mathbf{G}(\sigma_i) \\ \mathbf{l}_i^T\hat{\mathbf{H}}'(\sigma_i)\mathbf{r}_i &= \mathbf{l}_i^T\mathbf{G}'(\sigma_i)\mathbf{r}_i \end{cases} \tag{3.44}$$

for $i = 1, \dots, n$.

Theorem 3.16 is equivalent to Theorem 3.10 when the right and left interpolation points are equal. Hence, Theorem 3.16 enables to build a reduced-order model $\hat{\mathbf{H}}$ satisfying the the bitangential Hermite conditions as in Problem 3.7. This result will be particularly interesting

for the $\mathcal{H}_2$ optimal model approximation algorithm (**TF-IRKA** presented in Chapter 4). Notice that it is also possible to interpolate higher derivatives and the approach is described in [Mayo and Antoulas, 2007].

Now, we briefly present a procedure, also from [Mayo and Antoulas, 2007], to handle the case where more data than necessary is provided.

Theorem 3.17 (Redundant data [Mayo and Antoulas, 2007]). *Given tangential data as in Problem 3.12, let $\mathbb{L}$ and $\mathbb{L}_\sigma$ be the Loewner pencil. Then, if*

$$\textbf{\textit{rank}}\begin{bmatrix} x\mathbb{L} - \mathbb{L}_\sigma \end{bmatrix} = \textbf{\textit{rank}}\begin{bmatrix} \mathbb{L} & \mathbb{L}_\sigma \end{bmatrix} = \textbf{\textit{rank}}\begin{bmatrix} \mathbb{L} \\ \mathbb{L}_\sigma \end{bmatrix} = r, \tag{3.45}$$

for all $x = \{\lambda_i\} \cup \{\mu_i\}$, $i = 1, \ldots, n$. Consider the shorts SVDs

$$\begin{bmatrix} \mathbb{L} & \mathbb{L}_\sigma \end{bmatrix} = Y \Sigma_l \tilde{X}^* \text{ and } \begin{bmatrix} \mathbb{L} \\ \mathbb{L}_\sigma \end{bmatrix} = \tilde{Y} \Sigma_r X^*, \tag{3.46}$$

where $\Sigma_l, \Sigma_r = \mathbb{R}^{r \times r}$, $Y \in \mathbb{C}^{n \times r}$ and $X \in \mathbb{C}^{n \times r}$. Then, the system $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$, given by

$$\hat{E} = -Y^* \mathbb{L} X, \quad \hat{A} = -Y^* \mathbb{L}_\sigma X, \quad \hat{B} = Y^* V, \quad \hat{C} = W X, \tag{3.47}$$

is a decriptor system of an (approximate) interpolant of the data with McMillan degree $r = \textbf{\textit{rank}}(\mathbb{L})$.

For realistic applications, Theorem 3.17 is very useful to derive reduced-order models. In practice, one should use the numerical rank when verifying condition (3.45) and that is the reason why one has an approximate intepolant.

Some extensions of this framework were developed. Here is a non-exhaustive list of extensions and relative works:

- **Parametric systems:** in [Ionita and Antoulas, 2014a], authors extend the Loewner framework for parametric systems. In [Ionita and Antoulas, 2014b], authors present a study-case on the model approximation using Loewner framework and reduced-basis.
- **Time-delay systems:** in [Pontes Duff et al., 2015a] and [Schulze and Unger, 2015], the Loewner framework for a special class of time-delay systems is presented. They will be discussed later in Chapter 6.
- **Non-linear systems:** the Loewner framework has been used as well to treat non-linear model. See [Antoulas and Gosea, 2015] and [Gosea and Antoulas, 2015] for more details.

Conclusion

In this chapter, a non-exhaustive state-of-the-art on model approximation techniques was presented focusing on two family of methods: (i) the *projection-based* and (ii) the *data-based model approximation*. A great variety of model approximation methods can be described in this projection-based framework. Two projection-based model approximation methods are described: (i) model approximation by truncation in 3.2 and (ii) model approximation by tangential interpolation in 3.3. These techniques can be successfully applied in a wide range of applications. The major drawback is that they require that both the full-order model **G** and the reduced-order

model $\hat{\mathbf{H}}$ to be represented by a finite-dimensional realization. Hence, they are not directly applicable to infinite dimensional systems or irrational transfer functions and they do not allow the reduced-order model to have any particular structure. However, in Section 3.3, a projection-based method, able to preserve peculiar structure, is presented. It enables model approximation for a larger family of models. Finally, Section 3.4 presents the Loewner framework by performing data-driven tangential interpolation and, in the context of model approximation, it can be used to construct reduced-order models directly from the evaluation of the original model transfer function.

In this chapter, we do not discuss the choice of the interpolation points σ_k and tangential directions $\mathbf{r}_k$ and $\mathbf{l}_k$ for the interpolation based methods. 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.

Part II

Model approximation by structured time-delay reduced order models

Chapter 4

Optimal $\mathcal{H}_2$ model approximation

In this chapter, the standard optimal model approximation problem with respect to the $\mathcal{H}_2$ norm is introduced, motivated and contextualized in the literature. Following the philosophy of [Gugercin et al., 2008], the objective of this chapter is twofold. On one hand, to introduce the theoretical results which are the background and the main inspiration for the new results presented in this thesis, especially the results presented in Chapters 5 and 7. On the other hand, to revisit some of these results, *e.g.*, the $\mathcal{H}_2$ inner product computation based on a pole/residue decomposition.

The chapter is organized as follows. In Section 4.1, the use of the $\mathcal{H}_2$ norm is motivated, some historical literature results are also detailed and the problem is mathematically introduced in Problem 4.1. Section 4.2 presents the spectral $\mathcal{H}_2$ inner product formulation. Section 4.3, the necessary optimality conditions of Problem 4.1 are derived. Finally, in Section 4.4, two fixed-point algorithms are presented enabling the construction of a reduced-order model. We believe that some demonstrations are original and provide some new insights about the $\mathcal{H}_2$ approximation problem.

Contents

4.1 Motivation and problem statement

Motivation

The problem of determining "the best approximation" is meaningless if there is no metric associated. Indeed, different measures can lead to completely different optimal approximations. Consequently, when solving approximation problems, one should first of all precise how the approximation error is quantified.

In the context of model approximation, the well-defined system norms (Chapter 2) are largely used to quantify the quality of an approximation. Thus, a reduced order model $\hat{\mathbf{H}}$ will be considered to be the best approximation of $\mathbf{G}$ with respect to the norm $\mathcal{H}$ if $\|\mathbf{G}-\hat{\mathbf{H}}\|_{\mathcal{H}}$ is as small as possible. More precisely, for both $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms, the inequalities (2.11) and (2.12) from Chapter 2 allow to bound the output for a given input as follows :

- $\mathcal{H}_2$ **norm:** let $\mathbf{G} \in \mathcal{H}_2$ and suppose that $\hat{\mathbf{H}} \in \mathcal{H}_2$ is a good approximation of $\mathbf{G}$ in the $\mathcal{H}_2$ norm, *i.e.*, $\|\mathbf{G}-\hat{\mathbf{H}}\|_{\mathcal{H}_2} \ll 1$. In addition, assume that both models $\mathbf{G}$ and $\hat{\mathbf{H}}$ are subject to the same input $\mathbf{u}(t)$ and that $\|\mathbf{u}\|_{L_2} = 1$ has provided the outputs $\mathbf{y}$ and $\hat{\mathbf{y}}$. Then, by applying (2.11), the following estimate is obtained

$$\|\mathbf{y}(t)-\hat{\mathbf{y}}(t)\|_{L_\infty} \leq \|\mathbf{G}-\hat{\mathbf{H}}\|_{\mathcal{H}_2}\|\mathbf{u}\|_{L_2} \ll \|\mathbf{G}\|_{\mathcal{H}_2}.$$

  Hence, a good $\mathcal{H}_2$ approximation yields good L_∞ output approximation.

- $\mathcal{H}_\infty$ **norm:** in a similar way, if $\mathbf{G} \in \mathcal{H}_\infty$ and $\hat{\mathbf{H}} \in \mathcal{H}_\infty$ and $\|\mathbf{G}-\hat{\mathbf{H}}\|_{\mathcal{H}_\infty} \ll 1$, then, by applying (2.12), the following result is obtained

$$\|\mathbf{y}(t)-\hat{\mathbf{y}}(t)\|_{L_2} \leq \|\mathbf{G}-\hat{\mathbf{H}}\|_{\mathcal{H}_\infty}\|\mathbf{u}\|_{L_2} \ll 1.$$

  Hence, a good $\mathcal{H}_\infty$ approximation yields to a good L_2 output approximation.

It is worth recalling here that the $\mathcal{H}_\infty$ norm is the L_2-L_2 induced norm, *i.e.*, $\|\mathbf{G}\|_{\mathcal{H}_\infty} = \sup_{\mathbf{u}\in L_2} \frac{\|\mathbf{G}\mathbf{u}\|_{L_2}}{\|\mathbf{u}\|_{L_2}}$. However, $\mathcal{H}_2$ norm is not an induced system norm since it depends on the Frobenious matrix norm, which is not an induced matrix norm itself. Indeed, this norm is an unitarily invariant norm[1] and unitarily invariant norms are not induced, with the exception of the 2-norm[2] (see [Chellaboina and Haddad, 1995]).

Metric choice

One common system norm is the $\mathcal{H}_\infty$ norm (definition in Chapter 2), which measures the worst-case scenario. Hence, it gives valuable information about the worst amplification and it is well adapted in the context of control design, more specifically in robust control theory (see [Zhou and Doyle, 1998; Green and Limebeer, 2012]). However, a severe drawback of the $\mathcal{H}_\infty$ norm is that its computation is itself a hard optimization problem. This is the main reason why

[1] The norm $\|.\|$ is said to be an unitarily invariant norm if, for $U \in \mathbb{C}^{n\times n}$, $V \in \mathbb{C}^{m\times m}$, unitary matrices, then the following is satisfied : $\|UAV\| = \|A\|$, for all $A \in \mathbb{C}^{n\times m}$.

[2] The 2-norm, also now as the spectral norm, for $A \in \mathbb{C}^{n_y\times n_u}$, is defined as

$$\|A\|_2 = \sup_{u\in\mathbb{C}^{n_u}} \frac{\|Au\|_2}{\|u\|_2} = \sigma_{max}(A).$$

it is not adapted in the large-scale case (see [Antoulas and Astolfi, 2002] for some theoretical aspects about $\mathcal{H}_\infty$ approximation). Another measure that is particularly interesting in view of control design is the ν-gap metric, which quantifies the closed-loop behavior of a given model (see [Georgiou and Smith, 1990; Vinnicombe, 1993] for the definition and the theoretical results and [Cantoni, 2001; Sootla, 2014] for model approximation based on this metric). Similarly, ν-gap model approximation is generally considered to be computationally unfeasible, based on the same arguments provided for the $\mathcal{H}_\infty$ case.

In this thesis, the $\mathcal{H}_2$ norm is the one chosen to measure the quality of an approximation. Indeed, this norm has several properties that justify its use, *e.g.*,

- ➤ the $\mathcal{H}_2$ norm is numerically accessible even in the large-scale setting.
- ➤ it can be analytically characterized either as a solution of a Lyapunov equation, or based on the pole/residue decomposition of a transfer function.
- ➤ $\mathcal{H}_2$ space is a Hilbert space. Thus, it is equipped with an inner product enabling orthogonal projections. This geometric notion plays a very important role in $\mathcal{H}_2$ model approximation and will be largely explored from now on in the work.
- ➤ the approximation problem is a root mean-square problem [Riggs and Edgar, 1974] and it has a nice impulse response energy interpretation.

For some practical applications, the $\mathcal{H}_2$ norm might not be desirable (as an example, in the context of robust control design). Nevertheless, experience has shown that $\mathcal{H}_2$ reduced order model yields to a small error in the $\mathcal{H}_\infty$ norm as well (see [Poussot-Vassal et al., 2013; Vuillemin et al., 2016]).

Problem Statement

The standard $\mathcal{H}_2$ model approximation problem is stated as follows :

Problem 4.1 (Optimal $\mathcal{H}_2$ model approximation problem). *Given a strictly proper LTI model $\mathbf{G} \in \mathcal{H}_2$[a] and an order $n \in \mathbb{N}^*$, find a finite dimensional model $\hat{\mathbf{H}}^\star \in \mathcal{H}_2$ which minimizes the $\mathcal{H}_2$ norm of the approximation error,* i.e.

$$\hat{\mathbf{H}}^\star = \underset{\substack{\hat{\mathbf{H}} \in \mathcal{H}_2 \\ \dim(\hat{\mathbf{H}}) \leq n}}{\boldsymbol{arg\ min}} \ \|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{H}_2} \tag{4.1}$$

where $\mathcal{J}_{\mathcal{H}_2}(\hat{\mathbf{H}}) = \|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{H}_2}$.

[a] The order of $\mathbf{G}$ was not precised on purpose. In general, $\mathbf{G}$ is supposed to be an infinite dimensional model or a large-scale one. In the latter case, we denote the order of $\mathbf{G}$ by N and it is supposed that $n \ll N$.

Problem 4.1 has been widely studied. Finding a global solution of Problem 4.1 is a hard task and, so far, a common approach consist in finding a reduced order model which satisfies the first-order $\mathcal{H}_2$ necessary optimality conditions. Major results available in the literature are devoted to the case where $\mathbf{G}$ has a finite dimensional realization, *i.e.*, $\mathbf{G} = (E, A, B, C)$. A (non-exhaustive) list of references concerning this case is given hereafter:

- ❖ **Interpolation-based optimality conditions:** in the context of model approximation, $\mathcal{H}_2$ necessary first-order optimality conditions for SISO models were derived in [Meier III

and Luenberger, 1967] as interpolation conditions. Later, [Van Dooren et al., 2008] derived the optimality conditions as bitangential Hermite interpolation ones. In addition, [Van Dooren et al., 2010] characterizes the stationary points of Problem 4.1 when the reduced order model has higher order poles.

- ❖ **Lyapunov-based optimality conditions:** first-order necessary optimality conditions were as well derived in the context of Lyapunov equations [Wilson, 1970] for MIMO models. Later, those conditions were expressed in the projection framework in [Hyland and Bernstein, 1985]. Finally, frequency weighted model reduction was tackled by [Halevi, 1992] using the Lyapunov and the projection-based frameworks.

- ❖ **Iterative algorithms:** the well-known Iterative Rational Krylov Algorithm (**IRKA**) was proposed in [Gugercin et al., 2006] and [Gugercin et al., 2008]. This algorithm uses the Krylov-spaces methods to find a model satisfying the $\mathcal{H}_2$ optimality conditions. Later, [Beattie and Gugercin, 2012] proposes an algorithm based on the Loewner framework enabling the approximation of irrational transfer functions. These algorithms will be recalled later in this chapter.

- ❖ **Trust region methods:** a second order gradient algorithm was first proposed by [Bryson and Carrier, 1990]. Later, a trust-region method for optimal $\mathcal{H}_2$ model reduction problems was proposed by [Beattie and Gugercin, 2009b].

- ❖ **Approaches grounded on non-linear optimization:** an interpolation-based algorithm was proposed in [Lepschy et al., 1991]. A two-step algorithm based on the numerator-denominator parametrization of $\mathbf{G}$ was developed in [Spanos et al., 1992]. The $\mathcal{H}_2$ model approximation problem was formulated as an unconstrained optimization one over a manifold in [Yan and Lam, 1999] and some gradient-based algorithms were proposed. A descent algorithm using Krylov-spaces was presented in [Beattie and Gugercin, 2007]. More recently, non-linear optimization is used to solve model approximation and control design problems related to the frequency-limited $\mathcal{H}_2$ norm in [Petersson, 2013] and [Vuillemin et al., 2014a].

In addition, [Flagg et al., 2013] combines interpolatory techniques and $\mathcal{H}_\infty$ optimatization of the direct feedthrough term D in order to find good $\mathcal{H}_\infty$ approximation of SISO models.

In the context of infinite dimensional models, the recent article [Opmeer, 2015] gives a great survey of $\mathcal{H}_2$ optimal approximation methods for systems represented by non-rational transfer functions. It highlights that several methods proposed in the $\mathcal{H}_2$ literature for finite dimensional models can be applied with some minor modifications to the case of irrational transfer function.

In the context of discrete-time models, [Baratchart, 1986] has presented some theoretical results including the minimum existence, and the articles [Baratchart et al., 1991; Fulcheri and Olivi, 1998; Marmorat et al., 2002] present different algorithms to solve the problem using a dedicated parametrization. It should be notice that there is an isometry from $\mathcal{H}_2(\mathbb{C}^+)$ to the orthogonal of $\mathcal{H}_2(\mathbb{D})$ (Hardy space for discrete-time systems) so that methods used for discrete-time systems can also be used for continuous-time and vice versa (see [Olivi et al., 2013]). The case of discrete-time models is not treated in this thesis.

Hilbert projection and non-convexity

It is worth noticing that the optimization criterion $\mathcal{J}_{\mathcal{H}_2}(\hat{\mathbf{H}})$ is a convex function from $\mathcal{H}_2$ to $\mathbb{R}_+$. This can be easily verified, for $t \in [0,1]$, as follows

$$\begin{aligned}\mathcal{J}_{\mathcal{H}_2}(t\hat{\mathbf{H}}_1 + (1-t)\hat{\mathbf{H}}_2) &= \|\mathbf{G} - t\hat{\mathbf{H}}_1 - (1-t)\hat{\mathbf{H}}_2\|_{\mathcal{H}_2} \\ &= \|t(\mathbf{G} - \hat{\mathbf{H}}_1) + (1-t)(\mathbf{G} - \hat{\mathbf{H}}_2)\|_{\mathcal{H}_2} \\ &\leq \|t(\mathbf{G} - \hat{\mathbf{H}}_1)\|_{\mathcal{H}_2} + \|(1-t)(\mathbf{G} - \hat{\mathbf{H}}_2)\|_{\mathcal{H}_2} \\ &\leq t\mathcal{J}_{\mathcal{H}_2}(\hat{\mathbf{H}}_1) + (1-t)\mathcal{J}_{\mathcal{H}_2}(\hat{\mathbf{H}}_2).\end{aligned}$$

Additionally, since $\mathcal{H}_2$ is a Hilbert space, the non-familiar reader might think that Problem 4.1 could be solved by projection over a given basis, using the Hilbert projection theorem (see [Kreyszig, 1989, Chapter 3]). However, even if the optimization criterion $\mathcal{J}_{\mathcal{H}_2}(\hat{\mathbf{H}})$ is a convex function, Problem 4.1 is non-convex. Indeed, the constraints of Problem 4.1 is defined by the set

$$\mathcal{M} = \{\hat{\mathbf{H}} \in \mathcal{H}_2, order(\hat{\mathbf{H}}) \leq n\}.$$

The set $\mathcal{M}$ is non-convex[3] which justifies the non-convex nature of Problem 4.1.

As a consequence, $\mathcal{M}$ does not define a vector space and one cannot apply projection for solving Problem 4.1. We will see in Subsection 4.3.1 that if we assume that the reduced order model has assigned poles, then the optimization set becomes convex and the problem can be solved using the Hilbert projection theorem.

This chapter is dedicated to study the Problem 4.1. The procedure followed here can be summarized in three steps:

1. Spectral characterization of the $\mathcal{H}_2$ inner product, presented in Section 4.2.
2. Derivation of the $\mathcal{H}_2$ necessary optimality conditions of Problem 4.1 using the $\mathcal{H}_2$ inner product characterization. This is the subject of Section 4.3.
3. Recall a fixed-point algorithm based on the $\mathcal{H}_2$ necessary optimality conditions in order to construct reduced order models in Section 4.4.

In the following section, a spectral characterization of the $\mathcal{H}_2$ inner product is provided.

4.2 $\mathcal{H}_2$ inner product properties

4.2.1 Spectral $\mathcal{H}_2$ inner product

In Section 2.1 of Chapter 3, the notions of the $\mathcal{H}_2$ inner product and norm were introduced. Hence, for two models $\mathbf{G}$ and $\mathbf{H}$ in $\mathcal{H}_2$, their $\mathcal{H}_2$ inner product is given by

$$\begin{aligned}\langle \mathbf{G}, \mathbf{H} \rangle_{\mathcal{H}_2} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{trace}\Big(\overline{\mathbf{G}(i\omega)}\mathbf{H}(i\omega)^T\Big)d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \langle \mathbf{G}(i\omega), \mathbf{H}(i\omega)\rangle_F d\omega.\end{aligned}$$

In this section, a spectral characterization of this inner product will be presented. Firstly, let us start with a simple computation formula.

[3] The sum of two systems of order n is in general a system of order $2n$.

Lemma 4.2 (First order $\mathcal{H}_2$ inner product). *Let $\mathbf{G} \in \mathcal{H}_2$ be a strictly proper real model, $\hat{\phi} \in \mathbb{C}^{n_y \times n_u}$ and $\hat{\lambda} \in \mathbb{C}^-$. Then*

$$\left\langle \mathbf{G}, \frac{\hat{\phi}}{s - \hat{\lambda}} \right\rangle_{\mathcal{H}_2} = \text{trace}(\mathbf{G}(-\hat{\lambda})\hat{\phi}^T).$$

Moreover, if $\hat{\phi} = \hat{\mathbf{c}}\hat{\mathbf{b}}^T$, $\hat{\mathbf{c}} \in \mathbb{C}^{n_y}$ and $\hat{\mathbf{b}} \in \mathbb{C}^{n_u}$, then

$$\left\langle \mathbf{G}, \frac{\hat{\mathbf{c}}\hat{\mathbf{b}}^T}{s - \hat{\lambda}} \right\rangle_{\mathcal{H}_2} = \hat{\mathbf{c}}^T \mathbf{G}(-\hat{\lambda})\hat{\mathbf{b}}. \tag{4.2}$$

Proof. It is an implication of the Cauchy's residues theorem (see [Antoulas, 2005] and [Gugercin et al., 2008]). Observing that $\overline{\mathbf{G}(i\omega)} = \mathbf{G}(-i\omega)$ and that the only stable pole of the complex function $\text{trace}(\mathbf{G}(-s)\frac{\hat{\phi}^T}{s-\hat{\lambda}})$ is $\hat{\lambda} \in \mathbb{C}^-$ (the poles of $\mathbf{G}(-s)$ are all unstable). Let us consider the following semi-circular contour Γ_C located in the left half plane such that:

$$\Gamma_C = \Gamma_I \cup \Gamma_R,$$

with:

$$\begin{cases} \Gamma_I = \{s \in \mathbb{C} / s = i\omega \text{ and } \omega \in [-R; R],\ R \in \mathbb{R}_+\} \\ \Gamma_R = \{s \in \mathbb{C} / s = Re^{i\theta} \text{ where } \theta \in [\pi/2; 3\pi/2]\} \end{cases}.$$

This contour is sketched on Figure 4.1.

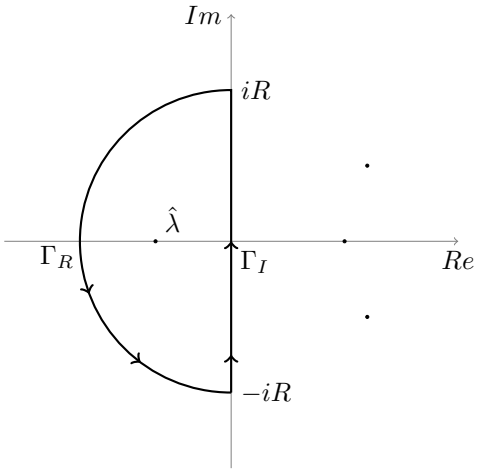

Figure 4.1: Complex integral path

For a sufficient large radius value R, the Γ_C contour will contain the pole $\hat{\lambda}$ only. Thus, by

applying the residues theorem, it follows that:

$$\begin{aligned}\left\langle \mathbf{G}, \frac{\hat{\phi}}{s-\hat{\lambda}} \right\rangle_{\mathcal{H}_2} &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \operatorname{trace}\left(\mathbf{G}(-i\omega)\frac{\hat{\phi}^T}{i\omega-\hat{\lambda}}\right) d\omega \\ &= \lim_{R\to+\infty}\frac{1}{2i\pi}\int_{\Gamma_C} \operatorname{trace}\left(\mathbf{G}(-s)\frac{\hat{\phi}^T}{s-\hat{\lambda}}\right) ds \\ &= \operatorname{trace}\big(\operatorname{Res}(\mathbf{G}(-s)\frac{\hat{\phi}^T}{s-\hat{\lambda}}, \hat{\lambda})\big) \\ &= \operatorname{trace}(\mathbf{G}(-\hat{\lambda})\hat{\phi}^T).\end{aligned}$$

where Res(.) denotes the residue operator. The second equality line holds true since, when $R \to +\infty$:

$$\left| \int_{\Gamma_R} \underbrace{\operatorname{trace}\left(\mathbf{G}(-s)\frac{\hat{\phi}^T}{s-\hat{\lambda}}\right)}_{\langle \overline{\mathbf{G}(-s)}, \frac{\hat{\phi}}{s-\hat{\lambda}} \rangle_F} ds \right| \leq \int_{\Gamma_R} \left\| \frac{\hat{\phi}^T}{s-\hat{\lambda}} \right\|_F \left\| \mathbf{G}(-s) \right\|_F ds \to 0^+,$$

which concludes the first part of the proof. Now, suppose that $\hat{\phi} = \hat{\mathbf{c}}\hat{\mathbf{b}}^T$. Then, from the inner product formula

$$\begin{aligned}\left\langle \mathbf{G}, \frac{\hat{\mathbf{c}}\hat{\mathbf{b}}^T}{s-\hat{\lambda}} \right\rangle_{\mathcal{H}_2} &= \operatorname{trace}\left(\mathbf{G}(-\hat{\lambda})(\hat{\mathbf{c}}\hat{\mathbf{b}}^T)^T\right) \\ &= \operatorname{trace}\left(\mathbf{G}(-\hat{\lambda})\hat{\mathbf{b}}\hat{\mathbf{c}}^T\right) \\ &= \hat{\mathbf{c}}^T\mathbf{G}(-\hat{\lambda})\hat{\mathbf{b}}.\end{aligned}$$

which concludes the proof. □

Lemma 4.2 provides a simple formula which characterizes the $\mathcal{H}_2$ inner product when one of the models is of order one. The following remark enables to turn a MIMO $\mathcal{H}_2$ inner product into a SISO $\mathcal{H}_2$ inner product.

Remark 4.3. *Let $\mathbf{G} \in \mathcal{H}_2$ be a MIMO model with n_u inputs and n_y outputs, $\hat{\mathbf{c}} \in \mathbb{C}^{n_y}$ and $\hat{\mathbf{b}} \in \mathbb{C}^{n_u}$ be two tangential directions and $\lambda \in \mathbb{C}^-$ be a pole. Then $\hat{\mathbf{H}}(s) = \frac{\hat{\mathbf{c}}\hat{\mathbf{b}}^T}{s-\hat{\lambda}} \in \mathbb{C}^{n_y \times n_u}$ is also a model with n_u inputs and n_y outputs. Moreover,*

$$\left\langle \mathbf{G}, \frac{\hat{\mathbf{c}}\hat{\mathbf{b}}^T}{s-\hat{\lambda}} \right\rangle_{\mathcal{H}_2} = \hat{\mathbf{c}}^T\mathbf{G}(-\hat{\lambda})\hat{\mathbf{b}} = \left\langle \mathbf{G}_{SISO}, \frac{1}{s-\hat{\lambda}} \right\rangle_{\mathcal{H}_2}, \tag{4.3}$$

where $\mathbf{G}_{SISO}$, is a SISO model defined by

$$\mathbf{G}_{SISO}(s) = \hat{\mathbf{c}}^T\mathbf{G}(s)\hat{\mathbf{b}} \in \mathbb{C}.$$

*In order words, the **tangential directions** $\hat{\mathbf{c}} \in \mathbb{C}^{n_y}$ and $\hat{\mathbf{b}} \in \mathbb{C}^{n_u}$ enable us to transform the MIMO $\mathcal{H}_2$ inner product into a SISO $\mathcal{H}_2$ inner product.*

In what follows, we present an alternative way to compute the $\mathcal{H}_2$ inner product based on the pole-residue decomposition of $\hat{\mathbf{H}}(s)$.

Theorem 4.4 (Spectral $\mathcal{H}_2$ inner product expression). *Let $\mathbf{G}$, $\hat{\mathbf{H}} \in \mathcal{H}_2$ be strictly proper real models. In addition, assume that $\hat{\mathbf{H}}$ has semi-simple poles, that is, it can be expressed by a pole-residue decomposition of the form*

$$\hat{\mathbf{H}} = \sum_{k=1}^{n} \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k}.$$

Then the $\mathcal{H}_2$ inner product can be expressed as follows:

$$\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} = \sum_{k=1}^{n} \hat{\mathbf{c}}_k^T \mathbf{G}(-\hat{\lambda}_k) \hat{\mathbf{b}}_k. \tag{4.4}$$

Moreover, the $\mathcal{H}_2$ norm of $\hat{\mathbf{H}}$ is given by:

$$\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2 = \sum_{k=1}^{n} \hat{\mathbf{c}}_k^T \hat{\mathbf{H}}(-\hat{\lambda}_k) \hat{\mathbf{b}}_k. \tag{4.5}$$

Proof. The result is a straightforward application of Lemma 4.2, as follows :

$$\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} = \left\langle \mathbf{G}, \sum_{k=1}^{n} \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k} \right\rangle_{\mathcal{H}_2} \tag{4.6}$$

$$= \sum_{k=1}^{n} \langle \mathbf{G}, \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k} \rangle_{\mathcal{H}_2} \tag{4.7}$$

$$\underbrace{=}_{\text{Lemma 4.2}} \sum_{k=1}^{n} \hat{\mathbf{c}}_k^T \mathbf{G}(-\hat{\lambda}_k) \hat{\mathbf{b}}_k \tag{4.8}$$

which completes the proof. □

Theorem 4.4 enables to express the $\mathcal{H}_2$ inner product as a function of the parametrization related to $\hat{\mathbf{H}}$. Hence, if $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ and the pencil $(\hat{E}, \hat{A})$ is diagonalizable, one is able to find the pole-residue decomposition and then apply Theorem 4.4 to obtain the $\mathcal{H}_2$ inner product. Notice that the expression has no restriction over the representation of $\mathbf{G}$ and should be valid even if it is represented by an irrational transfer function. Moreover, if $\mathbf{G}$ can also be represented by a pole residue decomposition, we have a symmetric expression of the $\mathcal{H}_2$ inner product. This result is presented in the following remark.

Remark 4.5 (Symmetric representation). *If both $\mathbf{G}, \hat{\mathbf{H}} \in \mathcal{H}_2$ are real models that can be represented by the following pole-residue decompositions*

$$\mathbf{G} = \sum_{k=1}^{N} \frac{\mathbf{l}_k \mathbf{r}_k^T}{s - \mu_k} \quad \text{and} \quad \hat{\mathbf{H}} = \sum_{k=1}^{n} \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k}, \tag{4.9}$$

then, the $\mathcal{H}_2$ inner product between $\mathbf{G}$ and $\hat{\mathbf{H}}$ can be expressed as follows:

$$\begin{aligned}\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} &= \sum_{k=1}^{n} \hat{\mathbf{c}}_k^T \mathbf{G}(-\hat{\lambda}_k)\hat{\mathbf{b}}_k \\ &= \sum_{k=1}^{N} \mathbf{l}_k^T \hat{\mathbf{H}}(-\mu_k)\mathbf{r}_k.\end{aligned}$$

Remark 4.5 shows that we have some kind of symmetry in the expression of the $\mathcal{H}_2$ inner product, *i.e.*, it can be evaluated using either the poles and residues of $\hat{\mathbf{H}}$ or $\mathbf{G}$ and evaluations of $\mathbf{G}$ or $\hat{\mathbf{H}}$, respectively. Let us illustrate this with a simple example:

Example 4.6. *Let $\mathbf{G}(s) = C(sI - A)^{-1}B$ and $\hat{\mathbf{H}}(s) = \hat{C}(sI - \hat{A})^{-1}\hat{B}$, where*

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \end{bmatrix} \text{ and } C = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{l}_1 & \mathbf{l}_2 \end{bmatrix}$$

and

$$\hat{A} = \begin{bmatrix} -7 & 1 \\ 1 & -7 \end{bmatrix}, \hat{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \hat{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Firstly, the eigenvalue decomposition of $\hat{A}$ is given by

$$\hat{A}X = X \begin{bmatrix} -6 & 0 \\ 0 & -8 \end{bmatrix}, \text{ where } X = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Hence,

$$\begin{aligned} \hat{\mathbf{c}}_1 = \hat{C}X\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \hat{\mathbf{c}}_2 = \hat{C}X\mathbf{e}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \\ \hat{\mathbf{b}}_1^T = \mathbf{e}_1^T X^{-1}\hat{B} = \frac{1}{\sqrt{2}} \quad \text{and} \quad \hat{\mathbf{b}}_2^T = \mathbf{e}_2^T X^{-1}\hat{B} = \frac{1}{\sqrt{2}} \end{aligned}$$

Then, the pole-residue decomposition of $\mathbf{G}(s)$ and $\hat{\mathbf{H}}(s)$ are

$$\mathbf{G}(s) = \frac{\mathbf{l}_1\mathbf{r}_1^T}{s + \underbrace{1}_{-\mu_1}} + \frac{\mathbf{l}_2\mathbf{r}_2^T}{s + \underbrace{2}_{-\mu_2}} \text{ and } \hat{\mathbf{H}}(s) = \frac{\hat{\mathbf{c}}_1\hat{\mathbf{b}}_1^T}{s + \underbrace{6}_{-\hat{\lambda}_1}} + \frac{\hat{\mathbf{c}}_2\hat{\mathbf{b}}_2^T}{s + \underbrace{8}_{-\hat{\lambda}_2}}$$

and the $\mathcal{H}_2$ inner product can be computed as follows:

$$\begin{aligned}\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} &= \sum_{k=1}^{2} \mathbf{l}_k^T \hat{\mathbf{H}}(-\mu_k)\mathbf{r}_k \\ &= \mathbf{l}_1^T \hat{\mathbf{H}}(-\mu_1)\mathbf{r}_1 + \mathbf{l}_2^T \hat{\mathbf{H}}(-\mu_2)\mathbf{r}_2 \approx 0.6337.\end{aligned}$$

and, similarly (see Remark 4.5),

$$\begin{aligned}\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} &= \sum_{k=1}^{2} \hat{\mathbf{c}}_k^T \mathbf{G}(-\hat{\lambda}_k)\hat{\mathbf{b}}_k \\ &= \hat{\mathbf{c}}_1^T \mathbf{G}(-\hat{\lambda}_1)\hat{\mathbf{b}}_1 + \hat{\mathbf{c}}_2^T \mathbf{G}(-\hat{\lambda}_2)\hat{\mathbf{b}}_2 \approx 0.6337.\end{aligned}$$

The reader should note that if $\hat{\mathbf{H}}$ is a stable finite dimensional model, *i.e.*, $\hat{\mathbf{H}} = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B}$, then one can compute the pole-residue decomposition of $\hat{\mathbf{H}}$ by diagonalizing the pencil $(\hat{E}, \hat{A})$. In the case it is not diagonalizable, higher order poles will appear in the pole-residue decomposition. The procedure to obtain the pole-residue decomposition is described in Chapter 2, Section 2.2.

In the following, we show that the derivative is Hermitian with respect to the $\mathcal{H}_2$ inner product. The $\mathcal{H}_2$ inner product expression for models having higher order poles will be a consequence of this property.

4.2.2 Hermitian derivative

In this subsection, we derive the $\mathcal{H}_2$ inner product expression based on pole-residue when the model has higher order poles. The key tool to derive this expression is the Hermitian derivative property in $\mathcal{H}_2$. In what follows, $\mathbf{G}'(s)$ denotes the derivative of $\mathbf{G}(s)$ with respect to s. The Hermitian derivative property follows:

Proposition 4.7 (Hermitian derivative). *Let $\mathbf{G}, \mathbf{H} \in \mathcal{H}_2$ be strictly proper real models. In addition, suppose that $\mathbf{G}', \mathbf{H}' \in \mathcal{H}_2$. Then the derivative with respect to s is an Hermitian operator,* i.e.,

$$\langle \mathbf{G}, \mathbf{H}' \rangle_{\mathcal{H}_2} = \langle \mathbf{G}', \mathbf{H} \rangle_{\mathcal{H}_2}.$$

Proof. This is obtained using integration by parts as follows. Let us consider the integral over $[-\omega, \omega]$. By applying integration by parts one obtains :

$$\begin{aligned}S_\omega &= \int_{-\omega}^{\omega} \operatorname{trace}\Big(\mathbf{G}(-i\omega)\mathbf{H}'(i\omega)^T\Big)d\omega \\ &= \underbrace{\operatorname{trace}\Big(\frac{1}{i}\mathbf{G}(-i\omega)\mathbf{H}(i\omega)^T\Big|_{-\omega}^{\omega}\Big)}_{R_\omega} + \underbrace{\int_{-\omega}^{\omega} \operatorname{trace}\Big(\mathbf{G}'(-i\omega)\mathbf{H}(i\omega)^T\Big)d\omega}_{T_\omega}.\end{aligned}$$

Since $\mathbf{H}$ and $\mathbf{G}$ are elements of $\mathcal{H}_2$, then $\|\mathbf{H}\|_{\mathcal{H}_2} < \infty$ and $\|\mathbf{G}\|_{\mathcal{H}_2} < \infty$, which implies that $\lim_{\omega \to \pm\infty} \overline{\mathbf{H}(i\omega)}\mathbf{G}(i\omega) = 0$. In addition, we can take the limit when ω goes to infinity because $\mathbf{H}, \mathbf{H}', \mathbf{G}$ and $\mathbf{G}' \in \mathcal{H}_2$. Hence, the result follows noticing that $S_\omega \to \langle \mathbf{H}', \mathbf{G} \rangle_{\mathcal{H}_2}$, $T_\omega \to \langle \mathbf{H}, \mathbf{G}' \rangle_{\mathcal{H}_2}$ and $R_\omega \to 0$ when $\omega \to \infty$. □

One of the consequences of Proposition 4.7 is given in the following Lemma:

Lemma 4.8 (Higher order $\mathcal{H}_2$ inner product). *Let $\mathbf{G} \in \mathcal{H}_2$ be a strictly proper real model, $\hat{\mathbf{c}} \in \mathbb{C}^{n_y}$, $\hat{\mathbf{b}} \in \mathbb{C}^{n_u}$, and $\hat{\lambda} \in \mathbb{C}^-$. Then*

$$\left\langle \mathbf{G}, \frac{\hat{\mathbf{c}}\hat{\mathbf{b}}^T}{(s - \hat{\lambda})^2} \right\rangle_{\mathcal{H}_2} = -\hat{\mathbf{c}}^T \mathbf{G}'(-\hat{\lambda})\hat{\mathbf{b}}.$$

Moreover, for higher order models, one has

$$\left\langle \mathbf{G}, \frac{\hat{\mathbf{c}}\hat{\mathbf{b}}^T}{(s-\hat{\lambda})^n} \right\rangle_{\mathcal{H}_2} = \frac{(-1)^{n-1}}{(n-1)!} \hat{\mathbf{c}}^T \mathbf{G}^{(n-1)}(-\hat{\lambda})\hat{\mathbf{b}},$$

where $\mathbf{G}^{(n)}$ denotes the nth derivative of $\mathbf{G}(s)$ with respect to s.

Proof. It is obtained by observing

$$\left(\frac{\phi}{s-\hat{\lambda}}\right)' = -\frac{\phi}{(s-\hat{\lambda})^2} \text{ and } \left(\frac{\phi}{s-\hat{\lambda}}\right)^{(n-1)} = (-1)^{n-1}(n-1)!\frac{\phi}{(s-\hat{\lambda})^n}.$$

Then,

$$\left\langle \mathbf{G}, \frac{\hat{\mathbf{c}}\hat{\mathbf{b}}^T}{(s-\hat{\lambda})^2} \right\rangle_{\mathcal{H}_2} = -\left\langle \mathbf{G}, \left(\frac{\hat{\mathbf{c}}\hat{\mathbf{b}}^T}{(s-\hat{\lambda})}\right)' \right\rangle_{\mathcal{H}_2} \tag{4.10}$$

$$\underbrace{=}_{\text{Hermitian derivative}} -\left\langle \mathbf{G}', \frac{\hat{\mathbf{c}}\hat{\mathbf{b}}^T}{(s-\hat{\lambda})} \right\rangle_{\mathcal{H}_2} \tag{4.11}$$

$$\underbrace{=}_{\text{Lemma 4.2}} -\hat{\mathbf{c}}^T \mathbf{G}'(-\hat{\lambda})\hat{\mathbf{b}}. \tag{4.12}$$

The result for higher derivative follows by induction. □

Lemma 4.8 enables us to compute the $\mathcal{H}_2$ inner product for models having poles with multiplicity bigger than one. This result is stated in the following theorem, which generalizes Theorem 4.4 when the model has higher order poles.

Theorem 4.9 (Spectral formulation of the $\mathcal{H}_2$ inner product, higher order poles). *Let $\mathbf{G}, \hat{\mathbf{H}} \in \mathcal{H}_2$ be two strictly proper real models. Let us suppose that the transfer function of $\hat{\mathbf{H}}$ can be written as*

$$\hat{\mathbf{H}}(s) = \sum_{k=1}^{n_J} \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{(s-\hat{\lambda})^k},$$

where $\hat{\mathbf{c}}_k \in \mathbb{C}^{n_y}$, $\hat{\mathbf{b}}_k \in \mathbb{C}^{n_u}$, for $k = 1, \ldots, n_J$, and $\hat{\lambda} \in \mathbb{C}^-$. Then,

$$\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} = \sum_{k=1}^{n_J} \frac{(-1)^{k-1}}{(k-1)!} \hat{\mathbf{c}}_k^T \mathbf{G}^{(k-1)}(-\hat{\lambda})\hat{\mathbf{b}}_k.$$

Moreover, the $\mathcal{H}_2$ norm of $\hat{\mathbf{H}}$ can be computed as

$$\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2 = \sum_{k=1}^{n_J} \frac{(-1)^{k-1}}{(k-1)!} \hat{\mathbf{c}}_k^T \hat{\mathbf{H}}^{(k-1)}(-\hat{\lambda})\hat{\mathbf{b}}_k.$$

Proof. By applying Proposition 4.8. It follows:

$$\begin{aligned}\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} &= \langle \mathbf{G}, \sum_{k=1}^{n_J} \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{(s-\lambda)^k} \rangle_{\mathcal{H}_2} \\ &= \sum_{k=1}^{n_J} \langle \mathbf{G}, \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{(s-\lambda)^k} \rangle_{\mathcal{H}_2} \\ \underbrace{=}_{\text{Lemma 4.8}} &\sum_{k=1}^{n_J} \frac{(-1)^{n-1}}{(n-1)!} \hat{\mathbf{c}}_k^T \mathbf{G}^{(n)}(-\lambda) \hat{\mathbf{b}}_k\end{aligned}$$

□

The Examples 4.10 and 4.11 illustrate the $\mathcal{H}_2$ inner product and norm computation via partial fraction decomposition.

Example 4.10 (Spectral $\mathcal{H}_2$ inner product application #1). *Let us consider the SISO model $\mathbf{H} = (E, A, B, C)$ from Example 2.11 whose transfer function is given by*

$$\mathbf{H}(s) = -\frac{3}{s+8} + \frac{6}{s+4} + \frac{18}{(s+4)^2}.$$

Then the $\mathcal{H}_2$ norm of $\mathbf{H}$ can be computed using Propositions 4.4 and 4.9 as follows:

$$\begin{aligned}\|\mathbf{H}\|_{\mathcal{H}_2}^2 &= \langle \mathbf{H}, \mathbf{H} \rangle_{\mathcal{H}_2} \\ &= \langle \mathbf{H}, -\tfrac{3}{s-8} + \tfrac{6}{s-4} + \tfrac{18}{(s-4)^2} \rangle_{\mathcal{H}_2} \\ &= -3\mathbf{H}(8) + 6\mathbf{H}(4) - 18\mathbf{H}'(4) \approx 5.9531.\end{aligned}$$

Example 4.11 (Spectral $\mathcal{H}_2$ inner product application #2). *In this example, the expression provided in Theorem 4.9 is used to compute the norm of a simple higher order model. Let us consider $n \in \mathbb{N}^*$, $\lambda \in \mathbb{R}^+$ and a model $\mathbf{G}_n \in \mathcal{H}_2$ whose transfer function is given by*

$$\mathbf{G}_n(s) = \frac{1}{(s+\lambda)^n}.$$

Hence, the $\mathcal{H}_2$ norm of $\mathbf{G}_n(s)$ can be computed by

$$\begin{aligned}\langle \mathbf{G}_n, \mathbf{G}_n \rangle_{\mathcal{H}_2} &= \langle \mathbf{G}_n, \frac{1}{(s+\lambda)^n} \rangle_{\mathcal{H}_2} \\ \underbrace{=}_{\text{Theorem 4.9}} &\frac{(-1)^{n-1}}{(n-1)!} \mathbf{G}_n^{(n-1)}(\lambda).\end{aligned}$$

Since,

$$\mathbf{G}_n^{(n-1)}(s) = \frac{(2n-2)!}{(n-1)!} \frac{(-1)^{n-1}}{(s+\lambda)^{2n-1}},$$

we obtain

$$\|\mathbf{G}_n\|_{\mathcal{H}_2}^2 = \langle \mathbf{G}_n, \mathbf{G}_n \rangle_{\mathcal{H}_2} = \frac{(2n-2)!}{(n-1)!^2} \frac{1}{(2\lambda)^{2n-1}}.$$

The reader should note that if $\hat{\mathbf{H}}$ is a stable finite dimensional model, *i.e.*, $\hat{\mathbf{H}} = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B}$, then one can characterize the $\mathcal{H}_2$ inner product by the pole-residue decomposition of $\hat{\mathbf{H}}$. The

procedure to obtain the pole-residue decomposition is described in Chapter 2, Section 2.2. Once the pole-residue decomposition is obtained, we are able to compute the $\mathcal{H}_2$ inner product provided in this section. It is worth noting that the computation of the pole-residue decomposition require the diagonlization of the pencil $(\hat{E}, \hat{A})$, which is an expensive task for a large-scale system. We have derived this spectral expression of the $\mathcal{H}_2$ inner product in order to parametrize it as a function of the reduced order model's poles and residues. This will help us to derive the optimality conditions of Problem 4.1.

To sum up, in this section we developed the pole-residue expression of the $\mathcal{H}_2$ inner product. This gives a parametrization of this inner product as a function of the poles and residues of the reduced order model $\hat{\mathbf{H}}$. In the next section, we tackle the $\mathcal{H}_2$ approximation Problem 4.1 using these expressions.

4.3 Formulation of the $\mathcal{H}_2$ first-order optimality conditions

This section is dedicated to the derivation of the necessary $\mathcal{H}_2$ optimality conditions of Problem 4.1. It is organized as follows:

1. In subsection 4.3.1, we consider the case where the poles of the ROM are fixed. In this case, the constraints defined a vector space (and, consequently a convex set) and the approximation problem can be solve by projection.
2. In subsection 4.3.2, the $\mathcal{H}_2$ inner product is used to characterize the approximation error.
3. In subsection 4.3.3, we suppose that $\mathbf{G}$ is a finite dimensional system. Then the gradients with respect to the parameters defining the reduced order model $\hat{\mathbf{H}}$ are explicitly computed and the $\mathcal{H}_2$ optimality condtions are derived.
4. In subsection 4.3.4, an extension of the result is presented and no particular structure is assumed for $\mathbf{G}$. This includes the case where $\mathbf{G}$ is represented by an irrational transfer function.

In the following subsection, we will present the model approximation problem when the reduced order model has assigned spectra. This is a much simpler problem which can be solved by Hilbert projection. Even if it is a simplification of Problem 4.1, we believe that this is the first step for a better understanding.

4.3.1 Reduced model with assigned eigenvalues in the SISO case

In Section 4.1, we show that Problem 4.1 is non-convex because the constrain $\mathcal{M} = \{\hat{\mathbf{H}} \in \mathcal{H}_2, order(\hat{\mathbf{H}}) = n\}$ is not-convex. However, if the constrains define a closed vector space, Problem 4.1 could be solved by the Hilbert projection theorem (see [Kreyszig, 1989, Chapter 3]). A very interesting example of constrains that define a closed vector space is the space of reduced order models with fixed semi-simple eigenvalues. More precisely, if $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, with $\lambda_k \in \mathbb{C}^-$, for $k = 1, \ldots, n$, then

$$\mathcal{M}_\Lambda = \left\{ \sum_{k=1}^{n} \frac{\hat{\phi}_k}{s - \lambda_k},\ \hat{\phi}_k \in \mathbb{C}, \text{for} \quad k = 1, \ldots, n \right\}$$

is the space of SISO models of order n having pre-assigned semi-simple poles from Λ. $\mathcal{M}_\Lambda$ is clearly a finite-dimensional vector space and thus closed. Moreover, $\{\frac{1}{s-\lambda_1}, \frac{1}{s-\lambda_2}, \ldots, \frac{1}{s-\lambda_n}\}$ is a basis of $\mathcal{M}_\Lambda$ and this proposition follows :

Proposition 4.12 (Model approximation with fixed eigenvalues - SISO case). *Let $\mathbf{G} \in \mathcal{H}_2$ be a real model and $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ be a set closed by conjugation (*i.e.*, $\overline{\Lambda} = \Lambda$). Then, the problem of finding real model $\hat{\mathbf{H}} \in \mathcal{M}_\Lambda$ which minimizes $\|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{H}_2}$ has an unique global minimizer $\hat{\mathbf{H}}^\star$ which is the orthogonal projection of $\mathbf{G} \in \mathcal{H}_2$ over the finite-dimensional subspace $\mathcal{M}_\Lambda$,* i.e., *$\hat{\mathbf{H}}^\star \in \mathcal{M}_\Lambda$ and*

$$\langle \mathbf{G} - \hat{\mathbf{H}}^\star, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} = 0, \text{ for every } \hat{\mathbf{H}} \in \mathcal{M}_\Lambda.$$

Moreover, $\hat{\mathbf{H}}^\star$ is the unique model from $\mathcal{M}_\Lambda$ satisfying the interpolatory conditions

$$\mathbf{G}(-\lambda_k) = \hat{\mathbf{H}}^\star(-\lambda_k), \text{ for } \lambda_k \in \Lambda. \tag{4.13}$$

Proof. This is a consequence of the Hilbert projection theorem. Hence, there exists an unique global minimizer $\hat{\mathbf{H}}$ of the approximation problem with assigned eigenvalues. If we denote by $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\} = \{\frac{1}{s-\lambda_1}, \frac{1}{s-\lambda_2}, \ldots, \frac{1}{s-\lambda_n}\}$, a basis of $\mathcal{M}_\Lambda$, then $\mathbf{H}^\star$ can be constructed as follows:

(i) $\mathbf{H}^\star = \sum_{k=1}^{n} \hat{\phi}_k \mathbf{v}_k$, for $\hat{\phi}_k \in \mathbb{C}$.

(ii) $\langle \mathbf{G} - \mathbf{H}^\star, \mathbf{v}_k \rangle_{\mathcal{H}_2} = 0$, for $k = 1, \ldots, n$.

Hence, from item (ii), we have

$$\begin{array}{rccll} & \langle \mathbf{G}, \mathbf{v}_k \rangle_{\mathcal{H}_2} & = & \langle \mathbf{H}^\star, \mathbf{v}_k \rangle_{\mathcal{H}_2} & \\ \text{hence,} & \mathbf{G}(-\lambda_k) & = & \mathbf{H}^\star(-\lambda_k), & \text{for } k = 1, \ldots, n, \end{array}$$

which proves the interpolatory conditions from (4.13). In addition, using item (i) and developing the inner product, one obtains

$$\underbrace{\begin{bmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle_{\mathcal{H}_2} & \cdots & \langle \mathbf{v}_1, \mathbf{v}_n \rangle_{\mathcal{H}_2} \\ \vdots & \ddots & \vdots \\ \langle \mathbf{v}_n, \mathbf{v}_1 \rangle_{\mathcal{H}_2} & \cdots & \langle \mathbf{v}_n, \mathbf{v}_n \rangle_{\mathcal{H}_2} \end{bmatrix}}_{:=M} \underbrace{\begin{bmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_n \end{bmatrix}}_{:=\Phi} = \underbrace{\begin{bmatrix} \langle \mathbf{G}, \mathbf{v}_1 \rangle_{\mathcal{H}_2} \\ \vdots \\ \langle \mathbf{G}, \mathbf{v}_n \rangle_{\mathcal{H}_2} \end{bmatrix}}_{:=\mathbf{g}}. \tag{4.14}$$

Thus, $\mathbf{H}^\star$ can be constructed by solving the linear system $M\Phi = \mathbf{g}$. Moreover, if we use the fact that

$$\langle \mathbf{G}, \mathbf{v}_k \rangle_{\mathcal{H}_2} = \langle \mathbf{G}, \frac{1}{s-\lambda_k} \rangle_{\mathcal{H}_2} = \mathbf{G}(-\lambda_k) \quad \text{and} \quad \langle \mathbf{v}_i, \mathbf{v}_j \rangle_{\mathcal{H}_2} = \frac{1}{-\overline{\lambda_i} - \lambda_j},$$

the matrices M and $\mathbf{g}$ can be expressed as

$$M = \begin{bmatrix} \frac{1}{-2Re(\lambda_1)} & \cdots & \frac{1}{-\lambda_n^* - \lambda_1} \\ \vdots & \ddots & \vdots \\ \frac{1}{-\lambda_1^* - \lambda_n} & \cdots & \frac{1}{-2Re(\lambda_n)} \end{bmatrix} \quad \text{and} \quad \mathbf{g} = \begin{bmatrix} \mathbf{G}(-\lambda_1) \\ \vdots \\ \mathbf{G}(-\lambda_n). \end{bmatrix} \tag{4.15}$$

□

It is worth noticing that the matrix M has a Cauchy structure (see [Antoulas, 2005, Chapter 9]). Reader should note that one could have chosen another basis of $\mathcal{M}_\Lambda$. For example, if $\{\tilde{\mathbf{v}}_\mathbf{1}, \ldots, \tilde{\mathbf{v}}_\mathbf{n}\}$ formed an orthogonal basis of $\mathcal{M}_\Lambda$, then the matrix M would be diagonal and the linear equation $M\Phi = \mathbf{g}$ would be easy to solve. In [Mi et al., 2012], a description of an efficient method to build this orthogonal basis is provided.

In order words, if the poles $\lambda_1, \ldots, \lambda_n$ are assigned, one can compute the unique optimum residues $\hat{\phi}_1, \ldots, \hat{\phi}_k$ through Proposition 4.12. Hence, for each $\mathbf{G} \in \mathcal{H}_2$, there is a function $\mathfrak{f}_{\mathbf{G}}$ which associates a set of poles to the unique optimum residues, *i.e.*,

$$\mathfrak{f}_{\mathbf{G}}(\lambda_1, \ldots, \lambda_n) = \begin{bmatrix} \hat{\phi}_1 & \hat{\phi}_2 & \ldots & \hat{\phi}_n \end{bmatrix}$$

and the model constructed such that $\hat{\mathbf{H}} = \sum_{k=1}^{n} \frac{\hat{\phi}_k}{s-\lambda_k}$ is the optimal approximation on $\mathcal{M}_\Lambda$ of $\mathbf{G}$. This result dates from the 30's in the complex analysis domain and it was first presented by Walsh in [Walsh, 1932]. In the monograph [Vuillemin, 2014, Chapter 4], this result is developed in a different way using optimization and it develops a nice simple example showing the non-convexity of Problem 4.1 numerically.

For the generalization of Proposition 4.12 to the MIMO case reader should refer to the monograph [Wolf, 2014]. Moreover, this result is used in [Wolf et al., 2013] to construct pseudo-optimal solutions to the non-convex Problem 4.1.

In what follows, we return back to the non-convex Problem 4.1, using the $\mathcal{H}_2$ inner product to characterize the approximation error.

4.3.2 $\mathcal{H}_2$ approximation error

The previous section gives a spectral expression of the $\mathcal{H}_2$ inner product from a general viewpoint, *i.e.*, the assumptions over $\mathbf{G}$ and $\hat{\mathbf{H}}$ are very weak. This expression will be useful to characterize some properties of Problem 4.1. We are now ready to describe the $\mathcal{H}_2$ error in this problem.

Lemma 4.13 ($\mathcal{H}_2$ **mismatch error characterization).** *Let us consider* $\mathbf{G}, \hat{\mathbf{H}} \in \mathcal{H}_2$*, two real strictly proper models. Then*

$$\mathcal{J}_{\mathcal{H}_2}(\hat{\mathbf{H}})^2 = \|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{H}_2}^2 = \|\mathbf{G}\|_{\mathcal{H}_2}^2 - 2\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} + \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2. \tag{4.16}$$

Proof. In order to simplify the notation, henceforth we denote by $\mathcal{J}_{\mathcal{H}_2}(\hat{\mathbf{H}})^2 := \mathcal{J}_2$. We should only write the norm as an inner product and develop the expression as follows :

$$\begin{aligned} \mathcal{J}_2 &= \|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{H}_2}^2 = \langle \mathbf{G} - \hat{\mathbf{H}}, \mathbf{G} - \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} \\ &= \|\mathbf{G}\|_{\mathcal{H}_2}^2 - 2\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} + \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2. \end{aligned}$$

□

Notice that the only terms from $\mathcal{J}_2$ which depend on $\hat{\mathbf{H}}$ are $2\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2}$ and $\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2$. Hence, $\|\mathbf{G}\|_{\mathcal{H}_2}^2$ does not need to be computed at all for optimization purposes.

Our goal here is to characterize the optimality conditions related to Problem 4.1. Thus, we should derive the gradient of $\mathcal{J}_2$ with respect to the parameters defining $\hat{\mathbf{H}}$. In what follows, we derive these optimality conditions in two different ways. In the first one, presented in Subsection 4.3.3, we assume that $\mathbf{G}$ is a finite dimensional system, and we analytically compute the terms $2\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2}$ and $\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2$. In the second one, presented in Subsection 4.3.4, we consider that the reduced order model depends on a vector of parameters, no special structure will be assumed on $\mathbf{G}$, and we derive the optimality conditions with respect to this latter. This second derivation extends the first one and includes the case where $\mathbf{G}$ is represented by an irrational transfer function.

4.3.3 $\mathcal{H}_2$ optimality conditions derivation

Let us assume that $\mathbf{G}$ is a finite dimensional representation and that both $\mathbf{G}$, $\hat{\mathbf{H}}$ have only semi-simple poles and can be expressed by the following pole-residue decompositions

$$\mathbf{G}(s) = \sum_{k=1}^{N} \frac{\mathbf{l}_k \mathbf{r}_k^T}{s - \mu_k} \quad \text{and} \quad \hat{\mathbf{H}}(s) = \sum_{k=1}^{n} \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k}.$$

Hence, by the pole-residue $\mathcal{H}_2$ inner product expression, the term $\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2$ can be written as

$$\begin{aligned} \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2 &= \sum_{k=1}^{n} \hat{\mathbf{c}}_k^T \hat{\mathbf{H}}(-\hat{\lambda}_k) \hat{\mathbf{b}}_k \\ &= \sum_{k=1}^{n} \hat{\mathbf{c}}_k^T \Big(\sum_{m=1}^{n} \frac{\hat{\mathbf{c}}_m \hat{\mathbf{b}}_m^T}{-\hat{\lambda}_k - \hat{\lambda}_m} \Big) \hat{\mathbf{b}}_k \\ &= \sum_{k=1}^{n} \sum_{m=1}^{n} \frac{(\hat{\mathbf{c}}_k^T \hat{\mathbf{c}}_m)(\hat{\mathbf{b}}_m^T \hat{\mathbf{b}}_k)}{-\hat{\lambda}_k - \hat{\lambda}_m} \end{aligned}$$

and the term $\langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2}$ can be written as

$$\begin{aligned} \langle \mathbf{G}, \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} &= \sum_{j=1}^{N} \mathbf{l}_j^T \hat{\mathbf{H}}(-\mu_j) \mathbf{r}_j \\ &= \sum_{j=1}^{N} \mathbf{l}_j^T \Big(\sum_{m=1}^{n} \frac{\hat{\mathbf{c}}_m \hat{\mathbf{b}}_m^T}{-\mu_j - \hat{\lambda}_m} \Big) \mathbf{r}_j \\ &= \sum_{j=1}^{N} \sum_{m=1}^{n} \frac{(\mathbf{l}_j^T \hat{\mathbf{c}}_m)(\hat{\mathbf{b}}_m^T \mathbf{r}_j)}{-\mu_j - \hat{\lambda}_m} \end{aligned}$$

In what follows, the gradients of $\mathcal{J}_2$ with respect to the parameters defining $\hat{\mathbf{H}}$ are then properly computed.

Gradient with respect to poles and residues

As explained above, $\mathcal{J}_2$ is a function of $\hat{\mathbf{b}}_k$, $\hat{\mathbf{c}}_k$ and $\hat{\lambda}_k$. Let us now compute the gradient of $\mathcal{J}_2$ with respect to these parameters.

Gradient with respect to $\hat{\mathbf{b}}_l$: Let us compute the gradient of $\mathcal{J}_2$ with respect to $\hat{\mathbf{b}}_l$. We are going to do this in multiple steps. Let δ_{lm} be the Kronecker δ_{lm}, *i.e.*,

$$\delta_{lm} = \begin{cases} 1 \text{ if } l = m \\ 0 \text{ otherwise} \end{cases}.$$

First, let's compute the gradient of $\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2$ with respect $\hat{\mathbf{b}}_l$ step by step:

$$\begin{aligned}
\nabla_{\hat{\mathbf{b}}_l}\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2 &= \sum_{k=1}^{n}\sum_{m=1}^{n}\frac{(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_m)}{-\hat{\lambda}_k-\hat{\lambda}_m}\nabla_{\hat{\mathbf{b}}_l}(\hat{\mathbf{b}}_m^T\hat{\mathbf{b}}_k)\\
&= \sum_{k=1}^{n}\sum_{m=1}^{n}\frac{(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_m)}{-\hat{\lambda}_k-\hat{\lambda}_m}(\hat{\mathbf{b}}_m^T\delta_{lk}+\hat{\mathbf{b}}_k^T\delta_{lm})\\
&= \sum_{k=1}^{n}\sum_{m=1}^{n}\frac{(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_m)}{-\hat{\lambda}_k-\hat{\lambda}_m}\hat{\mathbf{b}}_m^T\delta_{lk}+\sum_{k=1}^{n}\sum_{m=1}^{n}\frac{(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_m)}{-\hat{\lambda}_k-\hat{\lambda}_m}\hat{\mathbf{b}}_k^T\delta_{lm}\\
&= \sum_{m=1}^{n}\frac{(\hat{\mathbf{c}}_l^T\hat{\mathbf{c}}_m)}{-\hat{\lambda}_l-\hat{\lambda}_m}\hat{\mathbf{b}}_m^T+\sum_{k=1}^{n}\frac{(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_l)}{-\hat{\lambda}_k-\hat{\lambda}_l}\hat{\mathbf{b}}_k^T\\
&\underbrace{=}_{\text{Use }\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_l=\hat{\mathbf{c}}_l^T\hat{\mathbf{c}}_k} \hat{\mathbf{c}}_l^T\sum_{m=1}^{n}\frac{\hat{\mathbf{c}}_m\hat{\mathbf{b}}_m^T}{-\hat{\lambda}_l-\hat{\lambda}_m}+\hat{\mathbf{c}}_l^T\sum_{k=1}^{n}\frac{\hat{\mathbf{c}}_k\hat{\mathbf{b}}_k^T}{-\hat{\lambda}_k-\hat{\lambda}_l}\\
&= \hat{\mathbf{c}}_l^T\hat{\mathbf{H}}(-\hat{\lambda}_l)+\hat{\mathbf{c}}_l^T\hat{\mathbf{H}}(-\hat{\lambda}_l)=2\hat{\mathbf{c}}_l^T\hat{\mathbf{H}}(-\hat{\lambda}_l).
\end{aligned}$$

In a similar way, we compute the gradient of $\langle\mathbf{G},\hat{\mathbf{H}}\rangle_{\mathcal{H}_2}$ with respect to $\hat{\mathbf{b}}_l$ step by step:

$$\begin{aligned}
\nabla_{\hat{\mathbf{b}}_l}\langle\mathbf{G},\hat{\mathbf{H}}\rangle_{\mathcal{H}_2} &= \sum_{j=1}^{N}\sum_{m=1}^{n}\frac{(\mathbf{l}_j^T\hat{\mathbf{c}}_m)\mathbf{r}_j^T}{-\mu_j-\hat{\lambda}_m}\nabla_{\hat{\mathbf{b}}_l}\hat{\mathbf{b}}_m\\
&= \sum_{j=1}^{N}\sum_{m=1}^{n}\frac{(\mathbf{l}_j^T\hat{\mathbf{c}}_m)\mathbf{r}_j^T}{-\mu_j-\hat{\lambda}_m}\delta_{ml}\\
&= \sum_{j=1}^{N}\frac{(\mathbf{l}_j^T\hat{\mathbf{c}}_l)\mathbf{r}_j^T}{-\mu_j-\hat{\lambda}_l}\\
&\underbrace{=}_{\text{Use }\mathbf{l}_j^T\hat{\mathbf{c}}_l=\hat{\mathbf{c}}_l^T\mathbf{l}_j} \hat{\mathbf{c}}_l^T\sum_{j=1}^{N}\frac{\mathbf{l}_j\mathbf{r}_j^T}{-\mu_j-\hat{\lambda}_l}\\
&= \hat{\mathbf{c}}_l^T\mathbf{G}(-\hat{\lambda}_l).
\end{aligned}$$

Finally, the gradient of $\mathcal{J}_2$ with respect to $\hat{\mathbf{b}}_l$ is given by

$$\begin{aligned}
\nabla_{\hat{\mathbf{b}}_l}\mathcal{J}_2 &= -2\nabla_{\hat{\mathbf{b}}_l}\langle\mathbf{G},\hat{\mathbf{H}}\rangle_{\mathcal{H}_2}+\nabla_{\hat{\mathbf{b}}_l}\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2\\
&= -2\hat{\mathbf{c}}_l^T\mathbf{G}(-\hat{\lambda}_l)+2\hat{\mathbf{c}}_l^T\hat{\mathbf{H}}(-\hat{\lambda}_l)
\end{aligned}$$

Gradient with respect to $\hat{\mathbf{c}}_l$: similarly, let us compute the gradient of $\mathcal{J}_2$ with respect to $\hat{\mathbf{c}}_l$. We are going to do this in multiple steps. Once again, let's compute the gradient of $\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2$ with respect $\hat{\mathbf{c}}_l$ step by step:

$$\begin{aligned}
\nabla_{\hat{\mathbf{c}}_l}\|\hat{\mathbf{H}}\|^2_{\mathcal{H}_2} &= \sum_{k=1}^{n}\sum_{m=1}^{n}\frac{(\hat{\mathbf{b}}_m^T\hat{\mathbf{b}}_k)}{-\hat{\lambda}_k-\hat{\lambda}_m}\nabla_{\hat{\mathbf{c}}_l}(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_m)\\
&= \sum_{k=1}^{n}\sum_{m=1}^{n}\frac{(\hat{\mathbf{b}}_k^T\hat{\mathbf{b}}_m)}{-\hat{\lambda}_k-\hat{\lambda}_m}(\hat{\mathbf{c}}_m^T\delta_{lk}+\hat{\mathbf{c}}_k^T\delta_{lm})\\
&= \sum_{k=1}^{n}\sum_{m=1}^{n}\frac{(\hat{\mathbf{b}}_k^T\hat{\mathbf{b}}_m)}{-\hat{\lambda}_k-\hat{\lambda}_m}\hat{\mathbf{c}}_m^T\delta_{lk}+\sum_{k=1}^{n}\sum_{m=1}^{n}\frac{(\hat{\mathbf{b}}_k^T\hat{\mathbf{b}}_m)}{-\hat{\lambda}_k-\hat{\lambda}_m}\hat{\mathbf{c}}_k^T\delta_{lm}\\
&= \sum_{m=1}^{n}\frac{(\hat{\mathbf{b}}_l^T\hat{\mathbf{b}}_m)}{-\hat{\lambda}_l-\hat{\lambda}_m}\hat{\mathbf{c}}_m^T+\sum_{k=1}^{n}\frac{(\hat{\mathbf{b}}_k^T\hat{\mathbf{b}}_l)}{-\hat{\lambda}_k-\hat{\lambda}_l}\hat{\mathbf{c}}_k^T\\
&= \sum_{m=1}^{n}\frac{(\hat{\mathbf{b}}_l^T\hat{\mathbf{b}}_m)\hat{\mathbf{c}}_m^T}{-\hat{\lambda}_l-\hat{\lambda}_m}+\sum_{k=1}^{n}\frac{(\hat{\mathbf{b}}_l^T\hat{\mathbf{b}}_k)\hat{\mathbf{c}}_k^T}{-\hat{\lambda}_k-\hat{\lambda}_l}\\
&= \hat{\mathbf{b}}_l^T\sum_{m=1}^{n}\frac{\hat{\mathbf{b}}_m\hat{\mathbf{c}}_m^T}{-\hat{\lambda}_l-\hat{\lambda}_m}+\hat{\mathbf{b}}_l^T\sum_{k=1}^{n}\frac{\hat{\mathbf{b}}_k\hat{\mathbf{c}}_k^T}{-\hat{\lambda}_k-\hat{\lambda}_l}\\
&= \hat{\mathbf{b}}_l^T\hat{\mathbf{H}}^T(-\hat{\lambda}_l)+\hat{\mathbf{b}}_l^T\hat{\mathbf{H}}^T(-\hat{\lambda}_l).
\end{aligned}$$

and

$$\begin{aligned}
\nabla_{\hat{\mathbf{c}}_l}\langle\mathbf{G},\hat{\mathbf{H}}\rangle_{\mathcal{H}_2} &= \sum_{j=1}^{N}\sum_{m=1}^{n}\frac{(\mathbf{r}_j^T\hat{\mathbf{b}}_m)\mathbf{l}_j^T}{-\mu_j-\hat{\lambda}_m}\nabla_{\hat{\mathbf{c}}_l}\hat{\mathbf{c}}_m\\
&= \sum_{j=1}^{N}\sum_{m=1}^{n}\frac{(\mathbf{r}_j^T\hat{\mathbf{b}}_m)\mathbf{l}_j^T}{-\mu_j-\hat{\lambda}_m}\delta_{ml}\\
&= \sum_{j=1}^{N}\frac{(\mathbf{r}_j^T\hat{\mathbf{b}}_l)\mathbf{l}_j^T}{-\mu_j-\hat{\lambda}_l}\\
&= \hat{\mathbf{b}}_l^T\sum_{j=1}^{N}\frac{\mathbf{r}_j\mathbf{l}_j^T}{-\mu_j-\hat{\lambda}_l}\\
&= \hat{\mathbf{b}}_l^T\mathbf{G}^T(-\hat{\lambda}_l)
\end{aligned}$$

Finally, the gradient of $\mathcal{J}_2$ with respect to $\hat{\mathbf{c}}_l$ is given by

$$\begin{aligned}
\nabla_{\hat{\mathbf{c}}_l}\mathcal{J} &= -2\nabla_{\hat{\mathbf{c}}_l}\langle\mathbf{G},\hat{\mathbf{H}}\rangle_{\mathcal{H}_2}+\nabla_{\hat{\mathbf{c}}_l}\|\hat{\mathbf{H}}\|^2_{\mathcal{H}_2}\\
&= -2\hat{\mathbf{b}}_l^T\mathbf{G}^T(-\hat{\lambda}_l)+2\hat{\mathbf{b}}_l^T\hat{\mathbf{H}}^T(-\hat{\lambda}_l).
\end{aligned}$$

Hence,

$$\nabla_{\hat{\mathbf{c}}_l}\mathcal{J}=2\hat{\mathbf{b}}_l^T\mathbf{G}^T(-\hat{\lambda}_l)+2\hat{\mathbf{b}}_l^T\hat{\mathbf{H}}^T(-\hat{\lambda}_l),$$

and by taking the transposition, this leads to have the tangential interpolation conditions on $\mathbf{G}$ and $\hat{\mathbf{H}}$, instead of $\mathbf{G}^T$ and $\hat{\mathbf{H}}^T$, as follows:

$$\nabla_{\hat{\mathbf{c}}_l}\mathcal{J}^T=2\mathbf{G}(-\hat{\lambda}_l)\hat{\mathbf{b}}_l+2\hat{\mathbf{H}}(-\hat{\lambda}_l)\hat{\mathbf{b}}_l$$

Gradient with respect to $\hat{\lambda}_l$: Finally, let us compute the gradient of $\mathcal{J}_2$ with respect to $\hat{\lambda}_l$. Hence, let us compute the gradient of $\|\hat{\mathbf{H}}\|^2_{\mathcal{H}_2}$ with respect $\hat{\lambda}_l$ step by step:

$$\begin{aligned}
\nabla_{\hat{\lambda}_l}\|\hat{\mathbf{H}}\|^2_{\mathcal{H}_2} &= \sum_{k=1}^{n}\sum_{m=1}^{n}(\hat{\mathbf{b}}_m^T\hat{\mathbf{b}}_k)(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_m)\nabla_{\hat{\lambda}_l}\left(\frac{1}{-\hat{\lambda}_k-\hat{\lambda}_m}\right)\\
&= \sum_{k=1}^{n}\sum_{m=1}^{n}(\hat{\mathbf{b}}_k^T\hat{\mathbf{b}}_m)(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_m)\left(\frac{1}{(-\hat{\lambda}_k-\hat{\lambda}_m)^2}\delta_{lk}+\frac{1}{(-\hat{\lambda}_k-\hat{\lambda}_m)^2}\delta_{lm}\right)\\
&= \sum_{k=1}^{n}\sum_{m=1}^{n}\frac{(\hat{\mathbf{b}}_k^T\hat{\mathbf{b}}_m)(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_m)}{(-\hat{\lambda}_k-\hat{\lambda}_m)^2}\delta_{lk}+\sum_{k=1}^{n}\sum_{m=1}^{n}\frac{(\hat{\mathbf{b}}_k^T\hat{\mathbf{b}}_m)(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_m)}{(-\hat{\lambda}_k-\hat{\lambda}_m)^2}\delta_{lm}\\
&= \sum_{m=1}^{n}\frac{(\hat{\mathbf{c}}_l^T\hat{\mathbf{c}}_m)(\hat{\mathbf{b}}_l^T\hat{\mathbf{b}}_m)}{(-\hat{\lambda}_l-\hat{\lambda}_m)^2}+\sum_{k=1}^{n}\frac{(\hat{\mathbf{c}}_k^T\hat{\mathbf{c}}_l)(\hat{\mathbf{b}}_k^T\hat{\mathbf{b}}_l)}{(-\hat{\lambda}_k-\hat{\lambda}_l)^2}\\
&= \hat{\mathbf{c}}_l^T\left(\sum_{m=1}^{n}\frac{\hat{\mathbf{c}}_m\hat{\mathbf{b}}_m^T}{(-\hat{\lambda}_l-\hat{\lambda}_m)^2}\right)\hat{\mathbf{b}}_l+\hat{\mathbf{c}}_l^T\left(\sum_{k=1}^{n}\frac{\hat{\mathbf{c}}_k\hat{\mathbf{b}}_k^T}{(-\hat{\lambda}_k-\hat{\lambda}_l)^2}\right)\hat{\mathbf{b}}_l\\
&= -\hat{\mathbf{c}}_l^T\hat{\mathbf{H}}'(-\hat{\lambda}_l)\hat{\mathbf{b}}_l-\hat{\mathbf{c}}_l^T\hat{\mathbf{H}}'(-\hat{\lambda}_l)\hat{\mathbf{b}}_l.
\end{aligned}$$

and

$$\begin{aligned}
\nabla_{\hat{\lambda}_l}\langle\mathbf{G},\hat{\mathbf{H}}\rangle_{\mathcal{H}_2} &= \sum_{j=1}^{N}\sum_{m=1}^{n}(\mathbf{r}_j^T\hat{\mathbf{b}}_m)(\mathbf{l}_j^T\hat{\mathbf{c}}_m)\ \nabla_{\hat{\lambda}_l}\left(\frac{1}{-\mu_j-\hat{\lambda}_m}\right)\\
&= \sum_{j=1}^{N}\sum_{m=1}^{n}\frac{(\mathbf{r}_j^T\hat{\mathbf{b}}_m)(\mathbf{l}_j^T\hat{\mathbf{c}}_m)}{(-\mu_j-\hat{\lambda}_m)^2}\delta_{ml}\\
&= \sum_{j=1}^{N}\frac{(\mathbf{l}_j^T\hat{\mathbf{c}}_l)(\mathbf{r}_j^T\hat{\mathbf{b}}_l)}{(-\mu_j-\hat{\lambda}_l)^2}\\
&= \hat{\mathbf{c}}_l^T\sum_{j=1}^{N}\frac{\mathbf{l}_j\mathbf{r}_j^T}{(-\mu_j-\hat{\lambda}_l)^2}\hat{\mathbf{b}}_l\\
&= -\hat{\mathbf{c}}_l^T\mathbf{G}'(-\hat{\lambda}_l)\hat{\mathbf{b}}_l.
\end{aligned}$$

Hence,

$$\nabla_{\hat{\lambda}_l}\mathcal{J} = 2\hat{\mathbf{c}}_l^T\mathbf{G}(-\hat{\lambda}_l)\hat{\mathbf{b}}_l-2\hat{\mathbf{c}}_l^T\hat{\mathbf{H}}^T(-\hat{\lambda}_l)\hat{\mathbf{b}}_l.$$

Now we are able to state the $\mathcal{H}_2$ optimality conditions for the Problem 4.1.

Theorem 4.14 (Necessary condition of $\mathcal{H}_2$ optimality). *Let $\mathbf{G},\hat{\mathbf{H}}\in\mathcal{H}_2$ be two systems with semi-simple poles. Assume further that the ROM expressed as*

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

is a local minimizer of $\mathcal{J}_2 = \|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{H}_2}^2$. *Then the* $\mathcal{H}_2$ *optimality conditions*

$$\mathbf{G}(-\hat{\lambda}_k)\hat{\mathbf{b}}_k = \hat{\mathbf{H}}(-\hat{\lambda}_k)\hat{\mathbf{b}}_k,\ \hat{\mathbf{c}}_k^T\mathbf{G}(-\hat{\lambda}_k) = \hat{\mathbf{c}}_k^T\hat{\mathbf{H}}(-\hat{\lambda}_k), \tag{4.17}$$

and

$$\hat{\mathbf{c}}_k^T\mathbf{G}'(-\hat{\lambda}_k)\hat{\mathbf{b}}_k = \hat{\mathbf{c}}_k^T\hat{\mathbf{H}}'(-\hat{\lambda}_k)\hat{\mathbf{b}}_k, \tag{4.18}$$

hold.

Proof. The result is simply obtained by setting $\nabla_{\hat{\mathbf{c}}_l}\mathcal{J} = 0$, $\nabla_{\hat{\mathbf{b}}_l}\mathcal{J} = 0$ and $\nabla_{\hat{\lambda}_l}\mathcal{J} = 0$. □

Theorem 4.14 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 $\mathbf{G}$ at the opposite of the reduced-order model poles (see [Gugercin et al., 2008]). In the SISO case, the $\mathcal{H}_2$ optimality conditions are simpler and can be stated as follows:

Corollary 4.15 (Necessary condition of $\mathcal{H}_2$ optimality(SISO)). *Let* $\mathbf{G}, \hat{\mathbf{H}} \in \mathcal{H}_2$ *be two SISO models with semi-simple poles. Assume further that the reduced-system is expressed as*

$$\hat{\mathbf{H}}(s) = \sum_{k=1}^{n} \frac{\hat{\phi}_k}{s - \hat{\lambda}_k}$$

is a local minimizer of $\mathcal{J}_2 = \|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{H}_2}^2$. *Then the* $\mathcal{H}_2$ *optimality conditions*

$$\mathbf{G}(-\hat{\lambda}_k) = \hat{\mathbf{H}}(-\hat{\lambda}_k), \quad \text{and} \quad \mathbf{G}'(-\hat{\lambda}_k) = \hat{\mathbf{H}}'(-\hat{\lambda}_k) \tag{4.19}$$

hold, for $k = 1, \ldots, n$.

Hence, the optimality conditions for Problem 4.1 can be stated as a bitangential interpolation problem (see 3, Section 3.3). However, the parameters $\hat{\mathbf{c}}_l$, $\hat{\mathbf{b}}_l$ and $\hat{\lambda}_l$ are still unknowns.

In the following, we provide another demonstration of Theorem 4.14. In this proof, no particular structure is imposed on $\mathbf{G}$ and on $\hat{\mathbf{H}}$. This is a more general result and one consequence is to prove that Theorem 4.14 is also true when $\mathbf{G}$ is an irrational transfer function.

4.3.4 Optimality with respect to a general parameterization

In Subsection 4.3.3, we assumed that the reduced order model $\hat{\mathbf{H}}$ could be written as

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

Instead of doing that, let us suppose that $\hat{\mathbf{H}}$ depends on a vector of parameters $\hat{\mathbf{p}} = (\hat{p}_1, \ldots, \hat{p}_q) \in \mathbb{C}^q$. Here, the parameter dependency will not be explicit analytically. Then, the gradient of $\mathcal{J}_{\mathcal{H}_2}(\hat{\mathbf{H}})^2$ with respect to $\hat{\mathbf{p}}$ is given by

$$\begin{aligned} \nabla_{\hat{\mathbf{p}}}\|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{H}_2}^2 &= \nabla_{\hat{\mathbf{p}}}\langle \mathbf{G} - \hat{\mathbf{H}}, \mathbf{G} - \hat{\mathbf{H}}\rangle_{\mathcal{H}_2} \\ &= 2\langle \mathbf{G} - \hat{\mathbf{H}}, \nabla_{\hat{\mathbf{p}}}(\mathbf{G} - \hat{\mathbf{H}})\rangle_{\mathcal{H}_2} \\ &= -2\langle \mathbf{G} - \hat{\mathbf{H}}, \nabla_{\hat{\mathbf{p}}}\hat{\mathbf{H}}\rangle_{\mathcal{H}_2} \end{aligned} \tag{4.20}$$

where $\langle \mathbf{G} - \hat{\mathbf{H}}, \nabla_{\hat{\mathbf{p}}} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2}$ is an abuse of notation and stands for

$$\langle \mathbf{G} - \hat{\mathbf{H}}, \nabla_{\hat{\mathbf{p}}} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} = \left[\langle \mathbf{G} - \hat{\mathbf{H}}, \frac{\partial}{\partial \hat{p}_1} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} \quad \langle \mathbf{G} - \hat{\mathbf{H}}, \frac{\partial}{\partial \hat{p}_2} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} \quad \dots \quad \langle \mathbf{G} - \hat{\mathbf{H}}, \frac{\partial}{\partial \hat{p}_q} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} \right].$$

Thus, the local necessary optimization conditions with respect to $\hat{\mathbf{p}}$ are given by

$$\nabla_{\hat{\mathbf{p}}} \| \mathbf{G} - \hat{\mathbf{H}} \|_{\mathcal{H}_2}^2 = 0, \tag{4.21}$$

and the following result holds:

Proposition 4.16 (Necessary optimality condition with respect to a parameter). *Let $\mathbf{G} \in \mathcal{H}_2$ be a strictly proper model. In addition, let us consider for each parameter $\hat{\mathbf{p}} \in \mathbb{C}$ a model $\hat{\mathbf{H}}(\hat{\mathbf{p}})$, and the function that associates $\hat{\mathbf{p}} \to \hat{\mathbf{H}}(\hat{p})$ is differentiable. Then, if $\mathbf{p}^\star$ is the minimizer of $\| \mathbf{G} - \hat{\mathbf{H}}(\hat{\mathbf{p}}) \|$, it satisfies*

$$\langle \mathbf{G}, \nabla_{\hat{\mathbf{p}}} \hat{\mathbf{H}} \big|_{\hat{\mathbf{p}} = \mathbf{p}^\star} \rangle_{\mathcal{H}_2} = \langle \hat{\mathbf{H}}(\mathbf{p}^\star), \nabla_{\hat{\mathbf{p}}} \hat{\mathbf{H}} \big|_{\hat{\mathbf{p}} = \mathbf{p}^\star} \rangle_{\mathcal{H}_2} \tag{4.22}$$

Proof. This result follows by setting $\nabla_{\hat{\mathbf{p}}} \| \mathbf{G} - \hat{\mathbf{H}} \|_{\mathcal{H}_2}^2 = 0$ and using the equation (4.20). □

Proposition 4.16 enables us to derive the optimality conditions for a general parameter $\hat{\mathbf{p}}$. Notice that no structure was assumed on $\mathbf{G}$. In addition, it enables to obtain the $\mathcal{H}_2$ optimality conditions in a simpler way. First, we can assume that

$$\hat{\mathbf{H}} = \sum_{k=1}^{n} \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k}.$$

Thus, if one computes the gradient of $\hat{\mathbf{H}}$ with respect to $\hat{\mathbf{b}}_k$, we have

$$\begin{aligned} \nabla_{\hat{\mathbf{b}}_k} \hat{\mathbf{H}} &= \nabla_{\hat{\mathbf{b}}_k} \left(\frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k} \right) \\ &= \left[\frac{\hat{\mathbf{c}}_k e_1^T}{s - \hat{\lambda}_k} \quad \frac{\hat{\mathbf{c}}_k e_2^T}{s - \hat{\lambda}_k} \quad \dots \quad \frac{\hat{\mathbf{c}}_k e_{n_u}^T}{s - \hat{\lambda}_k} \right]. \end{aligned}$$

Hence, by writing the first order optimality conditions with respect to $\hat{\mathbf{b}}_k$ using Proposition 4.16, one should have

$$\begin{aligned} & \langle \mathbf{G}, \nabla_{\hat{\mathbf{b}}_k} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} = \langle \hat{\mathbf{H}}, \nabla_{\hat{\mathbf{b}}_k} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} \\ \Leftrightarrow \quad & \langle \mathbf{G}, \tfrac{\hat{\mathbf{c}}_k e_i^T}{s - \hat{\lambda}_k} \rangle_{\mathcal{H}_2} = \langle \hat{\mathbf{H}}, \tfrac{\hat{\mathbf{c}}_k e_i^T}{s - \hat{\lambda}_k} \rangle_{\mathcal{H}_2} \text{ for } i = 1, 2, \dots n_u \\ \Leftrightarrow \quad & \hat{\mathbf{c}}_k^T \mathbf{G}(-\hat{\lambda}_k) \mathbf{e}_i = \hat{\mathbf{c}}_k^T \hat{\mathbf{H}}(-\hat{\lambda}_k) \mathbf{e}_i \text{ for } i = 1, 2, \dots n_u \\ \Leftrightarrow \quad & \hat{\mathbf{c}}_k^T \mathbf{G}(-\hat{\lambda}_k) = \hat{\mathbf{c}}_k^T \hat{\mathbf{H}}(-\hat{\lambda}_k). \end{aligned}$$

In a very similar way, we obtain the right optimization conditions as follows :

$$\begin{aligned} \nabla_{\hat{\mathbf{c}}_k^T} \hat{\mathbf{H}} &= \nabla_{\hat{\mathbf{c}}_k^T} \left(\frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k} \right) \\ &= \left[\frac{\mathbf{e}_1 \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k} \quad \frac{\mathbf{e}_2 \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k} \quad \dots \quad \frac{\mathbf{e}_{n_y} \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k} \right], \end{aligned}$$

which implies

$$\begin{array}{rrcl} & \langle \mathbf{G}, \nabla_{\hat{\mathbf{c}}_k^T} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} & = & \langle \hat{\mathbf{H}}, \nabla_{\hat{\mathbf{c}}_k^T} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} \\ \Leftrightarrow & \langle \mathbf{G}, \frac{\mathbf{e}_i \hat{\mathbf{b}}_k^T}{s-\hat{\lambda}_k} \rangle_{\mathcal{H}_2} & = & \langle \hat{\mathbf{H}}, \frac{\mathbf{e}_i \hat{\mathbf{b}}_k^T}{s-\hat{\lambda}_k} \rangle_{\mathcal{H}_2} \text{ for } i = 1, 2, \dots n_y \\ \Leftrightarrow & \mathbf{e}_i^T \mathbf{G}(-\hat{\lambda}_k) \hat{\mathbf{b}}_k & = & \mathbf{e}_i^T \hat{\mathbf{H}}(-\hat{\lambda}_k) \hat{\mathbf{b}}_k \text{ for } i = 1, 2, \dots n_u \\ \Leftrightarrow & \mathbf{G}(-\hat{\lambda}_k) \hat{\mathbf{b}}_k & = & \hat{\mathbf{H}}(-\hat{\lambda}_k) \hat{\mathbf{b}}_k. \end{array}$$

Finally, since

$$\begin{aligned} \nabla_{\hat{\lambda}_k} \hat{\mathbf{H}} &= \nabla_{\hat{\lambda}_k} \left(\frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{s - \hat{\lambda}_k} \right) \\ &= -\frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{(s - \hat{\lambda}_k)^2}. \end{aligned}$$

the condition related to the derivative with respect to $\hat{\lambda}_k$ is obtained:

$$\begin{array}{rrcl} & \langle \mathbf{G}, \nabla_{\hat{\lambda}_k} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} & = & \langle \hat{\mathbf{H}}, \nabla_{\hat{\lambda}_k} \hat{\mathbf{H}} \rangle_{\mathcal{H}_2} \\ \Leftrightarrow & \langle \mathbf{G}, -\frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{(s-\hat{\lambda}_k)^2} \rangle_{\mathcal{H}_2} & = & \langle \hat{\mathbf{H}}, -\frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{(s-\hat{\lambda}_k)^2} \rangle_{\mathcal{H}_2} \\ \underbrace{\Leftrightarrow}_{\text{Hermitian derivative}} & \langle \mathbf{G}', \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{(s-\hat{\lambda}_k)} \rangle_{\mathcal{H}_2} & = & \langle \hat{\mathbf{H}}', \frac{\hat{\mathbf{c}}_k \hat{\mathbf{b}}_k^T}{(s-\hat{\lambda}_k)} \rangle_{\mathcal{H}_2} \\ \underbrace{\Leftrightarrow}_{\textit{Lemma 4.2}} & \hat{\mathbf{c}}_k^T \mathbf{G}'(-\hat{\lambda}_k) \hat{\mathbf{b}}_k & = & \hat{\mathbf{c}}_k^T \hat{\mathbf{H}}'(-\hat{\lambda}_k) \hat{\mathbf{b}}_k. \end{array}$$

Therefore, Theorem 4.14 was proved once again as a consequence of Proposition 4.16. At this time, no restriction was supposed on $\mathbf{G}$. Hence, this result is valid even if $\mathbf{G}$ is given by an irrational transfer function. In addition, Proposition 4.16 does not require any structure on the reduced order model $\hat{\mathbf{H}}$. This will be explored in Chapter 7.

In the next section, based on the optimality conditions provided in the tangential interpolation problem (see Chapter 3, Section 3.3 and 3.4), we recall one algorithm to find a reduced order model satisfying the optimality conditions (4.1) from Problem 4.1.

4.4 Fixed-point algorithms for $\mathcal{H}_2$ approximation

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

- ➤ the MIMO version of the Iterative Rational Krylov Algorithm (**IRKA**, see [Gugercin et al., 2008]).
- ➤ the MIMO version of the transfer function (**TF-IRKA**, see [Beattie and Gugercin, 2012]).

Both algorithms are based on the fixed point iteration. As it is shown in Theorem 4.14, the optimality conditions of Problem 4.1 are stated as bitangential interpolation conditions. The only inconvenient is that the optimal interpolation points and the tangential directions are not known a priori. In order to cope with this problem, one can choose randomly the interpolation points σ_i and the tangential directions $\mathbf{r}_i$ and $\mathbf{l}_i$ and repeat the following steps iteratively:

(i) Find $\hat{\mathbf{H}}$ satisfying

$$\begin{cases} \mathbf{G}(\sigma_i)\mathbf{r}_i &= \hat{\mathbf{H}}(\sigma_i)\mathbf{r}_i \\ \mathbf{l}_i^T\mathbf{G}(\sigma_i) &= \mathbf{l}_i^T\hat{\mathbf{H}}(\sigma_i) \\ \mathbf{l}_i^T\mathbf{G}'(\sigma_i)\mathbf{r}_i &= \mathbf{l}_i^T\hat{\mathbf{H}}'(\sigma_i)\mathbf{r}_i \end{cases}$$

(ii) Decompose $\hat{\mathbf{H}}(s) = \sum_{k=1}^{n} \frac{\hat{\mathbf{c}}_k\hat{\mathbf{b}}_k^T}{s-\hat{\lambda}_k}$.

(iii) Reset interpolation data: shift points as $\sigma_i \rightarrow= -\hat{\lambda}_i$, and tangential directions as $\mathbf{r}_i \leftarrow \hat{\mathbf{b}}_i$ and $\mathbf{l}_i \leftarrow \hat{\mathbf{c}}_i$, for $i = 1, \ldots, n$.

These steps enforce that the reduced order model should satisfy some similar conditions as those presented in Theorem 4.14. Hence, upon convergence, the reduced order model $\hat{\mathbf{H}}$ obtained satisfies the $\mathcal{H}_2$ optimality conditions of Theorem 4.14. Moreover, in [Krajewski et al., 1995], the authors show (for the SISO case) that saddle points and local maxima of the $\mathcal{H}_2$ minimization problem are know to be repellent to this iterative scheme. Hence, upon converge to a ROM $\hat{\mathbf{H}} \in \mathcal{H}_2$, this model is likely to be a local minimum.

The main difference between **IRKA** and **TF-IRKA** is that the former solves the tangential interpolation problem using projectors (see Theorem 3.10) while the second one uses the Loewner framework. **IRKA** is well adapted to find reduced order models for finite dimensional large-scale systems, because it exploits the projection framework to reduce the computational cost without computing explicitly the evaluations of the transfer function. In contrast, **TF-IRKA** is well adapted to find ROM for irrational transfer functions, requiring the evaluation of the transfer function on the interpolation points only. Both algorithms are detailed in what follows.

4.4.1 Iterative Rational Krylov Algorithm

This method has been proposed in [Gugercin et al., 2006] for SISO models based on the rational interpolation framework developed in [Grimme, 1997]. The MIMO extension has then been suggested in [Gugercin et al., 2008] based on the tangential interpolation framework developed in [Gallivan et al., 2004a]. 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 fulfills the first-order optimality conditions presented in Theorem 4.14. The method, called *Iterative Rational Krylov Algorithm* (**IRKA**), is recalled hereafter in Algorithm 1.

Algorithm 1 IRKA : [Gugercin et al., 2008]

1: **Input:** $\mathbf{G} = (E, A, B, C)$ of order N and reduced order n.
2: Make an initial choice for the shift points $\{\sigma_1^0, \ldots, \sigma_n^0\} \in \mathbb{C}$ initial interpolation points and tangential directions $\{\hat{\mathbf{b}}_{1,0}, \ldots, \hat{\mathbf{b}}_{n,0}\} \in \mathbb{C}^{n_u \times 1}$ and $\{\hat{\mathbf{c}}_{1,0}, \ldots, \hat{\mathbf{c}}_{n,0}\} \in \mathbb{C}^{n_y \times 1}$ closed by conjugation.
3: **while** not convergence **do**
4: Built $V_k = \begin{bmatrix}(A - \sigma_1^k E)^{-1} B \hat{\mathbf{b}}_{1,k}^T & \ldots & (A - \sigma_n^k E)^{-1} B \hat{\mathbf{b}}_{n,k}^T\end{bmatrix}$.
5: Built $W_k = \begin{bmatrix}(A^T - \sigma_1^k E^T)^{-1} C^T \hat{\mathbf{c}}_{1,k}^T & \ldots & (A^T - \sigma_n^k E^T)^{-1} C^T \hat{\mathbf{c}}_{n,k}^T\end{bmatrix}$.
6: $\hat{E} = W_k^T E V_k$, $\hat{A} = W_k^T A V_k$, $\hat{B} = W_k^T B$ and $\hat{C} = C V_k$
7: $k \leftarrow k + 1$.
8: Compute eigenvalue decomposition of $\hat{A}X = \hat{E}X\Sigma$, with $\Sigma = \mathbf{diag}(\hat{\lambda}_{1,k+1}, \ldots, \hat{\lambda}_{n,k+1})$.
9: Set $\sigma_i^{k+1} = -\hat{\lambda}_{i,k+1}$, new shift points.
10: Set $\begin{bmatrix}\hat{\mathbf{b}}_{1,k+1}^T, \ldots, \hat{\mathbf{b}}_{n,k+1}^T\end{bmatrix} = (\hat{E}X)^{-1}\hat{B}$ and $\begin{bmatrix}\hat{\mathbf{c}}_{1,k+1}, \ldots, \hat{\mathbf{c}}_{n,k+1}\end{bmatrix} = \hat{C}X$, new tangential directions.
11: **end while**
12: **Output:** $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ satisfying the $\mathcal{H}_2$ optimality conditions from Theorem 4.14.

Steps 1 and 2 define the inputs of the algorithm, which are the N-th order large-scale model $\mathbf{G} = (E, A, B, C)$, the order of the ROM n and an initial guess of both interpolation shifts and tangential directions.

Steps 3 to 9 correspond to the loop where the $\mathcal{H}_2$ optimality conditions are enforced. Firstly, the initial projectors V_k and W_k are constructed from the initial interpolation data from the previous iteration. Then, until the convergence is not reached, *i.e.*, the difference between $|\sigma_{i,k+1} - \sigma_{i,k}| \ll \varepsilon$, a reduced-order model is built by projection of the large-scale one in lines 5, 6 and 7. The eigenvalues and right eigenvectors of this reduced-order model are computed in step 7 and used in steps 8 in order to build the next interpolation points. At each iteration k, based on Theorem 3.10 (tangential interpolation), the reduced-order model $\hat{\mathbf{H}}_k = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ (step 6) tangentially interpolates the large-scale model at the opposite of the previous reduced-order model poles as it was sketched previously.

If Algorithm 1 converges, the model obtained satisfies the $\mathcal{H}_2$ first-order necessary conditions from Theorem 4.14. However, in general, there is neither any guarantee of convergence of the algorithm nor any certificate that the reduced order model is stable if the original was. Fortunately, extensive numerical applications of the algorithm tend to suggest that most of the time, it converges towards a stable system. Moreover, for SISO symmetric models, the article [Flagg et al., 2012] proves the convergence of the algorithm.

Notice that Algorithm 1 is only applicable in the case where $\mathbf{G}$ is a finite dimensional model. The next algorithm shows a procedure that extends this notion to the case of irrational transfer functions.

4.4.2 Fixed point algorithm based on transfer function evaluations

Based on the same fixed-point idea, an algorithm, named **TF-IRKA**, has been proposed in [Beattie and Gugercin, 2012] in order to obtain a reduced-order model that locally satisfies the optimality conditions for irrational transfer functions.

Algorithm 2 TF-IRKA [Beattie and Gugercin, 2012]

Initialization: Transfer function $\mathbf{G}(s)$ and reduced order n.
Make an initial choice for the shift points $\{\sigma_1^0, \ldots, \sigma_n^0\} \in \mathbb{C}$ initial interpolation points and tangential directions $\{\hat{\mathbf{b}}_{1,0}, \ldots, \hat{\mathbf{b}}_{n,0}\} \in \mathbb{C}^{n_u \times 1}$ and $\{\hat{\mathbf{c}}_{1,0}, \ldots, \hat{\mathbf{c}}_{n,0}\} \in \mathbb{C}^{n_y \times 1}$ closed by conjugation.
while not convergence **do**
    Build $\hat{E}$, $\hat{A}$, $\hat{B}$ and $\hat{C}$ satisfying Hermitian interpolation conditions using Loewner framework for the tangential interpolation data $\sigma_i^k, \hat{\mathbf{b}}_{i,k}^T, \hat{\mathbf{c}}_{i,k}$ (Theorem 3.16 from Chapter 3).
    $k \leftarrow k + 1$.
    Compute eigenvalue decomposition of $\hat{A}X = \hat{E}X\Sigma$, with $\Sigma = \mathbf{diag}(\hat{\lambda}_{1,k+1}, \ldots, \hat{\lambda}_{n,k+1})$.
    Set $\sigma_i^{k+1} = -\hat{\lambda}_{i,k+1}$, new shift points.
    Set $\left[\hat{\mathbf{b}}_{1,k+1}^T, \ldots, \hat{\mathbf{b}}_{n,k+1}^T\right] = (\hat{E}X)^{-1}\hat{B}$ and $\left[\hat{\mathbf{c}}_{1,k+1}, \ldots, \hat{\mathbf{c}}_{n,k+1}\right] = \hat{C}X$, new tangential directions.
end while
Output: $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ satisfying the $\mathcal{H}_2$ optimality conditions from Theorem 4.14.

This algorithm can be applied to any system for which the corresponding transfer function is available, regardless its dimension. When the transfer function is unknown, finite difference methods can be used to evaluate the derivative of the transfer function [Pontes Duff et al., 2015b]. However, there is neither any guarantee of convergence of the algorithm nor any certificate that the reduced order model is stable. Fortunately, extensive numerical applications of the algorithm tend to suggest that most of the time, it converges towards a stable system; see *e.g.* [Pontes Duff et al., 2015b].

Conclusion

In this chapter, the $\mathcal{H}_2$ optimal approximation problem has been introduced together with a non-exhaustive state-of-the-art. Additionally, the expression of the $\mathcal{H}_2$ inner product based on the pole-residue decomposition has been developed. Then, based on the properties of the $\mathcal{H}_2$ inner product, the $\mathcal{H}_2$ optimality conditions are derived in two different ways: (i) by developing the $\mathcal{H}_2$ approximation error with respect to $\hat{\mathbf{b}}_k, \hat{\mathbf{c}}_k$ and $\hat{\lambda}_k$; (2) and then, by deriving the $\mathcal{H}_2$ optimality conditions with respect to a general parametrization $\hat{\mathbf{p}}$ and then by applying the results to $\hat{\mathbf{b}}_k, \hat{\mathbf{c}}_k$ and $\hat{\lambda}_k$. In the last case, the Hermitian derivative property plays a very important role and the results are valid assuming very weak hypothesis over $\mathbf{G}$, including the case were $\mathbf{G}$ is an irrational function strictly proper in $\mathcal{H}_2$. Even if some of the results provided here are already well known, we believe that some of the proofs and development are new and relevant for the comprehension. Finally the two fixed-point iterative algorithms are presented in the last part of the chapter.

It is worth noticing that this chapter plays a pivotal role in this manuscript, since it links recent results on $\mathcal{H}_2$ model approximation while introducing the necessary elements required in the developments of the main contributions reported in this PhD, detailed in Chapters 5, 6 and 7.

Chapter 5

Optimal $\mathcal{H}_2$ model approximation by input/output-delay structured reduced order models

In this chapter, the $\mathcal{H}_2$ optimal approximation problem by a finite dimensional model including input/output delays, is addressed. Firstly, the $\mathcal{H}_2$ inner product formulas are revisited in the case where the reduced order models include input/output delays. Secondly, the approximation error is formulated as a function of the pole/residue decomposition. Then, by taking the gradient of the error, the $\mathcal{H}_2$-optimality conditions of the approximation problem are obtained as an extension of the tangential interpolatory conditions in the delay-free case. It is also demonstrated that for fixed delay values, this problem can be recast as a delay-free one. This approach followed in this first part is similar to what has been done in the delay-free case in Chapter 4. The similarities and the key differences from the delay-free case are highlighted. The results are compared with some simpler interpolation conditions. Finally, an iterative algorithm is sketched out and numerical results assess the theoretical contributions.

The chapter is organized as follows: first, the interest for the use of input/output delay reduced order models is exposed and the optimization problem is stated in Section 5.1. Initially, the theoretical results will be presented in the SISO case and a new expression of the $\mathcal{H}_2$ inner product in the presence of input and output delays is derived in Section 5.2. Then, Section 5.3 establishes the $\mathcal{H}_2$ optimality conditions solving the input/output delay dynamical model approximation problem. Then, results are extended to MIMO systems in Section 5.4. Finally, a fixed point algorithm is proposed allowing to compute such an approximation in practice. Section 5.5 details the results obtained after treating an academic example. The elements presented throughout this Chapter have been submitted in [Pontes Duff et al., 2016b].

Contents

5.1 Context and problem statement

Context

Input and output delays are generally used to model information lag, measurements and computational delays. In addition, for several physical models, *e.g.* transport phenomena, the information needs a finite amount of time to be transmitted from a point to another. Hence, these models have an intrinsic delay behavior even though they might not have an explicit delay in their representation.

In this context, an input/output delay approximation can be considered whenever the full order model represents a transport phenomenon. Examples of this kind of models are present in the domains of fluid mechanics, electronics and thermodynamics. Moreover, there exist several delay-dependent identification methods in the domains of chemistry [Bresch-Pietri et al., 2014] and hydroelectric channels [Litrico et al., 2010; Litrico and Fromion, 2009].

Another interesting academic examples are the models having very high order poles. This is shown in the following example :

Example 5.1. *Let us consider a SISO model* $\mathbf{G}_n \in \mathcal{H}_2$ *whose transfer function is given by*

$$\mathbf{G}_n(s) = \frac{C_n}{(s+1)^n}, \text{ for a given } n \in \mathbb{N},$$

where C_n *is chosen such that* $\|\mathbf{G}_n(s)\|_{\mathcal{H}_2}^2 = 1$. *Its impulse response is given by*

$$\mathbf{g}_n(t) = \frac{C_n}{n!} e^{-t} t^{n-1}.$$

The following figure sketches some of these impulse responses for $n = \{1, 6, 10, 20\}$.

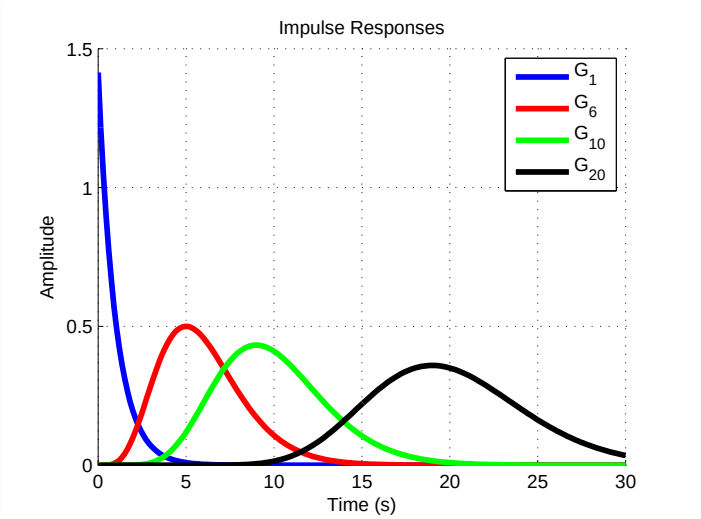

Figure 5.1: Impulse responses of $\mathbf{G}_n$ when $n = 1$ (blue line), $n = 6$ (red line), $n = 10$ (green line) and $n = 20$ (black line)

Hence, for n large enough, the model $\mathbf{G}_n$ has an intrinsic delay behavior and an input/output delay approximation is worth to be considered.

In this Chapter, the ($\mathcal{H}_2$ optimal) problem of approximating a given large-scale model by a low order one including (a priori unknown) input/output delays is addressed. To this aim an alternative pole/residue-based approach is developed, which enables to derive the $\mathcal{H}_2$-inner product expression in the presence of input and output delays, and then the $\mathcal{H}_2$ optimality conditions, treated as interpolation ones.

The presence of input/output delays in the approximation model was firstly tackled in [Halevi, 1996] (exploiting both Lyapunov equations and grammians properties derived in [Hyland and Bernstein, 1985] for the free-delay case). The bottleneck of this approach is that it requires to solve Lyapunov equations which might be costly in the large-scale context. From the moment matching side, [Scarciotti and Astolfi, 2014] proposed a problem formulation that enables the construction of an approximation which contains very rich delay structure (including state delay), but where the delays and the interpolation points are supposed to be a priori known. From the Loewner framework side, [Pontes Duff et al., 2015a] and after [Schulze and Unger, 2015] generalize the Loewner framework from [Mayo and Antoulas, 2007] to the state delay case enabling data-driven interpolation. The next chapter will be dedicated to this topic, *i.e.*, data interpolation with delay structure.

The main contribution of this chapter consists in extending the interpolation results of [Gugercin et al., 2008] to the approximation with an extended structure, namely, including non-zero input/output delays. Last but not least, an iterative algorithm and some numerical simulations illustrate the theoretical results developed.

Problem statement - SISO case

For simplicity, the derivation presented in this part is given in the context of SISO dynamical models, which already captures the idea of this result. Section 5.4 extends the results to the MIMO case. The main objective addressed in this chapter is to solve the following $\mathcal{H}_2$ approximation problem :

Problem 5.2. Input delay model $\mathcal{H}_2$-optimal approximation (SISO) *Given a stable N^{th} order SISO system $\mathbf{G} \in \mathcal{H}_2$, find a reduced n^{th} order (*s.t. *$n \ll N$) input-delay model $\hat{\mathbf{H}}_d^\star = \hat{\mathbf{H}} e^{-s\tau}$ s.t.:*

$$\hat{\mathbf{H}}_d^\star = \underset{\substack{\hat{\mathbf{H}}_d \in \mathcal{H}_2 \\ \dim(\hat{\mathbf{H}}_d) \leq n \\ \tau \in \mathbb{R}_+}}{\textbf{\textit{arg min}}} \ \|\mathbf{G} - \hat{\mathbf{H}}_d\|_{\mathcal{H}_2},$$

where $\hat{\mathbf{H}}_d = \hat{\mathbf{H}} e^{-s\tau} \in \mathcal{H}_2$. Once again we define the $\mathcal{H}_2$ mismatch error as $\mathcal{J}_{\mathcal{H}_2}(\hat{\mathbf{H}}) = \|\mathbf{G} - \hat{\mathbf{H}}_d\|_{\mathcal{H}_2}$.

As in the delay-free case, this search for an optimal solution will be carried out assuming that both $\mathbf{G}$ and $\hat{\mathbf{H}}$ from Problem 5.2 have semi-simple poles, *i.e.,* such that their respective transfer function matrix can be decomposed as follows:

$$\mathbf{G}(s) = \sum_{j=1}^{N} \frac{\psi_j}{s - \mu_j} \text{ and } \hat{\mathbf{H}}(s) = \sum_{k=1}^{n} \frac{\hat{\phi}_k}{s - \hat{\lambda}_k}, \tag{5.1}$$

where $\forall j = 1 \dots N,\ \forall k = 1 \dots n,\ \psi_j, \hat{\phi}_k \in \mathbb{C}$ and the poles $\mu_j, \hat{\lambda}_k$ are elements of $\mathbb{C}^-$ so that $\mathbf{G}$ and $\hat{\mathbf{H}}$ belong to $\mathcal{H}_2$ and $\dim(\mathbf{G}) = N$, $\dim(\hat{\mathbf{H}}) = n$.

It is worth mentioning that for a SISO system there is no difference between an input or output delay, due to the fact that the transfer function $\hat{\mathbf{H}}$ commutes with $e^{-s\tau}$, *i.e.*,

$$\hat{\mathbf{H}}(s) e^{-s\tau} = e^{-s\tau} \hat{\mathbf{H}}(s).$$

As in Chapter 4, this chapter will follow the hereafter three steps :

(i) Spectral characterization of the $\mathcal{H}_2$ inner product in the presence of input/output delays, presented in Section 5.2.

(ii) Derivation of the $\mathcal{H}_2$ necessary optimality conditions of Problem 5.2 using the $\mathcal{H}_2$ inner product characterization. This is the subject of Section 5.3.

(iii) Development of a fixed-point algorithm based on the $\mathcal{H}_2$ necessary optimality conditions in order to construct reduced order models in Section 5.5.

Let us firstly derive some $\mathcal{H}_2$ inner product properties specific to models having input delay.

5.2 Input/output delay $\mathcal{H}_2$ inner product

In this section, some elementary but important results, which will be useful along this chapter, are derived. First of all, a fundamental result dealing with the $\mathcal{H}_2$ norm invariance in case of input/output delay systems is presented.

Proposition 5.3. ($\mathcal{H}_2$ norm invariance) *Let $\hat{\mathbf{H}} \in \mathcal{H}_2$ be a stable dynamical system and $\sigma \in \mathcal{H}_\infty$ such that:*

$$\forall \omega \in \mathbb{R},\ \overline{\sigma(i\omega)}\sigma(i\omega) = 1. \tag{5.2}$$

If $\hat{\mathbf{H}}_d = \hat{\mathbf{H}}\sigma$ then $\|\hat{\mathbf{H}}_d\|_{\mathcal{H}_2} = \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}$.

Proof. If $\hat{\mathbf{H}}_d = \hat{\mathbf{H}}\sigma$, the $\mathcal{H}_2$ norm can be expressed as :

$$\begin{aligned}\|\hat{\mathbf{H}}_d\|^2_{\mathcal{H}_2} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\hat{\mathbf{H}}_d(i\omega)}\hat{\mathbf{H}}_d(i\omega)d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\hat{\mathbf{H}}(i\omega)}\,\overline{\sigma(i\omega)}\hat{\mathbf{H}}(i\omega)\sigma(i\omega)d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\hat{\mathbf{H}}(i\omega)}\,\underbrace{\overline{\sigma(i\omega)}\sigma(i\omega)}_{=1}\hat{\mathbf{H}}(i\omega)d\omega = \|\hat{\mathbf{H}}\|^2_{\mathcal{H}_2}\end{aligned}$$

□

One can easily check that condition (5.2) appearing in Proposition 5.14 is satisfied by an input/output delay, *i.e.*, if we take $\sigma(s) = e^{-s\tau}$, then $\overline{\sigma(i\omega)}\sigma(i\omega) = e^{i\omega\tau}e^{-i\omega\tau} = 1$. In other words, the $\mathcal{H}_2$ norm does not depend neither on the input, nor on the output delays. This is also true for MIMO systems if we have multiple input and output delays as stated in Section 5.4. Now let us derive the following lemma, analogous to Lemma 4.2 from Chapter 4, but in the presence of input/output delays.

Lemma 5.4 (First order $\mathcal{H}_2$ inner product). *Let $\mathbf{G} \in \mathcal{H}_2$ be a strictly proper real model whose transfer function can be expressed by*

$$\mathbf{G}(s) = \sum_{j=1}^{N} \frac{\psi_j}{s - \mu_j}.$$

In addition, suppose that $\hat{\lambda} \in \mathbb{C}^-$ and $\tau > 0$ is a delay. Then

$$\langle \mathbf{G}, \frac{1}{s-\hat{\lambda}}e^{-s\tau}\rangle_{\mathcal{H}_2} = \sum_{j=1}^{N} \frac{\psi_j e^{\mu_j \tau}}{-\hat{\lambda} - \mu_j}. \tag{5.3}$$

Proof. The proof uses similar ideas from the one of Lemma 4.2. Since, $\overline{\mathbf{G}(i\omega)} = \mathbf{G}(-i\omega)$ and that the only stable pole of the complex function $\mathbf{G}(-s)\frac{1}{s-\hat{\lambda}}e^{-s\tau}$ is $\hat{\lambda} \in \mathbb{C}^-$ (indeed, the poles of $\mathbf{G}(-s)$ are all unstable). However, in this case, one cannot use the same contour we have used in Lemma 4.2, *i.e.*,

$$\Gamma_C^{(1)} = \Gamma_I^{(1)} \cup \Gamma_R^{(1)}, \quad \text{where} \quad \begin{cases} \Gamma_I^{(1)} = \{s \in \mathbb{C}, s = i\omega \text{ and } \omega \in [-R; R],\ R \in \mathbb{R}_+\} \\ \Gamma_R^{(1)} = \{s \in \mathbb{C}, s = Re^{i\theta} \text{ where } \theta \in [\pi/2; 3\pi/2]\} \end{cases}.$$

The main reason is that $\mathbf{G}(-s)\frac{1}{s-\hat{\lambda}}e^{-s\tau}$ goes exponentially to infinite over Γ_R when R goes to infinity (because of the term $e^{-s\tau} \to \infty$ in the left half plane) . Therefore, the integral over this contour does not converge to the $\mathcal{H}_2$ inner product. In order to deal with this problem, let us consider an alternative contour Γ_C located in the right half plane encircling the poles of $\mathbf{G}(-s)$, *i.e.*,

$$\Gamma_C^{(2)} = \Gamma_I^{(2)} \cup \Gamma_R^{(2)},$$

with:

$$\begin{cases} \Gamma_I^{(2)} = \{s \in \mathbb{C}, s = i\omega \text{ and } \omega \in [-R; R],\ R \in \mathbb{R}_+\} \\ \Gamma_R^{(2)} = \{s \in \mathbb{C}, s = Re^{i\theta} \text{ where } \theta \in [\pi/2; -\pi/2]\} \end{cases}.$$

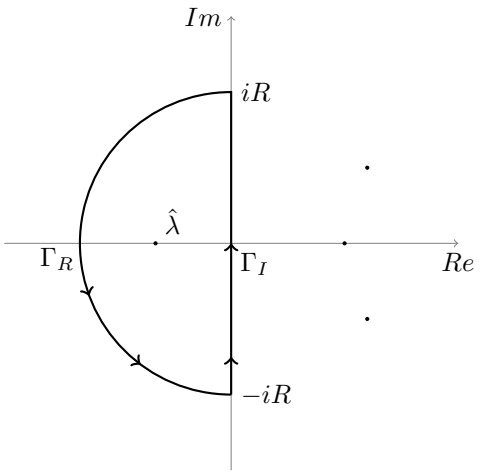

Figure 5.2: Wrong complex integral contour encircling the poles of $\dfrac{1}{s-\hat{\lambda}}e^{-s\tau}$

This contour is sketched on Figure 5.3. Now the function $\mathbf{G}(-s)\dfrac{1}{s-\hat{\lambda}}e^{-s\tau}$ goes to zero expo-

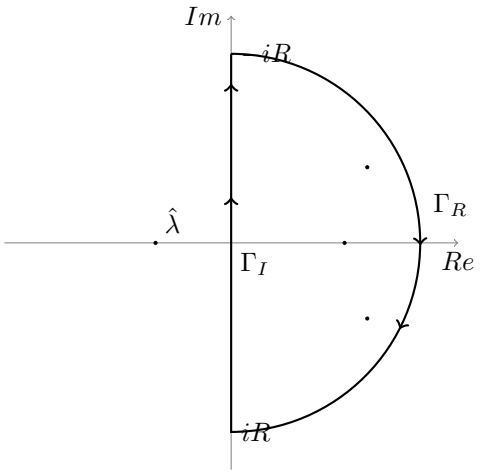

Figure 5.3: Complex integral contour encircling poles of $\mathbf{G}(-s)$.

nentially over Γ_R, when R goes to infinity. Once again, for a sufficient large radius value R, the Γ_C contour will contain all the poles of $\mathbf{G}(-s)$. Thus, by applying the residues theorem, it

follows that:

$$\begin{aligned}\langle \mathbf{G}, \frac{1}{s-\hat{\lambda}}\rangle_{\mathcal{H}_2} &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathbf{G}(-i\omega)\frac{1}{i\omega-\hat{\lambda}}e^{-i\omega\tau}\,d\omega \\ &= \lim_{R\to+\infty}\frac{1}{2i\pi}\int_{\Gamma_C}\mathbf{G}(-s)\frac{1}{s-\hat{\lambda}}e^{-s\tau}\,ds \\ &= \sum_{j=1}^{N}\mathrm{Res}\left(\mathbf{G}(-s)\frac{1}{s-\hat{\lambda}}e^{-s\tau}, -\mu_j\right) \\ &= \sum_{j=1}^{N}\frac{\psi_j e^{-\mu_j\tau}}{-\mu_j-\hat{\lambda}}.\end{aligned}$$

where Res(.) denotes the residue operator. Once again, the second equality line holds true since, when $R\to+\infty$:

$$\left|\int_{\Gamma_R}\mathbf{G}(-s)\frac{1}{s-\hat{\lambda}}e^{-s\tau}\,ds\right| \le \int_{\Gamma_R}\left\|\frac{1}{s-\hat{\lambda}}\right\|_F\left\|\mathbf{G}(-s)\right\|_F ds \to 0^+,$$

which concludes the first part of the proof. □

Lemma 5.4 provides a pole/residue formula for the $\mathcal{H}_2$ inner product in the presence of input/output delays. The main difference between the delay-free formula and the delayed one is that expression (5.3) involves poles and residues of $\mathbf{G}$, while the delay-free expression (4.2) only involves evaluations of $\mathbf{G}$. This difference comes from the fact that we are not able to use the same integral contour from Lemma 4.2 and we have to find an appropriate integral contour to handle the delay. It is worth to note that if $\tau = 0$, i.e. $e^{-s\tau} = 1$ on Lemma 5.4 , the result of Lemma 4.2 is recovered. Now, let us develop a more general formula as follows:

Theorem 5.5. Input/output delay $\mathcal{H}_2$ inner product computation *Let* $\mathbf{G}$, $\hat{\mathbf{H}}$ *be two SISO systems in* $\mathcal{H}_2$ *whose respective transfer function reads*

$$\mathbf{G}(s) = \sum_{j=1}^{N}\frac{\psi_j}{s-\mu_j} \text{ and } \hat{\mathbf{H}}(s) = \sum_{k=1}^{n}\frac{\hat{\phi}_k}{s-\hat{\lambda}_k},$$

and let $\tau > 0$. *If* $\hat{\mathbf{H}}_d = \hat{\mathbf{H}}e^{-s\tau}$, *the inner product* $\langle \mathbf{G}, \hat{\mathbf{H}}_d\rangle_{\mathcal{H}_2}$ *is given by:*

$$\langle \mathbf{G}, \hat{\mathbf{H}}_d\rangle_{\mathcal{H}_2} = \sum_{j=1}^{N}\hat{\mathbf{H}}(-\mu_j)\psi_j e^{\tau\mu_j}. \tag{5.4}$$

Proof. The result is a straightforward application of Lemma 5.4, as follows :

$$\langle \mathbf{G}, \hat{\mathbf{H}}e^{-s\tau}\rangle_{\mathcal{H}_2} = \langle \mathbf{G}, \sum_{k=1}^{n}\frac{\hat{\phi}_k}{s-\hat{\lambda}_k}e^{-s\tau}\rangle_{\mathcal{H}_2} \tag{5.5}$$

$$= \sum_{k=1}^{n}\langle \mathbf{G}, \frac{\hat{\phi}_k}{s-\hat{\lambda}_k}e^{-s\tau}\rangle_{\mathcal{H}_2} \tag{5.6}$$

$$\underbrace{=}_{\text{Lemma 5.4}} \sum_{k=1}^{n}\sum_{j=1}^{N}\hat{\phi}_k\frac{\psi_j e^{-\mu_j\tau}}{-\hat{\lambda}_k-\mu_j} = \sum_{j=1}^{N}\hat{\mathbf{H}}(-\mu_j)\psi_j e^{\tau\mu_j}. \tag{5.7}$$

which completes the proof. □

Theorem 5.5 enables to express the $\mathcal{H}_2$ inner product in the presence of input/output delays. The reader should remark that if $\hat{\mathbf{H}}_d$ is the model containing the input delay structure, then the inner product formula (5.4) will depend on the poles μ_k and residues ψ_k of $\mathbf{G}$, and on the evaluations of $\hat{\mathbf{H}}$. However, a symmetric formula involving the residues of $\hat{\mathbf{H}}$ and evaluations of $\mathbf{G}$ is **only** valid in the delay-free case (see Remark 4.5, Chapter 4). This is illustrated in the following Example.

Example 5.6 (**Non-symmetric expression of the $\mathcal{H}_2$ inner product**). ***Delay-free case :*** *Let* $\mathbf{G}(s) = \frac{1}{s+1} = \frac{\psi}{s-\mu}$, $\hat{\mathbf{H}}(s) = \frac{1}{s+2} = \frac{\hat{\phi}}{s-\hat{\lambda}}$. *One is able to compute the inner product between* $\mathbf{G}$ *and* $\hat{\mathbf{H}}$ *using both of these symmetric pole/residue expressions (see Remark 4.5, Chapter 4).*

$$\begin{aligned}\langle \hat{\mathbf{H}}, \mathbf{G}\rangle_{\mathcal{H}_2} &= \psi\hat{\mathbf{H}}(-\mu) = \psi\frac{\hat{\phi}}{-\mu-\hat{\lambda}} = \frac{1}{3} \text{ and}\\ \langle \mathbf{G}, \hat{\mathbf{H}}\rangle_{\mathcal{H}_2} &= \hat{\phi}\mathbf{G}(-\hat{\lambda}) = \hat{\phi}\frac{\psi}{-\hat{\lambda}-\mu} = \frac{1}{3}.\end{aligned}$$

Thus, in the delay-free context, the $\mathcal{H}_2$-inner product can be computed using either the poles-residues decomposition of $\mathbf{G}$ *or* $\hat{\mathbf{H}}$.

I/O delay case : *Now, let us define* $\hat{\mathbf{H}}_d(s) = \hat{\mathbf{H}}(s)e^{-s\tau}$ *with* $\tau = 1$. *Let us compute the $\mathcal{H}_2$-inner product between* $\hat{\mathbf{H}}_d$ *and* $\mathbf{G}$ *using Theorem 5.5 :*

$$\langle \mathbf{G}, \hat{\mathbf{H}}e^{-\tau s}\rangle_{\mathcal{H}_2} = \psi\hat{\mathbf{H}}(-\mu)e^{\tau\mu} = \psi\frac{\hat{\phi}}{-\mu-\hat{\lambda}}e^{\mu\tau} = \frac{1}{3}e^{-1}.$$

By noticing that, $\langle \mathbf{G}, \hat{\mathbf{H}}e^{-\tau s}\rangle_{\mathcal{H}_2} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{G}(-i\omega)\hat{\mathbf{H}}(i\omega)e^{-i\omega}d\omega = \langle \mathbf{G}e^{s}, \hat{\mathbf{H}}\rangle_{\mathcal{L}_2(i\mathbb{R})}$, *one applies the symmetric version as follows :*

$$\frac{1}{3}e^{-1} = \langle \mathbf{G}, \hat{\mathbf{H}}e^{-s}\rangle_{\mathcal{H}_2} = \underbrace{\langle \mathbf{G}e^{s}, \hat{\mathbf{H}}\rangle_{\mathcal{L}_2(i\mathbb{R})} \neq \hat{\phi}\mathbf{G}(-\hat{\lambda})e^{-\tau\hat{\lambda}}}_{\text{incorrect symmetric version}} = \hat{\phi}\frac{\psi}{-\hat{\lambda}-\mu}e^{-\hat{\lambda}\tau} = \frac{1}{3}e^{2}.$$

Hence a symmetric version of the $\mathcal{H}_2$-inner product does not provide the same result anymore. It is remarkable that $\mathbf{G}e^{s} \notin \mathcal{H}_2$ *because it is not analytic at infinite. This is the justification of the non-symmetry in the inner product expression.*

This non-symmetry in the expression of the $\mathcal{H}_2$-inner product in the presence of input/output delays is the crucial point in the derivation of the optimality conditions. Indeed, in the delay-free context (see [Gugercin et al., 2008]), it is assumed that we have this symmetric form and this property plays a crucial role in the $\mathcal{H}_2$-optimality conditions derivation.

To conclude the inner product discussion, the next example shows what happens when both models $\hat{\mathbf{H}}$ and $\mathbf{G}$ carry an input/output delay.

Example 5.7 (**Both models carrying an input/output delay**). *Let us consider* $\mathbf{H}(s) = \frac{1}{s+1}$ *and* $\mathbf{G}(s) = \frac{1}{s+2}$ *and one wants to compute the $\mathcal{H}_2$ inner product* $\langle \mathbf{H}e^{-s}, \mathbf{G}e^{-s\tau}\rangle_{\mathcal{H}_2}$

as a function of $\tau \geq 0$. *Firstly, let us look at the computation when* $\tau < 1$. *In this case,*

$$\langle \mathbf{H}e^{-s}, \mathbf{G}e^{-s\tau} \rangle_{\mathcal{H}_2} = \langle \mathbf{H}e^{-(1-\tau)s}, \mathbf{G} \rangle_{\mathcal{H}_2} = \mathbf{H}(2)e^{-(1-\tau)2},$$

and the inner product can be expressed using the partial fraction decomposition of $\mathbf{G}$.

However, we would like to stress that this formula is valid if $\tau \leq 1$ *only. In the case where* $\tau > 1$, *we have*

$$\langle \mathbf{H}e^{-s}, \mathbf{G}e^{-s\tau} \rangle_{\mathcal{H}_2} = \langle \mathbf{H}, \mathbf{G}e^{-s(\tau-1)} \rangle_{\mathcal{H}_2} = \mathbf{G}(1)e^{-(\tau-1)}.$$

Hence, in this case, the inner product can only be expressed using the partial fraction decomposition of $\mathbf{H}$ *(because of the exponential domination).*

The following Figure 5.4 sketches the plot of $f(\tau) = \langle \mathbf{H}e^{-s}, \mathbf{G}e^{-s\tau} \rangle_{\mathcal{H}_2}$, *for* $0 \leq \tau \leq 2$.

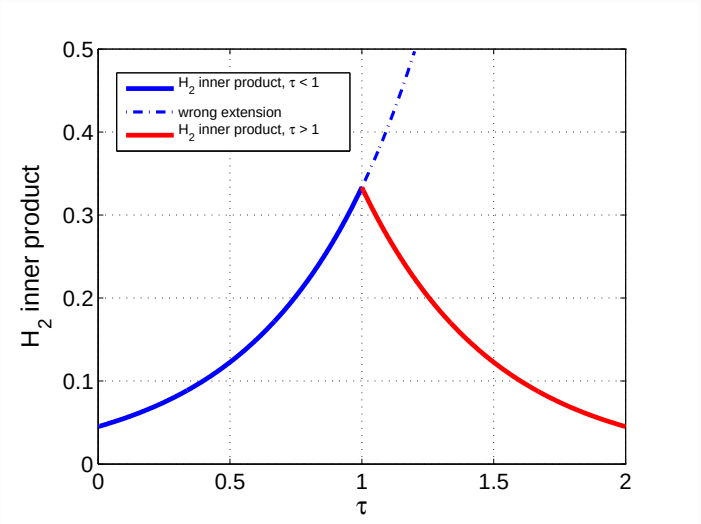

Figure 5.4: Graph of $f(\tau) = \langle \mathbf{H}e^{-s}, \mathbf{G}e^{-s\tau} \rangle_{\mathcal{H}_2}$ for $\tau \in [0, 2]$. Notice that when $\tau < 1$, the $f(\tau)$ is increasing, while when $\tau > 1$, the $f(\tau)$ is decreasing.

With reference to Figure 5.4, one should note that $f(\tau)$ *is increasing when* $\tau \leq 1$ *and decreasing when* $\tau > 1$. *Indeed, this can be easily observed by looking at the time-domain inner product and the action of the delay operator as shown Figure 5.5* .

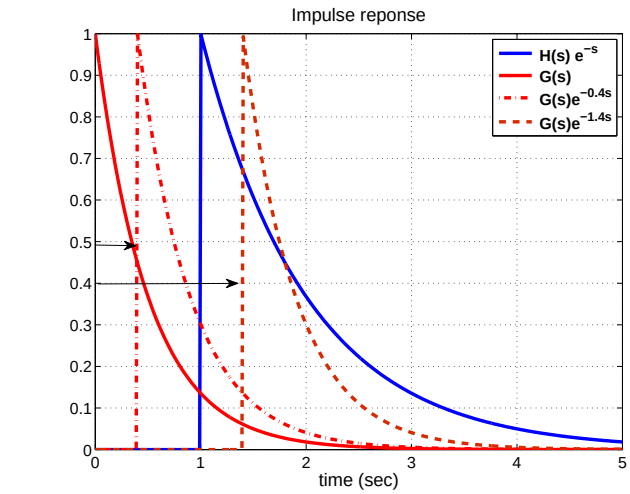

Figure 5.5: Impulse response of $\mathbf{H}e^{-s}$, $\mathbf{G}$, $\mathbf{G}e^{-0.4s}$ and $\mathbf{G}e^{-1.4s}$.

Figure 5.5 sketches what happens when the value of τ varies. Increasing τ shifts the impulse response $\mathbf{g}(t)$ to the right. Since, from Plancherel's theorem,

$$f(\tau) = \langle \mathbf{H}e^{-s}, \mathbf{G}e^{-s\tau} \rangle_{\mathcal{H}_2} = \langle \mathbf{h}(t-1), \mathbf{g}(t-\tau) \rangle_{L_2}.$$

*Hence, one can conclude graphically that (i) the maximum value of the inner product is obtained when $\mathbf{g}$ is exactly the same as the delay on $\mathbf{h}$ (*i.e.*, $\tau = 1$), (ii) the inner product must decrease for $\tau > 1$ and (iii) the inner product must increase for $\tau < 1$.*

To sum up, this section has developed the pole-residue expression of the $\mathcal{H}_2$ inner product in the presence of input/output delays and highlighted the main differences with the delay-free case. In the next section, we tackle the input delay $\mathcal{H}_2$ approximation Problem 5.2 using these expressions.

5.3 Formulation of the input/output delay $\mathcal{H}_2$ first order optimality conditions

This section aims at presenting the SISO $\mathcal{H}_2$ model approximation problem when the reduced order model has an input-output delay structure and at expressing the $\mathcal{H}_2$-approximation error as a function of the poles, residues and delay. Section 5.4 provides the generalization of these results in the MIMO context.

5.3.1 $\mathcal{H}_2$ approximation error

The following proposition makes now explicit the calculation of the $\mathcal{H}_2$ norm associated with the dynamical mismatch gap $\|\mathbf{G} - \hat{\mathbf{H}}_d\|_{\mathcal{H}_2}$, which characterizes Problem 5.2 criterion. In order to simplify the notations, let us denote $\mathcal{J}_2 = \mathcal{J}_{\mathcal{H}_2}^2 = \|\mathbf{G} - \hat{\mathbf{H}}_d\|_{\mathcal{H}_2}^2$.

Proposition 5.8. *Let* $\mathbf{G},\ \hat{\mathbf{H}}_d \in \mathcal{H}_2$, *such that* $\hat{\mathbf{H}}_d = \hat{\mathbf{H}}e^{-s\tau}$. *The* $\mathcal{H}_2$ *norm of* $\mathcal{J}_2$, *can be expressed as:*

$$\begin{aligned} \mathcal{J}_2 &= \|\mathbf{G} - \hat{\mathbf{H}}_d\|_{\mathcal{H}_2}^2 \\ &= \|\mathbf{G}\|_{\mathcal{H}_2}^2 - 2\langle \mathbf{G}, \hat{\mathbf{H}}e^{-s\tau}\rangle_{\mathcal{H}_2} + \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2. \end{aligned} \tag{5.8}$$

Proof. Simply develop the $\mathcal{H}_2$ norm using the $\mathcal{H}_2$ inner product definition and exploit the previous result $\|\hat{\mathbf{H}}e^{-s\tau}\|_{\mathcal{H}_2} = \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}$. □

Obviously, regarding (5.8), minimizing $\mathcal{J}_2$ is equivalent to minimize $-2\langle \mathbf{G}, \hat{\mathbf{H}}e^{-s\tau}\rangle_{\mathcal{H}_2} + \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2$ and thus look for the optimal values of the decision variables contained in both the realization $\hat{\mathbf{H}} \in \mathcal{H}_2$ and the delay $\tau > 0$. The pole/residue expression of the $\mathcal{H}_2$ inner product given in Section 5.2 enables to parametrize the mismatch error $\mathcal{J}_2$ as s function of the delay $\tau > 0$ and the minimal parameters defining the reduced order model $\hat{\mathbf{H}}$, *i.e.*, its poles and residues. This expression is given in the following proposition.

Proposition 5.9. *Let us assume the same hypothesis from Proposition 5.8. In addition,* $\mathbf{G}$ *and* $\hat{\mathbf{H}}$ *are given by*

$$\mathbf{G}(s) = \sum_{j=1}^{N} \frac{\psi_j}{s - \mu_j} \text{ and } \hat{\mathbf{H}}(s) = \sum_{k=1}^{n} \frac{\hat{\phi}_k}{s - \hat{\lambda}_k}.$$

Then, $\mathcal{H}_2$*-mismatch error can be expressed as*

$$\mathcal{J}_2 = \|\mathbf{G}\|_{\mathcal{H}_2}^2 - 2\sum_{k=1}^{N} \hat{\mathbf{H}}(-\mu_k)e^{\mu_k\tau}\psi_k + \sum_{k=1}^{n} \hat{\phi}_k\hat{\mathbf{H}}(-\hat{\lambda}_k). \tag{5.9}$$

Proof. One should use the spectral formulation of the $\mathcal{H}_2$ inner product in the presence of delays from Section 5.2 and the result follows. □

It is worth noticing that the last two terms of the right hand side in the expression of $\mathcal{J}_2$ depends on the evaluations of $\hat{\mathbf{H}}$, while in the delay-free case, one of them depends on the evaluations of $\mathbf{G}$ (see Section 4.3 in Chapter 4). This is a consequence of the nonsymmetric expression for the $\mathcal{H}_2$ inner product in the presence of input/output delays (see Example 5.6), which states that the $\mathcal{H}_2$ inner product $\langle \mathbf{G}, \hat{\mathbf{H}}e^{-s\tau}\rangle_{\mathcal{H}_2}$ can only be expressed as a function of the poles and residues of $\mathbf{G}$ and the evaluations of $\hat{\mathbf{H}}$. This remark, which plays a very important role here, is the origin of the main difference between Problem 5.2 and the delay-free model approximation Problem 4.1 from Chapter 4.

5.3.2 Gradient with respect to poles, residues and delays

Now, we are able to derive the input/output first-order $\mathcal{H}_2$-optimality conditions for Problem 5.2. Let us compute the gradient of $\mathcal{J}_2$ with respect to the parameters of $\hat{\mathbf{H}}_d$.

Gradient with respect to $\hat{\phi}_l$: let us compute the gradient of $\mathcal{J}_2$ with respect to $\hat{\phi}_l$. First, let's recall the gradient of $\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2$ with respect $\hat{\phi}_l$ from Chapter 4, Section 4.3 :

$$\nabla_{\hat{\phi}_l}\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2 = \nabla_{\hat{\phi}_l}\left(\sum_{k=1}^{n} \hat{\phi}_k\hat{\mathbf{H}}(-\hat{\lambda}_k)\right) = 2\hat{\mathbf{H}}(-\hat{\lambda}_l)$$

In a similar way, we now compute the gradient of $\langle \mathbf{G}, \hat{\mathbf{H}} e^{-s\tau} \rangle_{\mathcal{H}_2}$ with respect to $\hat{\phi}_l$ as:

$$\nabla_{\hat{\phi}_l} \langle \mathbf{G}, \hat{\mathbf{H}} e^{-s\tau} \rangle_{\mathcal{H}_2} = \nabla_{\hat{\phi}_l} \left(\sum_{k=1}^{N} \hat{\mathbf{H}}(-\mu_k) e^{\mu_k \tau} \psi_k \right) = \sum_{k=1}^{N} \frac{\psi_k}{-\mu_k - \lambda_l} e^{\mu_k \tau}.$$

Hence, the gradient of $\mathcal{J}_2$ with respect to $\hat{\phi}_l$ is given by:

$$\nabla_{\hat{\phi}_l} \mathcal{J}_2 = -2 \sum_{k=1}^{N} \frac{\psi_k}{-\mu_k - \lambda_l} e^{\mu_k \tau} + 2\hat{\mathbf{H}}(-\hat{\lambda}_l). \tag{5.10}$$

Gradient with respect to $\hat{\lambda}_l$: now, let us compute the gradient of $\mathcal{J}_2$ with respect to $\hat{\lambda}_l$. First, let's recall the gradient of $\|\mathbf{H}\|^2_{\mathcal{H}_2}$ with respect $\hat{\lambda}_l$ from Chapter 4, Section 4.3 :

$$\nabla_{\hat{\lambda}_l} \|\hat{\mathbf{H}}\|^2_{\mathcal{H}_2} = \nabla_{\hat{\lambda}_l} \left(\sum_{k=1}^{n} \hat{\phi}_k \hat{\mathbf{H}}(-\hat{\lambda}_k) \right) = -2\hat{\phi}_l \hat{\mathbf{H}}'(-\hat{\lambda}_l)$$

In a similar way, we compute the gradient of $\langle \mathbf{G}, \hat{\mathbf{H}} e^{-s\tau} \rangle_{\mathcal{H}_2}$ with respect to $\hat{\lambda}_l$ as:

$$\nabla_{\hat{\lambda}_l} \langle \mathbf{G}, \hat{\mathbf{H}} e^{-s\tau} \rangle_{\mathcal{H}_2} = \nabla_{\hat{\lambda}_l} \left(\sum_{k=1}^{N} \hat{\mathbf{H}}(-\mu_k) e^{\mu_k \tau} \psi_k \right) = \sum_{k=1}^{N} \hat{\phi}_l \frac{\psi_k}{(-\mu_k - \lambda_l)^2} e^{\mu_k \tau}.$$

Hence, the gradient of $\mathcal{J}_2$ with respect to $\hat{\lambda}_l$ is given by:

$$\nabla_{\hat{\lambda}_l} \mathcal{J}_2 = 2\hat{\phi}_l \left[-\sum_{k=1}^{N} \frac{\psi_k}{(-\mu_k - \lambda_m)^2} e^{\mu_k \tau} - \hat{\mathbf{H}}'(-\hat{\lambda}_l) \right]. \tag{5.11}$$

Gradient with respect to τ : Finally, the only term of $\mathcal{J}_2$ depending on the delay is $\langle \mathbf{G}, \hat{\mathbf{H}} e^{-s\tau} \rangle_{\mathcal{H}_2}$. Thus, the gradient of $\mathcal{J}_2$ with respect to τ is computed as:

$$\nabla_\tau \mathcal{J}_2 = -2\nabla_\tau \langle \mathbf{G}, \hat{\mathbf{H}} e^{-s\tau} \rangle_{\mathcal{H}_2} = -2 \sum_{k=1}^{N} \mu_k \hat{\mathbf{H}}(-\mu_k) e^{\mu_k \tau} \psi_k$$

Now we are ready to state the input/output delay $\mathcal{H}_2$ optimality conditions in the SISO case.

Theorem 5.10. **($\mathcal{H}_2$ optimality conditions in the SISO case)** *Considering* $\mathbf{G}(s) = \sum_{j=1}^{N} \frac{\psi_j}{s - \mu_j}$, $\hat{\mathbf{H}}(s) = \sum_{k=1}^{n} \frac{\hat{\phi}_k}{s - \hat{\lambda}_k}$, *such that* $\hat{\mathbf{H}}_d = \hat{\mathbf{H}} e^{-\tau s}$ *is a local optimum of Problem 5.2, then the following conditions hold:*

$$\begin{cases} \hat{\mathbf{H}}(-\hat{\lambda}_k) = \tilde{\mathbf{G}}(-\hat{\lambda}_k), \\ \hat{\mathbf{H}}'(-\hat{\lambda}_k) = \tilde{\mathbf{G}}'(-\hat{\lambda}_k), \end{cases} \tag{5.12}$$

$$\sum_{j=1}^{N} \mu_j \psi_j \left(\sum_{k=1}^{n} \frac{\hat{\phi}_k}{\mu_j + \hat{\lambda}_k} \right) e^{\tau \mu_j} = 0, \tag{5.13}$$

for all $k = 1 \dots n$, *and where*

$$\tilde{\mathbf{G}}(s) = \sum_{j=1}^{N} \frac{\psi_j}{s - \mu_j} e^{\tau \mu_j}. \tag{5.14}$$

Proof. The stationary conditions hold as soon as we set $\nabla_{\hat{\phi}_m} \mathcal{J}_2 = 0, \nabla_{\hat{\lambda}_m} \mathcal{J}_2 = 0$ and $\nabla_\tau \mathcal{J}_2 = 0$. In addition, by noticing that

$$\tilde{\mathbf{G}}(-\hat{\lambda}_l) = \sum_{k=1}^{N} \frac{\psi_k}{-\mu_k - \hat{\lambda}_l} e^{\mu_k \tau} \quad \text{and} \quad \tilde{\mathbf{G}}'(-\hat{\lambda}_l) = -\sum_{k=1}^{N} \frac{\psi_k}{(-\mu_k - \hat{\lambda}_l)^2} e^{\mu_k \tau},$$

the results follow. □

Theorem 5.10 asserts that any solution of the $\mathcal{H}_2$ model approximation Problem 5.13, denoted by $\hat{\mathbf{H}}_d = \hat{\mathbf{H}} e^{-s\hat{\tau}}$ is such that $\hat{\mathbf{H}}$ satisfies, at the same time, a set of interpolation conditions detailed in (5.12) and an another relation on the delay given by (5.13). Moreover, as in the delay-free case, (5.12) corresponds to interpolation conditions. However, the main difference is that instead of interpolating the full order model $\mathbf{G}$, a new model $\tilde{\mathbf{G}}$ is here considered. This new model possesses the same poles as $\mathbf{G}$, but differs from $\mathbf{G}$ by its residues, computed as in (5.14). One should note that this new model is intrinsically related to the non-symmetric expression of the $\mathcal{H}_2$-inner product (see Example 5.6).

Some considerations

The following remark shows that a simple straightforward reasoning can lead to false optimality conditions.

Remark 5.11. *(***False optimality conditions***)* *Let us consider:*

$$\mathcal{J} = \|\mathbf{G} - \hat{\mathbf{H}} e^{-s\tau}\|_{\mathcal{H}_2} = \|\overbrace{\mathbf{G} e^{s\tau}}^{\mathbf{G}_{NC}} - \hat{\mathbf{H}}\|_{\mathcal{H}_2},$$

where $\mathbf{G}_{NC} = \mathbf{G} e^{s\tau}$ *is a new full order model (and it is a* ***N****on-****C****ausal model). Then the model approximation problem can be written as*

$$\hat{\mathbf{H}}^\star = \underset{\dim(\hat{\mathbf{H}}) \leq n}{\boldsymbol{arg\ min}} \|\mathbf{G}_{NC} - \hat{\mathbf{H}}\|_{\mathcal{H}_2}.$$

Since in the last expression there is no intervention of the delay, one could presume that the optimality conditions would be those presented in [Gugercin et al., 2008], i.e.

$$\hat{\mathbf{H}}(-\lambda_k) = \mathbf{G}_{NC}(-\lambda_k) \quad \text{and} \quad \hat{\mathbf{H}}'(-\lambda_k) = \mathbf{G}'_{NC}(-\lambda_k), \quad k = 1, \dots, n, \tag{5.15}$$

where λ_k *is a pole of* $\hat{\mathbf{H}}$ *and* $\mathbf{G}_{NC}(s) = \mathbf{G} e^{s\tau}$. *Unfortunately, this result is not true, since in the delay-free* $\mathcal{H}_2$ *optimality conditions derivation, the symmetry in the pole-residue* $\mathcal{H}_2$*-inner-product expression is required.*

The following example illustrates why conditions (5.15) do not provide a good approximation and compare them with those derived in Theorem 5.10.

Example 5.12. *Let us consider*

$$\mathbf{G}(s) = \frac{1}{(s+1)(s+2)} \in \mathcal{H}_2,$$

*a SISO dynamical model of order $n = 2$, for which we are seeking an input-delay model approximation. Let us consider the delay to be fixed at $\tau = 0.3$. By means of conditions (5.15) and using the fixed-point algorithm **TF-IRKA** [Beattie and Gugercin, 2012], one is able to construct an input delay approximation*

$$\tilde{\mathbf{H}}_d(s) = \frac{\tilde{\phi}}{(s-\tilde{\lambda})} e^{-\tau s}$$

where $\tilde{\phi} \approx 0.5165$ and $\tilde{\lambda} \approx -1.189$. Now, we construct

$$\hat{\mathbf{H}}_d(s) = \frac{\hat{\phi}}{(s-\hat{\lambda})} e^{-\tau s}$$

satisfying the optimality conditions from (5.12) where we obtained $\hat{\phi} \approx 0.3066$ and $\hat{\lambda} \approx -0.6292$. The impulse response of the original system and the two reduced order models are shown in the following Figure 5.12.

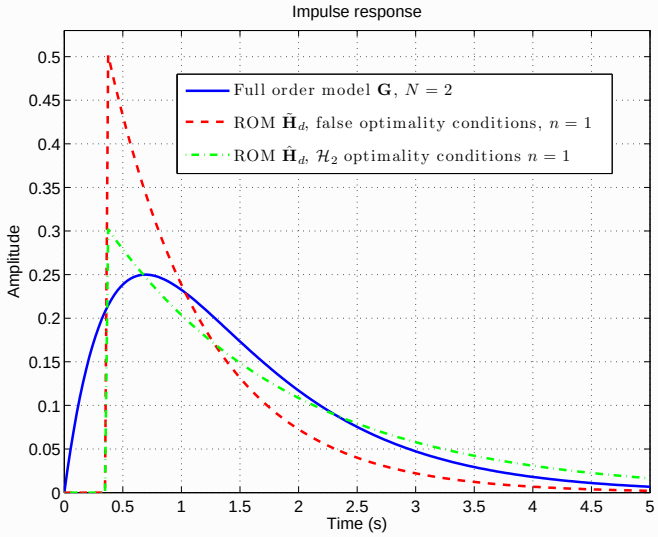

Figure 5.6: Impulse response of the original model $\mathbf{G}$ of order $N = 2$ (solid blue line), the input model approximation $\tilde{\mathbf{H}}_d$ of order $n = 1$ from Example 2 (dashed red line) and the optimal input model approximation $\hat{\mathbf{H}}_d$ of order $n = 1$ (dash-dotted green line) satisfying conditions (5.12).

It is clear that the model satisfying the optimality conditions derived in this chapter provides a better approximation of $\mathbf{G}$.

Up to now, both $\hat{\mathbf{H}}_d$ and $\mathbf{G}$ were consider to be SISO systems and only one input-delay was considered. In next section, all the results obtained here are generalized to the multiple input/output delay MIMO case.

5.4 Extension to MIMO models

This section generalizes the results obtained for SISO models to multiple input multiple output delay models. The results will be briefly detailed, since they are very similar to the SISO case.

$\mathcal{H}_2$ model approximation for MIMO systems

Let $\hat{\mathbf{H}}_d$ be a multiple-input/output delays **MIMO** system *s.t.* $\hat{\mathbf{H}}_d(s) \in \mathcal{H}_2$ and represented by:

$$\hat{\mathbf{H}}_d : \left\{ \begin{array}{rcl} \hat{E}\dot{\hat{\mathbf{x}}}(t) & = & \hat{A}\hat{\mathbf{x}}(t) + \hat{B}\hat{\mathbf{\Delta}}_i(\mathbf{u}(t)) \\ \hat{\mathbf{y}}(t) & = & \hat{\mathbf{\Delta}}_o(\hat{C}\hat{\mathbf{x}}(t)) \end{array} \right. , \tag{5.16}$$

where $\hat{E}, \hat{A} \in \mathbb{R}^{n\times n}$ (with state dimension $n \in \mathbb{N}^*$), $\hat{B} \in \mathbb{R}^{n\times n_u}$, $\hat{C} \in \mathbb{R}^{n_y\times n}$ and $\hat{\mathbf{\Delta}}_i$ and $\hat{\mathbf{\Delta}}_o$ are the delay operators defined as follows

$$\left\{ \begin{array}{rcl} \hat{\mathbf{\Delta}}_i(s) & = & \mathbf{diag}(e^{-s\hat{\tau}_1} \ \ldots \ e^{-s\hat{\tau}_{n_u}}) \in \mathcal{H}_\infty^{n_u\times n_u} \\ \hat{\mathbf{\Delta}}_o(s) & = & \mathbf{diag}(e^{-s\hat{\gamma}_1} \ \ldots \ e^{-s\hat{\gamma}_{n_y}}) \in \mathcal{H}_\infty^{n_y\times n_y}. \end{array} \right. \tag{5.17}$$

The matrix transfer functions $\hat{\mathbf{\Delta}}_i(s)$ and $\hat{\mathbf{\Delta}}_o(s)$ defined in (5.17) represent the frequency behavior of the delay operators $\mathbf{\Delta}_i$ and $\mathbf{\Delta}_o$ respectively. We recall that the transfer function of $\hat{\mathbf{H}}_d$ is given by

$$\hat{\mathbf{H}}_d(s) = \hat{\mathbf{\Delta}}_o(s)\hat{\mathbf{H}}(s)\hat{\mathbf{\Delta}}_i(s) \in \mathcal{H}_2. \tag{5.18}$$

Recall that we denote by $\hat{\mathbf{H}}_d = (\hat{\mathbf{E}}, \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{\Delta}}_i, \hat{\mathbf{\Delta}}_o)$ a **MIMO** input/output delayed system represented by the transfer matrix (5.18). $\hat{\mathbf{H}}_d$ will also be said to have order $n \ll N$ (where N is the original model order), shortly denoted by $\dim(\hat{\mathbf{H}}_d) = n$.

The main objective addressed is to solve the following $\mathcal{H}_2$ approximation problem:

Problem 5.13. (Delay model MIMO $\mathcal{H}_2$-optimal approximation) *Given a stable N^{th} order system $\mathbf{G} \in \mathcal{H}_2^{n_y\times n_u}$, find a reduced n^{th} order (*s.t. *$n \ll N$) multiple-input/output delays model $\hat{\mathbf{H}}_d^\star = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \hat{\mathbf{\Delta}}_i, \hat{\mathbf{\Delta}}_o)$ s.t.:*

$$\hat{\mathbf{H}}_d^\star = \underset{\substack{\hat{\mathbf{H}}_d \in \mathcal{H}_2 \\ \dim(\hat{\mathbf{H}}_d) \leq n}}{\boldsymbol{arg\ min}} \ \|\mathbf{G} - \hat{\mathbf{H}}_d\|_{\mathcal{H}_2}, \tag{5.19}$$

where $\hat{\mathbf{H}}_d = \hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i$, as in (5.18).

Once again, this search for an optimal solution will be carried out assuming that both $\mathbf{G}$ and $\hat{\mathbf{H}}$ from (5.19) have semi-simple poles, *i.e.,* such that their respective transfer function matrix can be decomposed as follows:

$$\mathbf{G}(s) = \sum_{j=1}^{N} \frac{\mathbf{l}_j\mathbf{r}_j^T}{s-\mu_j} \text{ and } \hat{\mathbf{H}}(s) = \sum_{k=1}^{n} \frac{\hat{\mathbf{c}}_k\hat{\mathbf{b}}_k^T}{s-\hat{\lambda}_k}, \tag{5.20}$$

where $\forall j = 1\ldots N$, $\forall k = 1\ldots n$, $\mathbf{r}_j, \hat{\mathbf{b}}_k \in \mathbb{C}^{n_u\times 1}$ and $\mathbf{l}_j, \hat{\mathbf{c}}_k \in \mathbb{C}^{n_y\times 1}$. The poles $\mu_j, \hat{\lambda}_k$ are elements of $\mathbb{C}^-$ so that $\mathbf{G}$ and $\hat{\mathbf{H}}$ belong to $\mathcal{H}_2$ and $\dim(\mathbf{G}) = N$, $\dim(\hat{\mathbf{H}}) = n$. Let us now first generalize the inner product results to the MIMO case.

5.4.1 $\mathcal{H}_2$ inner product for multiple input and output delays MIMO systems

First of all, a fundamental result dealing with the $\mathcal{H}_2$ norm invariance in case of input/output delayed systems is recalled.

Proposition 5.14. **$\mathcal{H}_2$ norm invariance** *Let $\hat{\mathbf{H}} \in \mathcal{H}_2^{n_y \times n_u}$ be a stable dynamical system and $\mathbf{M} \in \mathcal{H}_\infty^{n_u \times n_u}$, $\mathbf{N} \in \mathcal{H}_\infty^{n_y \times n_y}$ such that:*

$$\forall \omega \in \mathbb{R},\ \overline{\mathbf{M}(i\omega)}\mathbf{M}^T(i\omega) = \mathbf{I}_{n_u},\ \mathbf{N}^T(i\omega)\overline{\mathbf{N}(i\omega)} = \mathbf{I}_{n_y}. \tag{5.21}$$

If $\hat{\mathbf{H}}_d = \mathbf{N}\hat{\mathbf{H}}\mathbf{M}$ then $\|\hat{\mathbf{H}}_d\|_{\mathcal{H}_2} = \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}$.

Proof. If $\hat{\mathbf{H}}_d = \mathbf{N}\hat{\mathbf{H}}\mathbf{M}$, the scaled term $2\pi\|\hat{\mathbf{H}}_d\|_{\mathcal{H}_2}^2$ can be expressed as :

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \operatorname{trace}\left(\overline{\mathbf{N}(i\omega)}\overline{\hat{\mathbf{H}}(i\omega)}\overline{\mathbf{M}(i\omega)}\mathbf{M}^T(i\omega)\hat{\mathbf{H}}^T(i\omega)\mathbf{N}^T(i\omega)\right) d\omega \\
&= \int_{-\infty}^{+\infty} \operatorname{trace}\left(\overline{\mathbf{N}(i\omega)}\overline{\hat{\mathbf{H}}(i\omega)}\hat{\mathbf{H}}^T(i\omega)\mathbf{N}^T(i\omega)\right) d\omega \\
&= \int_{-\infty}^{+\infty} \operatorname{trace}\left(\overline{\hat{\mathbf{H}}(i\omega)}\hat{\mathbf{H}}^T(i\omega)\mathbf{N}^T(i\omega)\overline{\mathbf{N}(i\omega)}\right) d\omega \\
&= \int_{-\infty}^{+\infty} \operatorname{trace}\left(\overline{\hat{\mathbf{H}}(i\omega)}\,\hat{\mathbf{H}}^T(i\omega)\right) d\omega = 2\pi\|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2.
\end{aligned}$$

□

One can easily check that condition (5.21) appearing in Proposition 5.14 is satisfied by the delay matrices of the two last lines of (5.17) when $\mathbf{M} = \hat{\mathbf{\Delta}}_i$ and $\mathbf{N} = \hat{\mathbf{\Delta}}_o$. In other words, the $\mathcal{H}_2$ norm does not depend on the input, nor output delays. Now let us derive the $\mathcal{H}_2$ inner product computation with input/output delays in the MIMO context.

Theorem 5.15. **($\mathcal{H}_2$ inner product computation with input/output delays)** *Let $\mathbf{G}$, $\hat{\mathbf{H}}$ be two systems $\in \mathcal{H}_2^{n_y \times n_u}$ whose respective transfer functions $\mathbf{G}(s)$ and $\hat{\mathbf{H}}(s)$ can be expressed as in (5.20). Let $\hat{\mathbf{\Delta}}_i$, $\hat{\mathbf{\Delta}}_o$ be real, elements of $\mathcal{H}_\infty^{n_u \times n_u}$ and $\mathcal{H}_\infty^{n_y \times n_y}$ respectively, models satisfying $\sup_{s\in\mathbb{C}_+}\{\|\hat{\mathbf{\Delta}}_o(s)\|_F, \|\hat{\mathbf{\Delta}}_i(s)\|_F\} = M < +\infty$. By denoting $\hat{\mathbf{H}}_d = \hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i$, the inner product $\langle\hat{\mathbf{H}}_d, \mathbf{G}\rangle_{\mathcal{H}_2}$ is expressed as:*

$$\begin{aligned}
\langle\hat{\mathbf{H}}_d, \mathbf{G}\rangle_{\mathcal{H}_2} &= \sum_{j=1}^{N} \operatorname{trace}\left(\operatorname{Res}\left[\hat{\mathbf{H}}_d(-s)\mathbf{G}^T(s), \mu_j\right]\right) \\
&= \sum_{j=1}^{N} \mathbf{l}_j^T\hat{\mathbf{\Delta}}_o(-\mu_j)\hat{\mathbf{H}}(-\mu_j)\hat{\mathbf{\Delta}}_i(-\mu_j)\mathbf{r}_j.
\end{aligned} \tag{5.22}$$

Proof. The proof is similar to the SISO case one. Once again, one should take the contour encircling the poles of $\mathbf{G}(-s)$ in order to make the integral over Γ_R converges to 0 when $R \to \infty$. Then the result follows. □

Based on Theorem 5.15, let us derive the $\mathcal{H}_2$ optimality conditions in the MIMO case.

5.4.2 $\mathcal{H}_2$ optimality conditions

The following proposition makes now explicit the computation of the $\mathcal{H}_2$ norm associated with the dynamical mismatch gap $\|\mathbf{G} - \hat{\mathbf{H}}_d\|_{\mathcal{H}_2}$, which defines Problem 5.13 criterion.

Proposition 5.16. *Let* $\mathbf{G},\ \hat{\mathbf{H}}_d \in \mathcal{H}_2^{n_y \times n_u}$ *s.t.* $\hat{\mathbf{H}}_d$ *is given by Equation (5.18). The* $\mathcal{H}_2$ *norm of the approximation gap (or mismatch error), denoted by* $\mathcal{J}_2$*, can be expressed as:*

$$\begin{aligned} \mathcal{J}_2 &= \|\mathbf{G} - \hat{\mathbf{\Delta}}_o \hat{\mathbf{H}} \hat{\mathbf{\Delta}}_i\|_{\mathcal{H}_2}^2 \\ &= \|\mathbf{G}\|_{\mathcal{H}_2}^2 - 2\langle \mathbf{G}, \hat{\mathbf{\Delta}}_o \hat{\mathbf{H}} \hat{\mathbf{\Delta}}_i \rangle_{\mathcal{H}_2} + \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2. \end{aligned} \tag{5.23}$$

Proof. Simply develop the $\mathcal{H}_2$ norm using the inner product definition and exploit Proposition 5.14, *i.e.*, $\|\hat{\mathbf{\Delta}}_o \hat{\mathbf{H}} \hat{\mathbf{\Delta}}_i\|_{\mathcal{H}_2} = \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}$. □

Considering the mathematical formulation of Problem 5.13 and the reduced order system structure $\hat{\mathbf{H}}_d = \hat{\mathbf{\Delta}}_o \hat{\mathbf{H}} \hat{\mathbf{\Delta}}_i$, where $\hat{\mathbf{H}}(s)$ is given as in (5.20), the underlying optimization issue that must be solved is parameterized by $(k = 1, \ldots, n)$:

(i) the n pole(s) $\hat{\lambda}_k \in \mathbb{C}^-$.

(ii) the n bi-tangential directions $\{\hat{\mathbf{b}}_k, \hat{\mathbf{c}}_k\} \in \mathbb{C}^{n_u \times 1} \times \mathbb{C}^{n_y \times 1}$, for $k = 1, \ldots, n$.

(iii) the $n_u + n_y$ delay values $(\hat{\tau}_l, \hat{\gamma}_m),\ l = 1 \ldots n_u,\ m = 1 \ldots n_y$.

Our primary objective consists in rewriting the expression of the $\mathcal{H}_2$ gap $\mathcal{J}_2$ as a function of these latter parameters which will subsequently facilitate the derivation of the $\mathcal{H}_2$ optimality conditions for Problem 5.13. This forms the topic of the three following propositions and of Theorem 5.20, which stands as the main result of this chapter .

Proposition 5.17. *From the preliminary results, the mismatch* $\mathcal{H}_2$ *gap defined previously in Proposition 5.16 can be equivalently rewritten as:*

$$\mathcal{J}_2 = \|\mathbf{G}\|_{\mathcal{H}_2}^2 + \sum_{k=1}^{n} \hat{\mathbf{c}}_k^T \hat{\mathbf{H}}(-\hat{\lambda}_k) \hat{\mathbf{b}}_k - 2 \sum_{j=1}^{N} \mathbf{l}_j^T \hat{\mathbf{\Delta}}_o(-\mu_j) \hat{\mathbf{H}}(-\mu_j) \hat{\mathbf{\Delta}}_i(-\mu_j) \mathbf{r}_j. \tag{5.24}$$

Proof. The result is immediate. When developing the $\mathcal{H}_2$ norm expression showing the inner product and then exploiting Theorem 5.15. □

Gradient of the $\mathcal{H}_2$ approximation error

From the previous equation (5.24), the first-order optimality conditions related to the minimization of $\mathcal{J}_2$ can be analytically computed. The gradient expressions of the $\mathcal{H}_2$ gap *w.r.t.* each parameters (delays, tangential directions and poles) are detailed in the two following propositions. Starting with the simplest calculations, since the second term of the right-hand side part of (5.24) is delay-dependent only, we first derive the gradient of $\mathcal{J}_2$ *w.r.t.* the delays.

Proposition 5.18. *The gradients of the $\mathcal{H}_2$ gap $\mathcal{J}_2$ with respect to the delays $\forall l = 1 \dots n_u$, $\forall m = 1 \dots n_y$ read :*

$$\begin{cases} \nabla_{\hat{\tau}_l}\mathcal{J} &= -2\dfrac{\partial\langle\hat{\mathbf{H}}_d, \mathbf{G}\rangle_{\mathcal{H}_2}}{\partial\hat{\tau}_l} \\ &= -2\displaystyle\sum_{j=1}^{N} \mu_j \mathbf{l}_j^T \hat{\mathbf{\Delta}}_o(-\mu_j)\hat{\mathbf{H}}(-\mu_j)\mathbf{D}_l\hat{\mathbf{\Delta}}_i(-\mu_j)\mathbf{r}_j, \\ \nabla_{\hat{\gamma}_m}\mathcal{J} &= -2\dfrac{\partial\langle\hat{\mathbf{H}}_d, \mathbf{G}\rangle_{\mathcal{H}_2}}{\partial\hat{\gamma}_m} \\ &= -2\displaystyle\sum_{j=1}^{N} \mu_j \mathbf{l}_j^T \mathbf{D}_m\hat{\mathbf{\Delta}}_o(-\mu_j)\hat{\mathbf{H}}(-\mu_j)\hat{\mathbf{\Delta}}_i(-\mu_j)\mathbf{r}_j, \end{cases}$$

where elements of $\mathbf{D}_l \in \mathbb{R}^{n_u \times n_u}$, $\mathbf{D}_m \in \mathbb{R}^{n_y \times n_y}$, are defined as:

$$[\mathbf{D}_k]_{ij} = \delta_{ijk} = \begin{cases} 1 \text{ if } i = j = k \\ 0 \text{ otherwise} \end{cases}.$$

Proof. The proof is straightforward to establish by noticing that both $\hat{\mathbf{\Delta}}_i$ and $\hat{\mathbf{\Delta}}_o$ terms are diagonal matrices and the exponential derivative function is obvious. □

Now, let us compute the gradient with respect to the parameters defining $\hat{\mathbf{H}}$.

Proposition 5.19. *The gradients of the $\mathcal{H}_2$ gap $\mathcal{J}_2$ with respect to parameters $\hat{\mathbf{c}}_k$, $\hat{\mathbf{b}}_k$ and $\hat{\lambda}_k$, $\forall k = 1 \dots n$ read:*

$$\begin{cases} \nabla_{\hat{\mathbf{c}}_k}\mathcal{J} &= -2\dfrac{\partial\langle\hat{\mathbf{H}}_d, \mathbf{G}\rangle_{\mathcal{H}_2}}{\partial\hat{\mathbf{c}}_k} + \dfrac{\partial\|\hat{\mathbf{H}}\|^2_{\mathcal{H}_2}}{\partial\hat{\mathbf{c}}_k} \\ &= -2\hat{\mathbf{b}}_k^T\left(\tilde{\mathbf{G}}(-\hat{\lambda}_k) - \hat{\mathbf{H}}(-\hat{\lambda}_k)\right)^T, \\ \nabla_{\hat{\mathbf{b}}_k}\mathcal{J} &= -2\hat{\mathbf{c}}_k^T\left(\tilde{\mathbf{G}}(-\hat{\lambda}_k) - \hat{\mathbf{H}}(-\hat{\lambda}_k)\right), \\ \nabla_{\hat{\lambda}_k}\mathcal{J} &= 2\hat{\mathbf{c}}_k^T\left(\tilde{\mathbf{G}}'(-\hat{\lambda}_k) - \hat{\mathbf{H}}'(-\hat{\lambda}_k)\right)\hat{\mathbf{b}}_k, \end{cases}$$

where:

$$\tilde{\mathbf{G}}(s) = \sum_{j=1}^{N} \hat{\mathbf{\Delta}}_o(-\mu_j)\frac{\mathbf{l}_j^T\mathbf{r}_j}{s - \mu_j}\hat{\mathbf{\Delta}}_i(-\mu_j). \tag{5.25}$$

and where $\tilde{\mathbf{G}}'$ and $\hat{\mathbf{H}}'$ are the Laplace derivative of $\tilde{\mathbf{G}}$ and $\hat{\mathbf{H}}$, respectively.

Proof. By defining $\tilde{\mathbf{r}}_j = \hat{\mathbf{\Delta}}_i(-\mu_j)\mathbf{r}_j$ and $\tilde{\mathbf{l}}_j^T = \mathbf{l}_j^T\hat{\mathbf{\Delta}}_o(-\mu_j)$ with $j = 1 \dots N$, the $\mathcal{H}_2$ product can be written as:

$$\begin{aligned} \langle\mathbf{G}, \hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i\rangle_{\mathcal{H}_2} &= \sum_{j=1}^{N}\tilde{\mathbf{l}}_j^T\Big(\sum_{m=1}^{n}\frac{\hat{\mathbf{c}}_m\hat{\mathbf{b}}_m^T}{-\mu_j - \hat{\lambda}_m}\Big)\tilde{\mathbf{r}}_j \\ &= \sum_{m=1}^{n}\hat{\mathbf{c}}_m\Big(\sum_{j=1}^{N}\frac{\tilde{\mathbf{l}}_j\tilde{\mathbf{r}}_j^T}{-\mu_j - \hat{\lambda}_m}\Big)\hat{\mathbf{b}}_m \\ &= \langle\tilde{\mathbf{G}}, \hat{\mathbf{H}}\rangle_{\mathcal{H}_2} \end{aligned}$$

Hence, the $\mathcal{H}_2$ gap can be expressed as:

$$\begin{aligned}\mathcal{J}_2 &= \|\mathbf{G}\|_{\mathcal{H}_2}^2 - 2\langle \mathbf{G}, \hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i\rangle_{\mathcal{H}_2} + \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2 \\ &= \|\mathbf{G}\|_{\mathcal{H}_2}^2 - 2\langle \tilde{\mathbf{G}}, \hat{\mathbf{H}}\rangle_{\mathcal{H}_2} + \|\hat{\mathbf{H}}\|_{\mathcal{H}_2}^2\end{aligned}$$

Hence, this expression does not depend on the delay blocks anymore and the result follows as in the delay-free case. □

Theorem 5.20 gathers all the first-order optimality conditions related to Problem 5.13 and stands as the main result of this chapter.

Theorem 5.20. (Delay model approximation first-order $\mathcal{H}_2$ optimality conditions) *Let us consider $\mathbf{G} \in \mathcal{H}_2^{n_y\times n_u}$ whose transfer function is $\mathbf{G}(s) \in \mathbb{C}^{n_y\times n_u}$. Let $\hat{\mathbf{H}}_d = \hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i$ be a local optimum of Problem 5.13. It is assumed that $\hat{\mathbf{H}} \in \mathcal{H}_2^{n_y\times n_u}$ corresponds to a model with semi-simple poles only and whose transfer function is denoted by $\hat{\mathbf{H}}(s) = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B} \in \mathbb{C}^{n_y\times n_u}$. Let $\hat{\mathbf{\Delta}}_i$, $\hat{\mathbf{\Delta}}_o$ be elements of $\mathcal{H}_\infty^{n_u\times n_u}$ and $\mathcal{H}_\infty^{n_y\times n_y}$, respectively, s.t. Propositions 5.14 is verified. Then, the following equalities hold:*

$$\left\{\begin{array}{rcl}\hat{\mathbf{H}}(-\hat{\lambda}_k)\hat{\mathbf{b}}_k &=& \tilde{\mathbf{G}}(-\hat{\lambda}_k)\hat{\mathbf{b}}_k, \\ \hat{\mathbf{c}}_k^T\hat{\mathbf{H}}(-\hat{\lambda}_k) &=& \hat{\mathbf{c}}_k^T\tilde{\mathbf{G}}(-\hat{\lambda}_k), \\ \hat{\mathbf{c}}_k^T\hat{\mathbf{H}}'(-\hat{\lambda}_k)\hat{\mathbf{b}}_k &=& \hat{\mathbf{c}}_k^T\tilde{\mathbf{G}}'(-\hat{\lambda}_k)\hat{\mathbf{b}}_k,\end{array}\right. \tag{5.26}$$

$$\left\{\begin{array}{l}\displaystyle\sum_{j=1}^{N}\mu_j\mathbf{l}_j^T\hat{\mathbf{\Delta}}_o(-\mu_j)\hat{\mathbf{H}}(-\mu_j)\mathbf{D}_l\hat{\mathbf{\Delta}}_i(-\mu_j)\mathbf{r}_j = 0, \\ \displaystyle\sum_{j=1}^{N}\mu_j\mathbf{l}_k^T\mathbf{D}_m\hat{\mathbf{\Delta}}_o(-\mu_j)\hat{\mathbf{H}}(-\mu_j)\hat{\mathbf{\Delta}}_i(-\mu_j)\mathbf{r}_j = 0,\end{array}\right. \tag{5.27}$$

for all $k = 1\dots n, l = 1\dots n_u$ and $m = 1\dots n_y$ where $\tilde{\mathbf{G}}(s)$ is given by (5.25) and the terms $\mathbf{D}_l$ and $\mathbf{D}_m$ are defined in Proposition 5.18.

Proof. The interpolation conditions gathered in (5.26) are deduced by taking $\nabla_{\hat{\mathbf{c}}_l}\mathcal{J}_2 = 0$, $\nabla_{\hat{\mathbf{b}}_l}\mathcal{J}_2 = 0$ and $\nabla_{\hat{\lambda}_l}\mathcal{J}_2 = 0$. Conditions (5.27) are obtained similarly by taking $\nabla_{\hat{\tau}_l}\mathcal{J}_2 = 0$ and $\nabla_{\hat{\gamma}_m}\mathcal{J}_2 = 0$. □

Theorem 5.20 asserts that any solution of the $\mathcal{H}_2$ model approximation Problem 5.13, denoted by $\hat{\mathbf{H}}_d = \hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i$ is *s.t.* $\hat{\mathbf{H}}$ satisfies, at the same time, a set of $3n$ bi-tangential interpolation conditions detailed in (5.26) and another set of $n_u + n_y$ relations on the delays contained in the $\hat{\mathbf{\Delta}}_i$ and $\hat{\mathbf{\Delta}}_o$ diagonal matrices (5.27). Moreover, as in the delay-free case, (5.26) corresponds also to interpolation conditions. However, the main difference is that instead of interpolating the full order model $\mathbf{G}$, a new model $\tilde{\mathbf{G}}$ is here considered. This new system possesses the same poles as $\mathbf{G}$, but differs from $\mathbf{G}$ by its residues, computed as in (5.25). As in the SISO case, this new model is intrinsically related to the non-symmetric expression of the $\mathcal{H}_2$-inner product. The following remark shows that the approximation Problem 5.13 with fixed delays have the same stationnary points that the delay-free model approximation problem involving $\tilde{\mathbf{G}}$.

Remark 5.21. Optimality conditions equivalence *Let us consider* $\tilde{\mathbf{G}}$ *as in* (5.25)*. Let consider the following delay-free problem:*

Delay-free problem: *Find* $\hat{\mathbf{H}} \in \mathcal{H}_2$ *a reduced* n^{th} *order (delay-free) approximation which minimizes* $\|\tilde{\mathbf{G}} - \hat{\mathbf{H}}\|_{\mathcal{H}_2}$*.*

If $\hat{\mathbf{H}}^\star$ *is a local minimum of this problem, then it satisfies the interpolation conditions* (5.26)*. Thus, for fixed input and output delays, this problem and Problem 5.13 both lead to the same optimality conditions and have the same stationary points. Hence, for fixed blocks* $\hat{\mathbf{\Delta}}_i$ *and* $\hat{\mathbf{\Delta}}_o$*, we have the following equivalence :*

$$\left\{\begin{array}{l}\text{Find } \hat{\mathbf{H}}^\star \text{ such that}\\ \\ \hat{\mathbf{H}}^\star = \underset{\substack{\hat{\mathbf{H}} \in \mathcal{H}_2\\ \dim(\hat{\mathbf{H}}) \leq n}}{\textbf{\textit{arg min}}} \ \|\mathbf{G} - \hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i\|_{\mathcal{H}_2},\end{array}\right. \Leftrightarrow \left\{\begin{array}{l}\text{Find } \hat{\mathbf{H}}^\star \text{ such that}\\ \\ \hat{\mathbf{H}}^\star = \underset{\substack{\hat{\mathbf{H}} \in \mathcal{H}_2\\ \dim(\hat{\mathbf{H}}) \leq n}}{\textbf{\textit{arg min}}} \ \|\tilde{\mathbf{G}} - \hat{\mathbf{H}}\|_{\mathcal{H}_2},\end{array}\right.$$

5.5 Development of an algorithm and numerical applications

5.5.1 Practical considerations

In this section, three considerations about Problem 5.13 and Theorem 5.20 are discussed. These latter are relevant to sketch out an algorithm which enables the computation of the model $\hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i$ satisfying the optimality conditions appearing in Theorem 5.20. Let us consider that $\hat{\mathbf{H}}_d = \hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i$ is a local minimum of the $\mathcal{H}_2$ optimization Problem 5.13 where $\hat{\mathbf{H}}$ is given by (5.20), then:

- **Consideration ❶.** If the matrices $\hat{\mathbf{\Delta}}_o$, $\hat{\mathbf{\Delta}}_i$ and the reduced order model poles $\hat{\lambda}_1$, $\hat{\lambda}_2$, …, $\hat{\lambda}_n$ are assumed to be known, Problem 5.13 is reduced to a much simpler problem that can be solved, for example, by using the well-known Loewner framework such as in [Mayo and Antoulas, 2007];

- **Consideration ❷.** If the delay matrices $\hat{\mathbf{\Delta}}_o$, $\hat{\mathbf{\Delta}}_i$ are known, then Problem 5.13 can be solved by finding a model realization $\hat{\mathbf{H}}$ which satisfies the interpolation conditions (5.26) of Theorem 5.20, only. This can be done using, for instance, a very efficient iterative algorithm, *e.g.,* **IRKA** (see [Gugercin et al., 2008]);

- **Consideration ❸.** Assume that the system realization $\hat{\mathbf{H}}$ has already been determined. It follows that Problem 5.13 is equivalent to look for optimal delays matrices $(\hat{\mathbf{\Delta}}_o^\star, \hat{\mathbf{\Delta}}_i^\star) \in \mathcal{H}_\infty^{n_y \times n_y} \times \mathcal{H}_\infty^{n_u \times n_u}$ such that:

$$(\hat{\mathbf{\Delta}}_o^\star, \hat{\mathbf{\Delta}}_i^\star) = \max_{(\hat{\mathbf{\Delta}}_o, \hat{\mathbf{\Delta}}_i)} \langle \hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i, \mathbf{G}\rangle_{\mathcal{H}_2}. \tag{5.28}$$

  Interestingly, since $\langle \hat{\mathbf{\Delta}}_o\hat{\mathbf{H}}\hat{\mathbf{\Delta}}_i, \mathbf{G}\rangle_{\mathcal{H}_2} \to 0$ when the delays go to infinity, this problem can be restricted to a compact set and thus a global solution exists.

5.5.2 Computational considerations

An algorithm allowing to numerically compute a model $\hat{\mathbf{H}}_d$ satisfying the previous $\mathcal{H}_2$ optimality conditions is proposed in this subsection. It relies on the Remark 5.21 and the considerations discussed above (Section 5.5.1). Therefore, the proposed approach corresponds to an iterative algorithm in which each iteration can be decomposed in two steps. The first one aims at computing a realization $\hat{\mathbf{H}}$ which satisfies the interpolation conditions (5.26) while fixing the matrices $\hat{\mathbf{\Delta}}_o$, $\hat{\mathbf{\Delta}}_i$ at their values obtained from the previous iteration. This can be done using, for instance, **IRKA** (*Step 4*). In the second step, $\hat{\mathbf{H}}$ is therefore fixed and the optimal values for the $\hat{\mathbf{\Delta}}_o$, $\hat{\mathbf{\Delta}}_i$ matrices elements (*Step 5*) are determined. This sequential procedure can be summarized in Algorithm 3, and referred to as **MIMO IO-dIRKA** for **MIMO I**nput/**O**utput **d**elays **IRKA**.

Algorithm 3 MIMO IO-dIRKA

1: **Input:** A N^{th} order model $\mathbf{G} \in \mathcal{H}_2^{n_y \times n_u}$, dimension $n \in \mathbb{N}^*$ ($n \ll N$) and initial guesses for both $\hat{\mathbf{\Delta}}_i^{\mathsf{it}=0}$, $\hat{\mathbf{\Delta}}_o^{\mathsf{it}=0}$.
2: Make initial choice of the shift points $\sigma^0 = \{\sigma_1^0, \ldots, \sigma_n^0\} \in \mathbb{C}$ initial interpolation points and tangential directions $\hat{\mathbf{b}}_{1,0}, \ldots, \hat{\mathbf{b}}_{n,0} \in \mathbb{C}^{n_u \times 1}$ and $\hat{\mathbf{c}}_{1,0}, \ldots, \hat{\mathbf{c}}_{n,0} \in \mathbb{C}^{n_y \times 1}$ closed by conjugation.
3: **while** not convergence **do**
4: Set $\mathsf{it} \leftarrow \mathsf{it} + 1$
5: Build $\tilde{\mathbf{G}}^{\mathsf{it}}$ as in (5.25)
6: Build $\hat{\mathbf{H}}^{\mathsf{it}}$ satisfying the bi-tangential interpolation conditions (5.26) using **IRKA** [Gugercin et al., 2008] on $\tilde{\mathbf{G}}^{\mathsf{it}}$
7: Determine $(\hat{\mathbf{\Delta}}_i^\star, \hat{\mathbf{\Delta}}_o^\star)$ which solve (5.28) using $\hat{\mathbf{H}}^{\mathsf{it}}$
8: Set $\hat{\mathbf{\Delta}}_i^{\mathsf{it}} \leftarrow \hat{\mathbf{\Delta}}_i^\star$, $\hat{\mathbf{\Delta}}_o^{\mathsf{it}} \leftarrow \hat{\mathbf{\Delta}}_o^\star$
9: **end while**
10: **Output:** locally $\mathcal{H}_2$ optimal reduced model $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \hat{\mathbf{\Delta}}_i^\star, \hat{\mathbf{\Delta}}_o^\star)$.

This iterative algorithm is inspired from [Gugercin et al., 2008; Van Dooren et al., 2008] and, upon convergence, it provides input/output delay ROM satisfying the optimality conditions from Theorem 5.20.

To sum up, from a model reduction point of view, we believe that Theorem 5.20 is the great novelty from this chapter. One of the implication of this theorem is that, for a fixed delay, one can rewrite the input-delay approximation problem as a delay-free approximation problem. Hence, the new delay-free model approximation problem can be solved using dedicated tools. As it is drastically efficient on the large-scale framework, we have proposed an algorithm based on the fixed point iteration of **IRKA** (see [Gugercin et al., 2008]) to find a good approximation. The algorithm **MIMO IO-dIRKA** itself was not the main goal, but rather a way to exploit the main contribution using some well known results from the literature in order to find a solution of Problem 5.13.

Later in this section this algorithm is validated in some numerical examples.

Structured input/output delays

All the previous results are left unchanged in the case of structured input/output delays *i.e.,* if, for example, delays do not apply on given input(s) and/or output(s) of $\hat{\mathbf{H}}_d$. The results can be derived in a straightforward way, without any loss of generality, just by considering the following ordered delays matrices (where delays are present on the first $n_{d1} < n_u$ inputs and $n_{d2} < n_y$

outputs):

$$\begin{cases} \hat{\mathbf{\Delta}}_i(s) = \mathbf{diag}(e^{-s\hat{\tau}_1},\ e^{-s\hat{\tau}_2},\ \ldots,\ e^{-s\hat{\tau}_{n_{d1}}},\ 1,\ \ldots,\ 1) \\ \hat{\mathbf{\Delta}}_o(s) = \mathbf{diag}(e^{-s\hat{\gamma}_1},\ e^{-s\hat{\gamma}_2},\ \ldots,\ e^{-s\hat{\gamma}_{n_{d2}}},\ 1,\ \ldots,\ 1). \end{cases}$$

One can easily note that the preliminary results from Sections 5.2 and 5.4 still remain true when introducing these matrices. The main result stated in Theorem 5.20 thus remains unchanged. This can be particularly interesting when the original model has the delay behavior only for some inputs and outputs.

5.5.3 Numerical applications

This section is dedicated to the application of the results obtained in this chapter, namely, the input/output-delay optimal $\mathcal{H}_2$ model approximation and its first-order optimality conditions. We will emphasize the potential benefit and effectiveness of the proposed approach.

Example 1: Padé approximation

Let us consider a dynamical model $\mathbf{H}_{ref} \in \mathcal{H}_2$ whose transfer function is given by

$$\mathbf{H}_{ref}(s) = \frac{1}{s+1}e^{-s}. \tag{5.29}$$

The reference dynamical model (5.29) (which is obviously of order 1) has an input delay. This input delay is, then, approximated by a Padé of order 1, which leads to the approximation model $\mathbf{G}(s)$ of order $N = 2$, given by the following function

$$\mathbf{G}(s) = \frac{1}{s+1}\frac{2-s}{s+2} = \frac{3}{s+1} - \frac{4}{s+2}. \tag{5.30}$$

The purpose here is to approximate $\mathbf{G}$ (5.30) by a lower delayed model of order $n = 1$. For fixed values of τ in the interval $[0, 1]$, one now seek for a model $\hat{\mathbf{H}}_d = \hat{\mathbf{H}}e^{-\tau s}$, where $\hat{\mathbf{H}}(s) = \dfrac{\hat{\phi}}{s-\hat{\lambda}}$ which satisfies the interpolation conditions given in the SISO case as (5.12), where $\tilde{\mathbf{G}}(s)$ reads as follows:

$$\tilde{\mathbf{G}}(s) = \frac{3}{s+1}e^{-\tau} - \frac{4}{s+2}e^{-2\tau}.$$

To do so, the iterative algorithm **IRKA** from [Gugercin et al., 2008], which enables us to find a model satisfying (5.12), is applied. For each frozen value of the delay value τ, the $\mathcal{H}_2$-gap between $\hat{\mathbf{H}}_d$ and $\mathbf{G}$ was computed and the results are reported on Figure 5.7. It shows that the $\mathcal{H}_2$-optimal model is attained for a delay value $\tau \approx 0.48s$. One should first note that $0.48s < 1s$ (1 was the initial delay of the original model $\mathbf{H}_{ref}$ before Padé approximation), which at a first look, might be surprising.

However, by analyzing Figure 5.8, which shows the Impulse responses of $\mathbf{H}_{ref}$ (solid red line), of the model $\mathbf{G}$ obtained after Padé approximation (solid green line) and the optimal $\mathcal{H}_2$-approximation with delay $\hat{\mathbf{H}}_d$ ($n = 1$ dashed blue), it appears that $\hat{\mathbf{H}}_d$ well approximates $\mathbf{G}$. As a first remark, it is clear that the Padé approximation $\mathbf{G}$ degrades the original input delay from $\mathbf{H}$ and we are not able to recover it doing an $\mathcal{H}_2$-approximation. Indeed, since the Padé approximation impulse response concentrates a great amount of energy when $t < 1s$, then a $\mathcal{H}_2$ approximation with input delay $\tau \geq 1$ will not take this into account and one should obtain a better approximation when $\tau \approx 0.48s$.

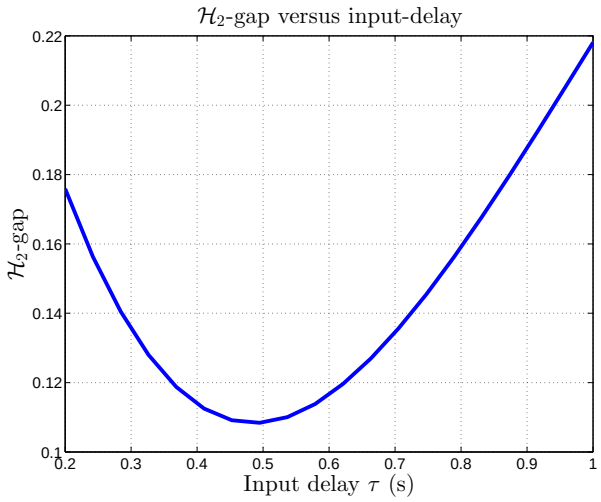

Figure 5.7: The $\mathcal{H}_2$-gap, *i.e.*, $\|\mathbf{G} - \hat{\mathbf{H}}e^{-s\tau}\|_{\mathcal{H}_2}$ as function of τ using **IRKA**.

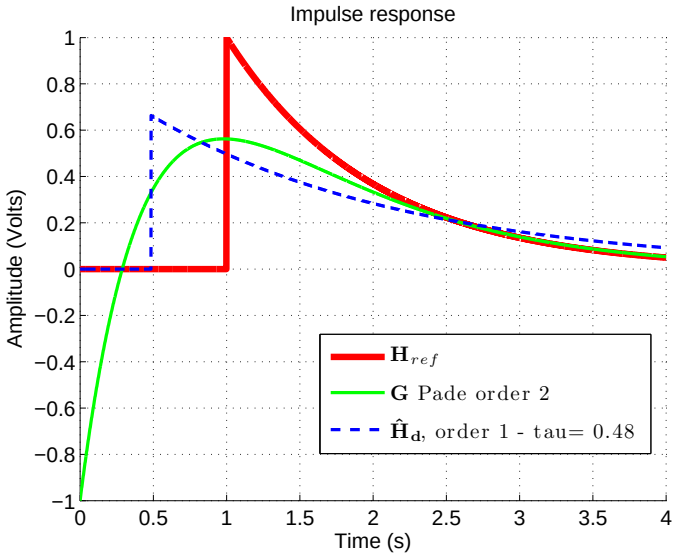

Figure 5.8: Impulse response of original model $\mathbf{H}$ (red solid line), Padé approximation $\mathbf{G}$ (green solid line) and input-delay $\mathcal{H}_2$-optimal model using interpolation conditions (5.12) of order $n = 1$ (blue dashed line).

This specific example shows how the Padé approximation degrades some characteristics of the original model and that performing a Padé approximation followed by a $\mathcal{H}_2$ approximation does not allow us to recover the original input delay.

Example 2: pseudo higher order poles

Let us consider a model $\mathbf{G}$ of order $N = 20$, given by the following transfer function

$$\mathbf{G}(s) = \prod_{j=1}^{20} \frac{\mu_j}{s - \mu_j}, \tag{5.31}$$

where $\mu_j \in \mathbb{R}_-$ $(j = 1, \ldots, N)$ are linearly spaced between $[-2 \ -1]$. The impulse response of $\mathbf{G}$ is given by the solid dotted blue line in Figure 5.9. Interestingly, it behaves like a system with an input delay. In order to fit the framework proposed in this chapter, input-delay $\mathcal{H}_2$ optimal model $\hat{\mathbf{H}}_d = \hat{\mathbf{\Delta}}_o \hat{\mathbf{H}} \hat{\mathbf{\Delta}}_i$ of order $n = 2$ (solid red) was obtained by applying Theorem 5.20 and **IO-dIRKA**, as described in Section 5.3. The obtained delay model is compared with delay-free approximations of order $n = \{2, 3, 4\}$, obtained with **IRKA**. All the results are reported on Figure 5.9.

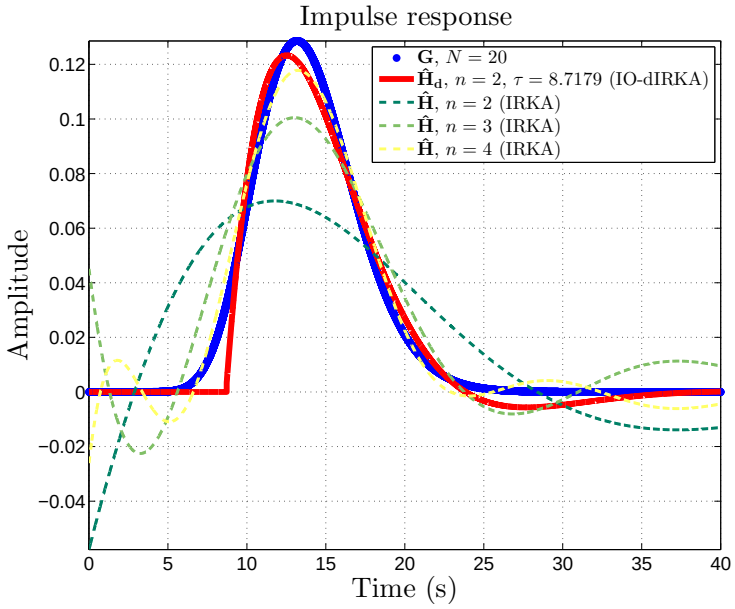

Figure 5.9: Impulse response of the original model $\mathbf{H}$ of order $N = 20$ (solid dotted blue line), the input-delay $\mathcal{H}_2$-optimal model $\hat{\mathbf{H}}_d$ of order $n = 2$ (solid red line) and the delay-free $\mathcal{H}_2$-optimal models $\hat{\mathbf{H}}$ of order $n = \{2, 3, 4\}$ (dashed dark green, light green and yellow lines).

As clearly shown on Figure 5.9, the proposed methodology allows to obtain an input-delay $\mathcal{H}_2$ approximation of model $\mathbf{G}$ that clearly provides a better matching than the delay-free cases, even for higher orders (here, **IRKA** with $n = 4$ still have a bad matching and exhibits difficulties in accurately catching the delay and main dynamics). Indeed, the delay-free cases exhibit an oscillatory behavior during the first seconds while the input-delay model $\hat{\mathbf{H}}_d$ takes benefit of the delay structure to focus on the main dynamical effect. Moreover, the approximation model of $\hat{\mathbf{H}}_d$ satisfies the conditions given in Theorem 5.20.

Remark 5.22 (**Numerical results (SISO case, $n = 2$)**). *For the sake of completeness, the optimal numerical values obtained with **MIMO IO-dIRKA** are:* $\hat{\lambda}_{1,2} = -2.0320 \times 10^{-1} \pm i\ 2.0700 \times 10^{-1}$, $\hat{\phi}_{1,2} = 1.5713 \times 10^{-3} \pm i\ 1.8691 \times 10^{-1}$ *and the optimal delay* $\tau = 8.7179$. *The interpolation conditions can be then easily checked:*

- *Condition* (5.12) *leads to* $\hat{\mathbf{H}}(-\hat{\lambda}_{1,2}) = \tilde{\mathbf{G}}(-\hat{\lambda}_{1,2}) = 2.3567 \times 10^{-1} \pm i\ 2.3614 \times 10^{-1}$ *and* $\hat{\mathbf{H}}'(-\hat{\lambda}_{1,2}) = \tilde{\mathbf{G}}'(-\hat{\lambda}_{1,2}) = 5.6466 \times 10^{-1} \pm i\ 1.1465$.
- *When evaluating* $\sum_{j=1}^{N} \mu_j \psi_j \left(\sum_{k=1}^{n} \frac{\hat{\phi}_k}{\mu_j + \hat{\lambda}_k} \right) e^{\tau \mu_j}$, *one obtains* 9.7284×10^{-5}, *which is close to zero, as stated by condition* (5.13).

With reference to Figure 5.10, similar results are obtained in the case of an input delay-dependent approximation of order $n = 4$ (using **IO-dIRKA**) and delay-free approximation of order $n = \{4, 5, 6\}$ (using **IRKA**). Then, Figure 5.11 shows the impulse response mismatch errors for these different configurations. For each reduced order models, the mean square absolute error ε of the impulse responses are computed. The main observation that can be made is that the mismatch error obtained for $\hat{\mathbf{H}}_d$ of order $n = 4$ is lower that the one obtained by a delay-free model $\hat{\mathbf{H}}$ of order $n = 6$ (a better result is obtained for a delay-free model with an order $n = 7$). This motivates the use of the specific approximation model delay structure.

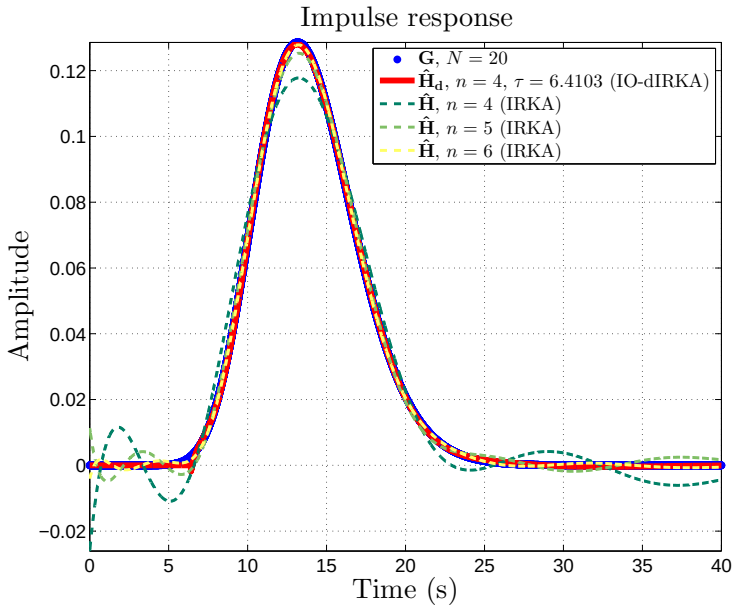

Figure 5.10: Impulse response of the original model $\mathbf{H}$ of order $N = 20$ (solid dotted blue line), the input-delay $\mathcal{H}_2$-optimal model $\hat{\mathbf{H}}_d$ of order $n = 4$ (solid red line) and the delay-free $\mathcal{H}_2$-optimal models $\hat{\mathbf{H}}$ of order $n = \{4, 5, 6\}$ (dashed dark green, light green and yellow lines).

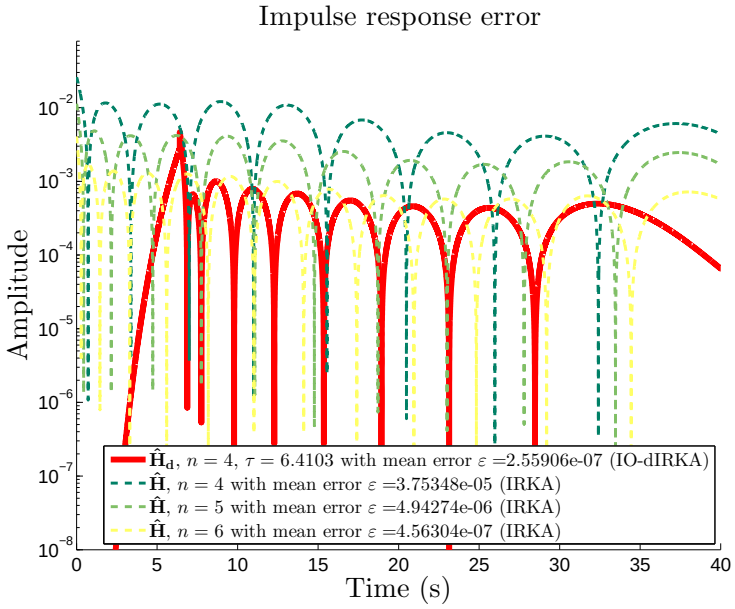

Figure 5.11: Impulse response error between the original model $\mathbf{H}$ of order $N = 20$ and the input-delay $\mathcal{H}_2$-optimal model $\hat{\mathbf{H}}_d$ of order $n = 4$ (solid red line) and the delay-free $\mathcal{H}_2$-optimal models $\hat{\mathbf{H}}$ of order $n = \{4, 5, 6\}$ (dashed dark green, light green and yellow lines).

Example 3: Loewner framework and pseudo-delay behavior

Let $\mathbf{G}$ be a SISO model of order $N = 34$ having complex poles which has an intrinsic input-delay behavior. This model was constructed by means of the Loewner framework (see [Mayo and Antoulas, 2007]) following these two steps:

1. Take $\mathbf{G}_{delay}(s) = \dfrac{\psi}{s^2 + 2\xi\omega_0 s + \omega_0^2} e^{-\tau s}$, where $\tau = 2, \omega_0 = 1$ and $\xi = 1/4$.

2. Using the Loewner framework for uniformly spaced interpolation points $i\omega_k$, $k = 1, \ldots, 100$, (see [Mayo and Antoulas, 2007]), we were able to construct $\mathbf{G} = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B}$ of order $N = 34$, a delay-free model interpolating $\mathbf{G}_{delay}$.

Hence, by construction, $\mathbf{G}$ is a delay-free model having an intrinsic input-delay behavior. Then, for the orders $n = 2, \ldots, 8$ delay-free model approximations were constructed using the **IRKA** algorithm. In addition, an input-delay approximation $\hat{\mathbf{H}}_d$ of order $n = 2$ was constructed using the proposed **IO-dIRKA** algorithm. The results are presented in Figure 5.5.3.

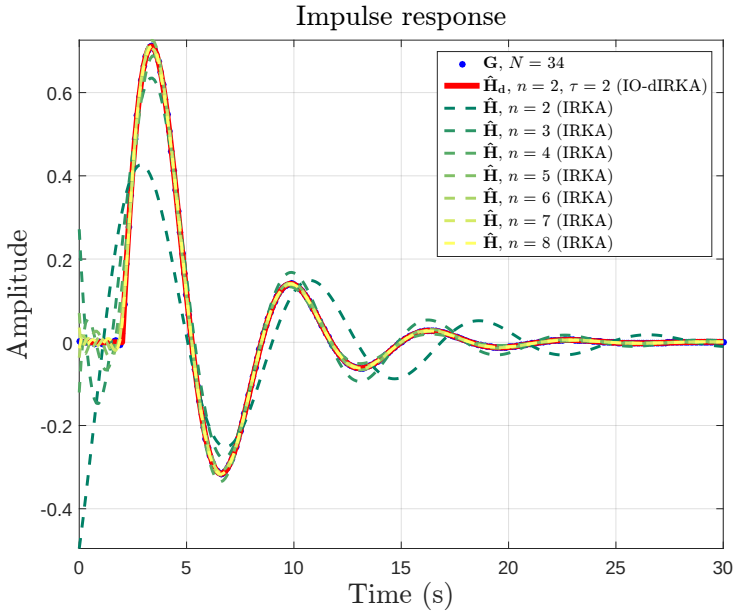

Figure 5.12: Impulse responses of the original model $\mathbf{G}$ of order $N = 34$ (solid blue dots), the optimal input model approximation $\hat{\mathbf{H}}_d$ of order $n = 2$ (solid red line) and optimal delay-free models of order 2 to 8 using **IRKA** (dashed lines).

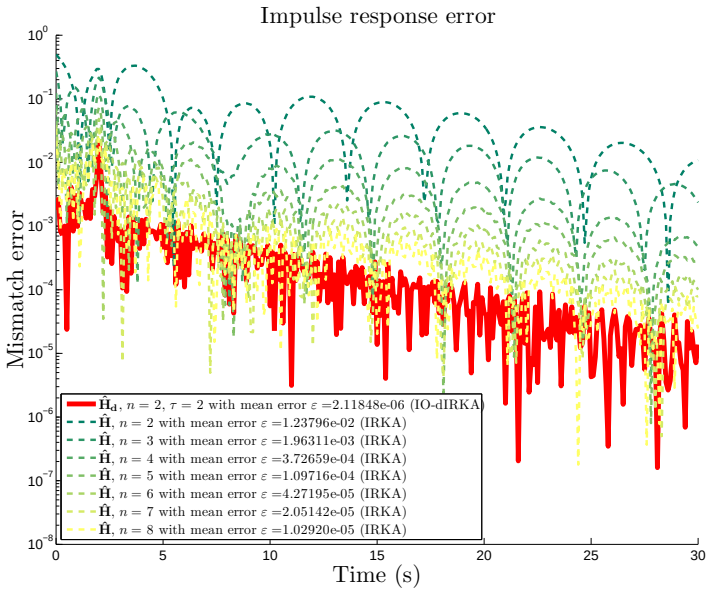

Figure 5.13: Impulse error responses for the optimal input model approximation $\hat{\mathbf{H}}_d$ of order $n = 2$ (solid red line) and the optimal delay-free models of order 2 to 8 using **IRKA** (dashed lines).

This example shows that one needs to increase the order of a delay-free approximation to capture the intrinsic delay behavior of $\mathbf{G}$. In this case, an input-delay approximation is suitable. Moreover, the proposed algorithm provides a good approximation and well captured the optimal delay $\hat{\tau} \approx 2$, the natural frequency $\hat{\omega}_0 \approx 1$ and damping $\hat{\xi} \approx 0.25005$, as they are very close to the initial parameters from $\mathbf{G}_{delay}$.

Example 4: Ladder Network system

Let $\mathbf{G}_{Ladder}$ be the ladder network system presented in Chapter 1 from [Beattie and Gugercin, 2011] and [Gugercin et al., 2012]. Even if this system is represented by a finite dimensional system, it has an intrinsic input-delay behavior (as shown by its impulse response in Chapter 1). Then, for the orders n = 6, 12 and 20, delay-free model approximations were constructed using **IRKA** and the results are reported in Figure 5.14.

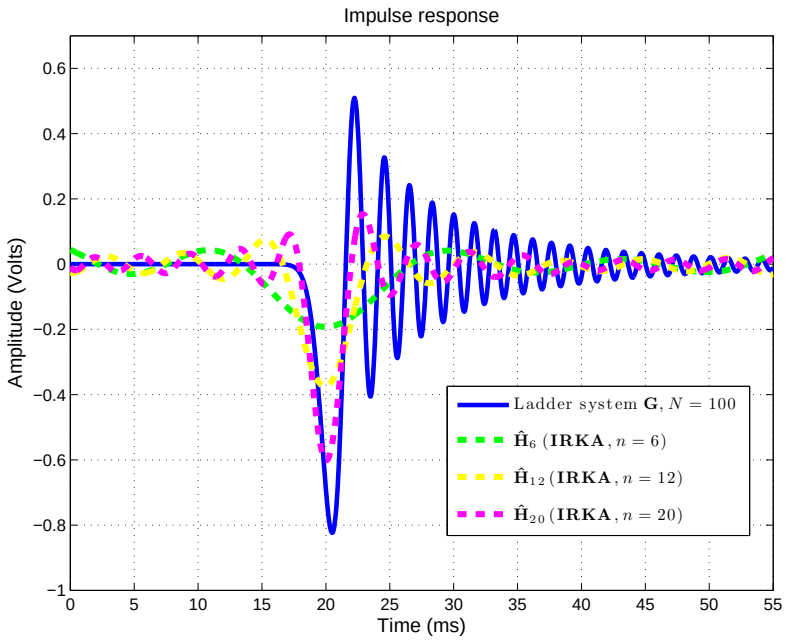

Figure 5.14: Impulse response of the Ladder system $\mathbf{G}_{Ladder}$ of order $N = 100$ (solid blue line) and the delay-free $\mathcal{H}_2$-optimal model $\hat{\mathbf{H}}_6$, $\hat{\mathbf{H}}_{12}$ and $\hat{\mathbf{H}}_{20}$ of orders $n = \{6, 12, 20\}$ (dashed green, dashed yellow and dashed magenta lines).

As the reader can see, even if we increase the order to $n = 20$, the delay-free approximation does not manage to follow the behavior of the impulse response of $\mathbf{G}_{Ladder}$. Let us now consider an input-delay approximation $\hat{\mathbf{H}}_d$ of order $n = 6$ constructed using the proposed **IO-dIRKA** algorithm. The optimal delay obtained was $\hat{\tau}^\star = 19.27$s. The results are presented in Figure 5.15 and compared with the delay-free ROM of order $n = 20$.

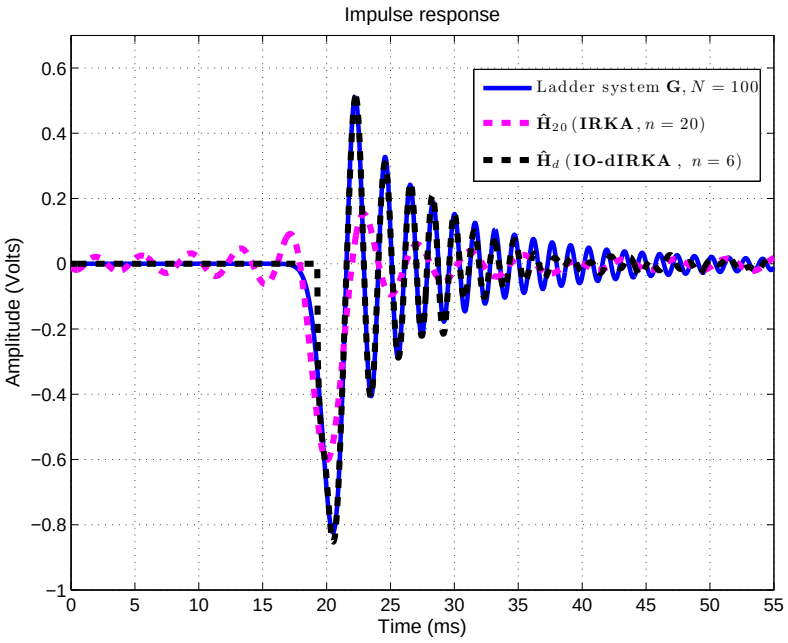

Figure 5.15: Impulse response of the Ladder system $\mathbf{G}_{Ladder}$ of order $N = 100$ (solid blue line), the delay-free $\mathcal{H}_2$-optimal model $\hat{\mathbf{H}}_{20}$ of order $n = 20$ (dashed magenta line) and the input-delay $\mathcal{H}_2$-optimal model $\hat{\mathbf{H}}_d$ of order $n = 6$, $\hat{\tau}^\star = 19.27$s (dashed black line).

As clearly shown on Figure 5.15, **IO-dIRKA** to obtain an input-delay approximation which is much more accurate than the delay-free approximation. Therefore, in this case, adding an input-delay in the reduced-order model enables to find better lower order approximations.

Example 5: Open channel flow for hydroelectricity

Let $\mathbf{H}_{flow}(s)$ be the open channel system presented in Chapter 1 from [Dalmas et al., 2016]. This model is represented by an irrational transfer function. Since all the results presented in this chapter suppose that the full order model is represented by a finite dimensional realization, they are not directly applicable to $\mathbf{H}_{flow}(s)$. In order to remedy this situation, let $\mathbf{G} \in \mathcal{H}_2$ be a finite dimensional approximation of order 103 which was built using the Loewner framework on $\mathbf{H}_{flow}$. Figure 5.16 shows the Bode magnitude plot of the irrational model $\mathbf{H}_{flow}$ and the Loewener approximation $\mathbf{G}$.

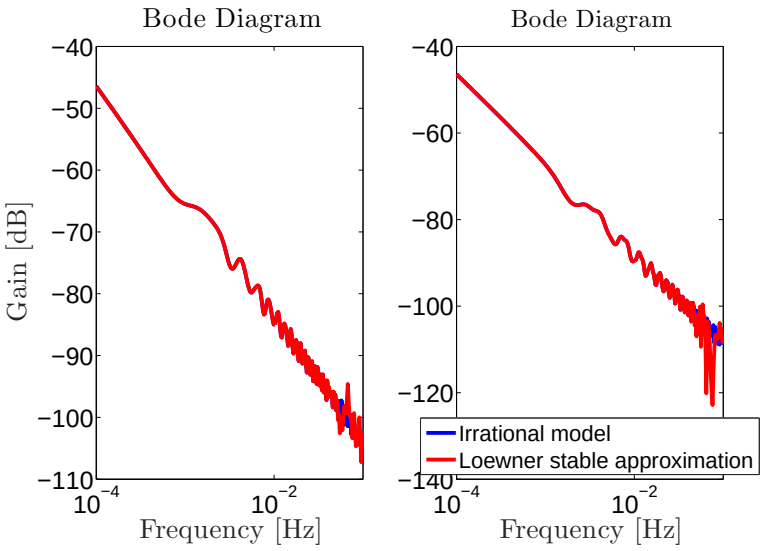

Figure 5.16: Bode magnitude diagram of the original irrational model $\mathbf{H}_{flow}$ (solid blue line) and Loewner stable approximation $\mathbf{G}$ of order 103.

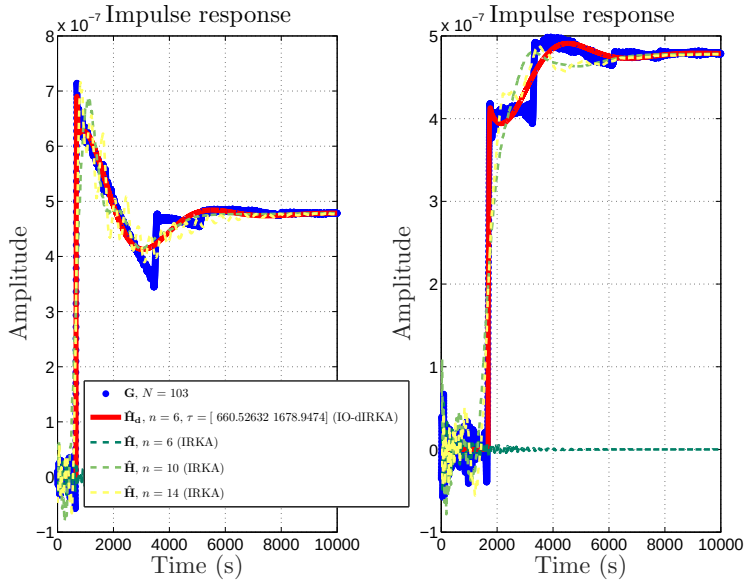

Figure 5.17: Impulse response of the full order model $\mathbf{G}$ (solid blue line), reduced-order model $\hat{\mathbf{H}}_d$ of order 6 with input delays $\tau_1 = 660.52$ and $\tau_2 = 1678.94$ (solid red line) and delay-free approximations of order $n = 6$, 10 and 14 (darker green, green and yellow dashed lines).

Hence, since $\mathbf{G}$ is a finite dimensional system, we are now able to apply the results from this chapter. An input-delay approximation $\hat{\mathbf{H}}_d$ of order $n = 6$ constructed using the proposed **IO-dIRKA** algorithm. The optimal delays obtained were $(\hat{\tau}_1^\star, \hat{\tau}_2^\star) = (660.52, 1678.94)$s. Then, for the orders n = 6, 10 and 14, delay-free model approximations were constructed using **IRKA** and the results are reported in Figure 5.17.

Once again, as shown on Figure 5.15, **IO-dIRKA** allows to obtain an input-delay approximation which reproduces the behavior of the full order model $\mathbf{G}$.

Conclusion

The main contribution of this chapter is the derivation of the first-order $\mathcal{H}_2$ optimality conditions for Problem 5.13. Moreover, to achieve this result, the expression of the $\mathcal{H}_2$-inner product in the presence of input and output delays was derived in Theorem 5.5 which forms another contribution. The crucial point is the introduction of the new interpolant model $\tilde{\mathbf{G}}$ in the $\mathcal{H}_2$ optimality conditions, which forms a direct extension of the bi-tangential interpolation conditions of the delay-free case derived in [Gugercin et al., 2008; Van Dooren et al., 2008]. Theorem 5.20 establishes that if $\hat{\mathbf{H}}_d = \hat{\mathbf{\Delta}}_o \hat{\mathbf{H}} \hat{\mathbf{\Delta}}_i$ is a local optimum, then the parameters of this latter verify an extended set of matrix equalities. These ones are of two types:

(i) a subset of interpolation conditions (5.26) satisfied by the rational part $\hat{\mathbf{H}}$ of $\hat{\mathbf{H}}_d$, which generalizes the delay-free case.

(ii) a subset of matricial relationships (5.27) focusing on the input/output delay blocks $\hat{\mathbf{\Delta}}_o$, $\hat{\mathbf{\Delta}}_i$.

These conditions are all dependent on the reduced order model parametrization described by $\hat{\mathbf{b}}_k$, $\hat{\mathbf{c}}_k$, $\hat{\lambda}_k$, $\hat{\tau}_l$ and $\hat{\gamma}_m$, and solving Problem 5.13 requires to tackle a non-convex optimization problem. An algorithm referred to as **IO-dIRKA**, has been suggested to practically address this issue and some numerical experiments have also been presented, illustrating the benefit of the proposed approximation which integrates delays with respect to standard delay-free approximation methods.

Chapter 6

Data-driven model approximation by single state-delay structure reduced order models

In this chapter, the realization-free model approximation problem, as stated in [Mayo and Antoulas, 2007; Beattie and Gugercin, 2012], is revisited in the case where the interpolating model is a single state-delay dependent one. To this aim, the Loewner framework, initially settled for delay-free realizations, is firstly extended to the single-delay one. Secondly, Finite dimensional inspired interpolation conditions are established through the use of the Lambert function. Finally, a numerically efficient iterative scheme, named **dTF-IRKA**, similar to the **TF-IRKA** [Beattie and Gugercin, 2012], is proposed to reach a part of the aforementioned optimality conditions. The proposed method validity and interest are proved on different numerical examples.

Contents

6.1 Problem statement

In Chapter 5, the model approximation problem when the reduced order model has an input/output delay structure was considered. As a next step, we wish to consider reduced order models with other kinds of structure. Hence, it seems natural to consider state-delay structures as the following step.

In this chapter, a data-driven model approximation by a single state-delay structure reduced order models is developed and used for approximation of any realization or realization-free linear dynamical model. More specifically, given some tangential interpolation frequency data (as in Chapter 3, Section 3.4), we are interested in finding a single delay finite-dimensional linear time-invariant descriptor system denoted $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ and defined by:

$$\hat{\mathbf{H}}_d = \left\{ \begin{array}{rcl} \hat{E}\dot{\mathbf{x}}(t) & = & \hat{A}\mathbf{x}(t-\tau) + \hat{B}\mathbf{u}(t) \\ \mathbf{y}(t) & = & \hat{C}\mathbf{x}(t), \end{array} \right. , \tag{6.1}$$

whose transfer function is $\hat{\mathbf{H}}_d(s) = \hat{C}(s\hat{E} - \hat{A}e^{-\tau s})^{-1}\hat{B}$ interpolates the tangential data. It is worth noticing that the approximation form (6.1) extends the delay-free one used in [Mayo and Antoulas, 2007; Beattie and Gugercin, 2012] given as $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, 0)$ (or simply $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$),

$$\hat{\mathbf{H}} = \left\{ \begin{array}{rcl} \hat{E}\dot{\mathbf{x}}(t) & = & \hat{A}\mathbf{x}(t) + \hat{B}\mathbf{u}(t) \\ \mathbf{y}(t) & = & \hat{C}\mathbf{x}(t), \end{array} \right. \tag{6.2}$$

In the model approximation context, inspired by finite dimensional $\mathcal{H}_2$ necessary interpolation conditions (see [Van Dooren et al., 2008; Gugercin et al., 2008] or Chapter 4), we are willing to tackle the following problem :

Problem 6.1 (**Finite dimensional inspired interpolation problem**). *Given* $\mathbf{G} \in \mathcal{H}_2$, *an order* n, *and a fixed delay* $\tau \geq 0$, *find* $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ *of order* n *such that*

$$\mathbf{G}(-\hat{\lambda}_k)\hat{\mathbf{b}}_k = \hat{\mathbf{H}}_d(-\hat{\lambda}_k)\hat{\mathbf{b}}_k, \;\; \hat{\mathbf{c}}_k^T\mathbf{G}(-\hat{\lambda}_k) = \hat{\mathbf{c}}_k^T\hat{\mathbf{H}}_d(-\hat{\lambda}_k) \tag{6.3}$$

and

$$\hat{\mathbf{c}}_k^T\mathbf{G}'(-\hat{\lambda}_k)\hat{\mathbf{b}}_k = \hat{\mathbf{c}}_k^T\hat{\mathbf{H}}_d'(-\hat{\lambda}_k)\hat{\mathbf{b}}_k \tag{6.4}$$

for all $\hat{\lambda}_k$ *pole of* $\hat{\mathbf{H}}_d$ *and* $\mathrm{Res}(\hat{\mathbf{H}}_d, \hat{\lambda}_k) = \hat{\mathbf{c}}_k\hat{\mathbf{b}}_k^T$.

In other words, if an evaluation of the transfer function $\mathbf{G}(s)$, for any $s \in \mathbb{C}$, is available (either from data or by simply evaluating $\mathbf{G}(s)$), our goal is to find a delay model of the form (6.1), that well approximates $\mathbf{G}$, and, which satisfies the finite dimensional inspired interpolation conditions (see Chapter 4, Theorem 4.14). Then, a main difference here compared to the delay-free case, is that the reduced order model is an infinite dimensional model having an infinity number of poles. Hence the conditions proposed in Problem 6.4 are infinitely many, while the number of parameters defining $\hat{\mathbf{H}}_d$ is finite. Hence, we shall not take into consideration all of them, but only a finite number of those conditions. Let us start our discussion with the extension of the Loewner framework (see [Mayo and Antoulas, 2007]).

6.2 Single state-delay data-driven framework

6.2.1 State-delay transformation

The main result of this section is based on a very simple representation argument, which will be useful along the rest of this chapter.

Lemma 6.2 (Single state-delay transformation). *Given* $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$, *its transfer function* $\hat{\mathbf{H}}_d(s)$ *can be decomposed as:*

$$\hat{\mathbf{H}}_d(s) = \hat{\mathbf{H}}\big(f(s)\big)e^{s\tau} \tag{6.5}$$

where

$$f(s) = se^{s\tau},$$

and $\hat{\mathbf{H}}(s)$ *is the transfer function of the delay-free model* $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ *as in* (6.2).

Proof. The result is straightforwardly obtained by using the expression of $f(s)$ combined with (6.2) as follows:

$$\begin{aligned} \hat{\mathbf{H}}_d(s) &= \hat{C}(s\hat{E} - \hat{A}e^{-s\tau})^{-1}\hat{B} \\ &= \hat{C}(se^{s\tau}\hat{E} - \hat{A})^{-1}\hat{B}e^{s\tau} = \hat{\mathbf{H}}(se^{s\tau})e^{s\tau}. \end{aligned}$$

□

Lemma 6.2 states that a single state-delay model $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ can be represented by a delay-free model $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ by means of the transformation (6.5). It is worth mentioning that the function $f(s) = se^{s\tau}$, which plays a very important role here, is the inverse of the Lambert function (see [Corless et al., 1996]). Next section will use this lemma to enable the construction of a single state-delay model $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ which interpolates some given tangential data.

6.2.2 Single state-delay Loewner framework

The Loewner framework enables to construct a delay-free model $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ which interpolates some given data. Then, one extension of the Loewner framework which makes feasible the interpolation with a single delay descriptor system as defined in (6.1) can be done by using the function $f(s) = se^{s\tau}$ as a variable substitution and applying the standard Loewner framework to the new transformed data. This first main result can be stated as follows:

Theorem 6.3 (Single state-delay Loewner framework). *Let us consider* $\tau \in \mathbb{R}$ *and given* $(\lambda_i, \mathbf{r}_i, \mathbf{w}_i)$ *and* $(\mu_j, \mathbf{l}_j, \mathbf{v}_j)$, *for* $1 \leq i, j \leq n$, *the* right and left interpolation data *respectively, as stated in Problem 3.12 from Chapter 3. Assuming that* $f(s) = se^{s\tau}$ *is one-to-one in the interpolation points domain*[a] *and let* $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ *be a realization satisfying right and left constraints from the data* $(f(\lambda_i), \mathbf{r}_i, \mathbf{w}_i e^{-\lambda_i \tau})$ *and* $(f(\mu_j), \mathbf{l}_j, \mathbf{v}_j e^{-\mu_i \tau})$ *constructed with the Loewner framework (see Theorem 3.15). Then* $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ *satisfies the* right*:*

$$\hat{\mathbf{H}}_d(\lambda_i)\mathbf{r}_i = \mathbf{w}_i, \;\; i = 1, \dots n \tag{6.6}$$

and left constraints*:*

$$\mathbf{l}_j^T \hat{\mathbf{H}}_d(\mu_j) = \mathbf{v}_j, \;\; j = 1, \dots n \tag{6.7}$$

for the given right and left interpolation data.

[a] This means that for any $h_1, h_2 \in \{\lambda_1, \dots, \lambda_n\} \cup \{\mu_1, \dots, \mu_n\}$, then $f(h_1) \neq f(h_2)$ if $h_1 \neq h_2$, where $f(s) = se^{s\tau}$.

Proof. The result for the right constraints (6.6) is obtained as follows: first note that if the delay-free model $\hat{\mathbf{H}}(s)$ satisfies the right constraints for $(f(\lambda_i), \mathbf{r}_i, \mathbf{w}_i e^{-\lambda_i \tau})$, then one obtains:

$$\hat{\mathbf{H}}(f(\lambda_i))\mathbf{r}_i = \mathbf{w}_i e^{-\lambda_i \tau}, \tag{6.8}$$

then, it equivalently follows that:

$$\hat{\mathbf{H}}(f(\lambda_i))e^{\lambda_i\tau}\mathbf{r}_i = \mathbf{w}_i, \tag{6.9}$$

and by invoking Lemma 6.2, we obtain the result:

$$\hat{\mathbf{H}}_d(\lambda_i)\mathbf{r}_i = \mathbf{w}_i. \tag{6.10}$$

The left data constraints (6.7) are similarly obtained. □

Theorem 6.3 provides a method to construct a model $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ whose transfer function $\hat{\mathbf{H}}_d(s) = \hat{C}(s\hat{E} - \hat{A}e^{-s\tau})^{-1}\hat{B}$ interpolates given right and left constraints. This is possible by noticing that the problem can be rewritten as right and left interpolation constraints for the delay-free case for which a realization is obtained by the standard Loewner framework as in Theorem 3.15. A similar reasoning enables the generalization of Theorem 3.16 as stated follows.

Theorem 6.4 (Derivative single delay Loewner framework). *Let us consider a given system represented by its transfer function $\mathbf{G}(s)$, n shift points $\{s_1, \ldots, s_n\} \in \mathbb{C}$ and n left and right tangential directions $\{\mathbf{l}_1, \ldots, \mathbf{l}_n\} \in \mathbb{C}^{n_y}$, $\{\mathbf{r}_1, \ldots, \mathbf{r}_n\} \in \mathbb{C}^{n_u}$. We assume that for all $k \neq m$, $f(s_k) \neq f(s_m)$, where $f(s) = se^{s\tau}$ (f is one-to-one in the interpolation points domain). The n-dimensional single delay model $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$, as in (6.1), interpolates $\mathbf{G}(s)$ as follows, for $k = 1, \ldots, n$:*

$$\mathbf{G}(s_k)\mathbf{r}_k = \hat{\mathbf{H}}_d(s_k)\mathbf{r}_k, \ \mathbf{l}_k^T\mathbf{G}(s_k) = \mathbf{l}_k^T\hat{\mathbf{H}}_d(s_k), \tag{6.11}$$

$$\mathbf{l}_k^T\mathbf{G}'(s_k)\mathbf{r}_k = \mathbf{l}_k^T\hat{\mathbf{H}}_d'(s_k)\mathbf{r}_k, \tag{6.12}$$

if only if the n-dimensional delay-free model $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ is constructed with the derivative Loewner framework as in Theorem 3.16 for the transformed shift points:

$$\big(\sigma_1, \ldots, \sigma_n\big) = \big(f(s_1), \ldots, f(s_n)\big), \tag{6.13}$$

and the transformed transfer function evaluation:

$$\big(\hat{\mathbf{H}}(\sigma_1), \ldots, \hat{\mathbf{H}}(\sigma_n)\big) = \big(\mathbf{G}(s_1)e^{-s_1\tau}, \ldots, \mathbf{G}(s_n)e^{-s_n\tau}\big) \tag{6.14}$$

and the transformed derivative transfer function evaluation:

$$\big(\hat{\mathbf{H}}'(\sigma_1), \ldots, \hat{\mathbf{H}}'(\sigma_n)\big) = \big(\mathbf{F}_1, \ldots, \mathbf{F}_n\big) \tag{6.15}$$

where, for $i = 1, \ldots n$:

$$\begin{aligned} \hat{\mathbf{H}}'(\sigma_i) &= \mathbf{F}_i \\ &= \big(\mathbf{G}'(s_i) - \tau\mathbf{G}(s_i)\big)\left(\frac{e^{-2s_i\tau}}{1+\tau s_i}\right) \end{aligned} \tag{6.16}$$

Proof. First, one can note that a single delay descriptor system can be expressed as

$$\hat{\mathbf{H}}_d(s) = \hat{C}(se^{s\tau}\hat{E} - \hat{A})^{-1}\hat{B}e^{s\tau} = \hat{\mathbf{H}}(f(s))e^{s\tau} \tag{6.17}$$

where $\hat{\mathbf{H}}(s)$ is a descriptor system whose representation is $(\hat{E}, \hat{A}, \hat{B}, \hat{C})$ and $f(s) = se^{s\tau}$. Thus one can use the Loewner matrices to construct the realization of system $\hat{\mathbf{H}}(s)$ for the shift points

$(\sigma_1, \dots \sigma_n) = (f(s_1), \dots, f(s_n))$ whose transfer function data are

$$(\hat{\mathbf{H}}(\sigma_1), \dots, \hat{\mathbf{H}}(\sigma_n)) = (\mathbf{G}(s_1)e^{-s_1\tau}, \dots, \mathbf{G}(s_n)e^{-s_n\tau}).$$

For the transfer function derivative data, one can take the derivative of (6.17) with respect to s written as $\hat{\mathbf{H}}(f(s)) = \mathbf{H}_d(s)e^{-s\tau}$ as follows

$$\hat{\mathbf{H}}(f(s))f'(s) = \hat{\mathbf{H}}'_d(s)e^{-s\tau} - \tau\hat{\mathbf{H}}_d(s)e^{-s\tau},$$

and by solving the equation for $\hat{\mathbf{H}}(\sigma_k)$ one obtains the result. □

This theorem allows to obtain a single delay descriptor system which interpolates any given transfer function $\mathbf{G}(s)$. This can also be used in the case of data obtained through a signal generator, considering that the derivative is accessible as well. Applications of this result can be found in Section 6.4.2.

Now that the state delay Loewner framework has been established, one might be interested in obtaining a good interpolant in the sense of the $\mathcal{H}_2$-norm as formulated in Problem 6.1. We will now formulate the mathematical conditions to select the finite dimensional inspired shift complex points s_i and tangential directions $\mathbf{r}_i$ and $\mathbf{l}_i$.

6.3 Finite dimensional inspired interpolation conditions

In this section, we tackle the Problem 6.1. We recall that this problem was inspired by the finite dimensional $\mathcal{H}_2$ optimality conditions in the delay-free case (see Chapter 4, Section 4.3). Hence, the first step is to characterize the poles of a single-delay reduced order model $\hat{\mathbf{H}}_d$.

6.3.1 Single state-delay model spectrum

Firstly, let us state some preliminary results about the spectrum of a single state-delay model as (6.1).

Proposition 6.5. *Assume that $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ is a single state-delay model and that $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ is its delay-free representation as in Lemma 6.2. Then, $\hat{\lambda}$ is a pole of $\hat{\mathbf{H}}_d$ if and only if $f(\hat{\lambda})$ is a pole of $\hat{\mathbf{H}}$, where $f(s) = se^{s\tau}$.*

Proof. This is a simple implication of Lemma 6.2, which says that $\hat{\mathbf{H}}_d(s) = \hat{\mathbf{H}}\big(f(s)\big)e^{s\tau}$. Since the function $f(s) = se^{s\tau}$ has no pole, then if $\hat{\lambda}$ is a pole of $\hat{\mathbf{H}}_d$, $f(\hat{\lambda})$ is a pole of $\hat{\mathbf{H}}$ and the converse is also true. □

We recall that the poles of $\hat{\mathbf{H}}$ can be simply computed as the generalized eigenvalues of the pencil $(\hat{E}, \hat{A})$. Moreover, the inverse of the function se^s is the scalar Lambert function. It is formally defined as follows

Definition 6.6 (Lambert function $\mathbf{W}_k$). *The Lambert function $\mathbf{W}_k$ is defined implicitly as the solution to the equation*

$$s = \mathbf{W}_k(s)e^{\mathbf{W}_k(s)}. \tag{6.18}$$

For every $s \in \mathbb{C}$ (except at 0), this equation has infinite solutions which are indexed by

$k \in \mathbb{Z}$. We refer to $\mathbf{W}_k(s)$ as the k-th branch of the Lambert function.

The reader might refer to [Corless et al., 1996] for more details and properties of the Lambert function. Hence, the poles of a single state-delay model can be computed as follows.

Corollary 6.7 (Poles of a single state-delay model). *Suppose that $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ is a single state-delay model. Suppose that $\hat{\alpha}_i$ is an eigenvalue of the pencil $(\hat{E}, \hat{A})$, for $i = 1, \ldots, n$. Then the poles of $\hat{\mathbf{H}}_d(s)$ are given by*

$$\hat{\lambda}_k^i = \frac{1}{\tau}\mathbf{W}_k(\tau\hat{\alpha}_i). \tag{6.19}$$

where $k \in \mathbb{Z}$.

Proof. This is a simple implication of Lemma 6.2, which says that $\hat{\mathbf{H}}_d(s) = \hat{\mathbf{H}}\big(f(s)\big)e^{s\tau}$. Since the function $f(s) = se^{s\tau}$ has no poles, then if $\hat{\lambda}$ is a pole of $\hat{\mathbf{H}}_d$, $f(\hat{\lambda})$ is a pole of $\hat{\mathbf{H}}$ and the converse is also true. □

Corollary 6.7 enables to characterize the poles of single state-delay model as function of the eigenvalues of the pencil $(\hat{E}, \hat{A})$. The following example illustrates the poles computation of a single state-delay model.

Example 6.8 (Spectra of an one dimensional time-delay model). *Let $\hat{\mathbf{H}}_d$ be a single delay model whose tranfer function is given by*

$$\hat{\mathbf{H}}_d(s) = \frac{1}{s + e^{-s}}. \tag{6.20}$$

Then, the spectra of $\hat{\mathbf{H}}_d$ can be computed using the Lambert function as follows:

$$\lambda_k \text{ is a pole of } \hat{\mathbf{H}}_d \text{ if and only if } \quad \lambda_k = \mathbf{W}_k(-1), \quad k \in \mathbb{Z}.$$

Figure 6.1 represents, in the complex plane, twenty poles of $\hat{\mathbf{H}}_d$ for $k = -10, \ldots, 9$.

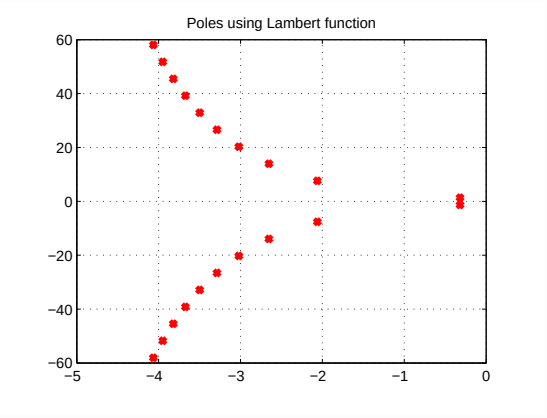

Figure 6.1: Some poles of the model $\hat{\mathbf{H}}_d$ in the complex plane

6.3.2 Truncating the interpolation conditions

Now, using Corollary 6.7, we are able to characterize the finite dimensional inspired interpolation conditions from Problem 6.1 as follows :

Theorem 6.9 (**Characterization of finite dimensional inspired interpolation conditions**). *Suppose that $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ is a n-th order model having only simple poles satisfying the interpolation conditions from Problem 6.1. Then, $\hat{\mathbf{H}}_d$ can be written as a pole-residue decomposition:*

$$\hat{\mathbf{H}}_d(s) = \hat{C}(s\hat{E} - \hat{A}e^{-s\tau})^{-1}\hat{B} = \sum_{i=1}^{n} \frac{\hat{\mathbf{c}}_i \hat{\mathbf{b}}_i^T}{s - \hat{\alpha}_i e^{-s\tau}}. \tag{6.21}$$

Moreover, if for a given $\mathbf{G} \in \mathcal{H}_2$ and a fixed delay $\tau \geq 0$, $\hat{\mathbf{H}}$ satisfies the finite dimensional inspired interpolation conditions from Problem 6.1, then they can be characterized by:

$$\mathbf{G}(-\hat{\lambda}_k^i)\hat{\mathbf{b}}_i = \hat{\mathbf{H}}_d(-\hat{\lambda}_k^i)\hat{\mathbf{b}}_i, \;\; \hat{\mathbf{c}}_i^T\mathbf{G}(-\hat{\lambda}_k^i) = \hat{\mathbf{c}}_i^T\hat{\mathbf{H}}_d(-\hat{\lambda}_k^i) \tag{6.22}$$

$$\hat{\mathbf{c}}_i^T\mathbf{G}(-\hat{\lambda}_{k,p})\hat{\mathbf{b}}_\mathbf{i} = \hat{\mathbf{c}}_i^T\hat{\mathbf{H}}_d(-\hat{\lambda}_{k,p})\hat{\mathbf{b}}_i, \tag{6.23}$$

for all $i = 1, \ldots, n$ and $k \in \mathbb{Z}$ and $\hat{\lambda}_k^i$ is defined by:

$$\hat{\lambda}_k^i = \frac{1}{\tau}\mathbf{W}_k(\tau\alpha_i) \tag{6.24}$$

where $\mathbf{W}_k$ is the k-th branch of the Lambert function.

Theorem 6.9 establishes a characterization of Problem 6.1 as an infinite number of interpolation conditions. This is a consequence that $\hat{\mathbf{H}}_d$ has an infinite number of poles.

Nevertheless, given $\tau \in \mathbb{R}_+$, as $\hat{\mathbf{H}}_d(s) = \hat{C}(s\hat{E} - \hat{A}e^{-s\tau})^{-1}\hat{B}$ is parametrized by a finite number of variables. Indeed, $\hat{\mathbf{H}}_d$ lives in a sub-manifold of dimension $n(n_u + n_y)$, *i.e.*, it can be shown that it is completely parametrized by $n(n_u + n_y)$ variables. This can be simply shown by noticing that there is a simple isomorphism between $(\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ and $(\hat{E}, \hat{A}, \hat{B}, \hat{C}, 0)$ and the last one is parametrized by $n(n_u + n_y)$ variables (see [Byrnes and Falb, 1979; Van Dooren et al., 2008]). Hence, all the interpolation conditions cannot be achieved in the general case, but a finite set only.

Hence, in order to remedy this inconvenient, instead of interpolating an infinite number of conditions, we are going to choose a finite number of conditions which are associated to the poles relative to the principal branch of the Lambert function, *i.e.*, $\mathbf{W}_0$. These conditions are detailed in the following problem:

Problem 6.10 (**Truncated finite dimensional inspired interpolation conditions**). *Given $\mathbf{G}$, a fixed delay $\tau > 0$ and an order n, find a n-th order model $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ having the following poles-residues decomposition:*

$$\hat{\mathbf{H}}_d(s) = \hat{C}(s\hat{E} - \hat{A}e^{-s\tau})^{-1}\hat{B} = \sum_{i=1}^{n} \frac{\hat{\mathbf{c}}_i \hat{\mathbf{b}}_i^T}{s - \hat{\alpha}_i e^{-s\tau}}, \tag{6.25}$$

satisfying the following finite interpolation conditions:

$$\mathbf{G}(-\hat{\lambda}_0^i)\hat{\mathbf{b}}_i = \hat{\mathbf{H}}_d(-\hat{\lambda}_0^i)\hat{\mathbf{b}}_i,\ \hat{\mathbf{c}}_i^T\mathbf{G}(-\hat{\lambda}_0^i) = \hat{\mathbf{c}}_i^T\hat{\mathbf{H}}_d(-\hat{\lambda}_0^i) \tag{6.26}$$

$$\hat{\mathbf{c}}_i^T\mathbf{G}(-\hat{\lambda}_0^i)\hat{\mathbf{b}}_\mathbf{i} = \hat{\mathbf{c}}_i^T\hat{\mathbf{H}}_d(-\hat{\lambda}_0^i)\hat{\mathbf{b}}_i, \tag{6.27}$$

for all $i = 1, \ldots, n$ *where, where* $\hat{\lambda}_0^i$ *is defined by:*

$$\hat{\lambda}_0^i = \frac{1}{\tau}\mathbf{W}_0(\tau\alpha_i) \tag{6.28}$$

where $\mathbf{W}_0$ *is the the principal branch of the Lambert function.*

Obviously, this is similar to a truncation and we cannot guaranty that the eliminated conditions will be satisfied. In Chapter 7, we derive the real $\mathcal{H}_2$ optimality conditions (in a simpler case) for single state-delay models and they are finitely many as well. In what follows, we build a point fixed algorithm to find such a $\hat{\mathbf{H}}_d$ solving Problem 6.10.

6.4 Development of a fixed-point algorithm and numerical applications

6.4.1 Iterative algorithm dTF-IRKA

The algorithm proposed in this section permits to derive a system which satisfies the truncated finite dimensional inspired interpolation conditions (Problem 6.10) for n shift points. The idea behind is based on **TF-IRKA** [Beattie and Gugercin, 2012] which finds a model satisfying the optimality conditions from Theorem 4.14 using a fixed point iteration. For each iteration the new shift points will be the poles located in the principal branch of the Lambert function, only. This algorithm is called **delay TF-IRKA** (or **dTF-IRKA**) and is summed up as follows:

Algorithm 4 dTF-IRKA

1: **Input:** Transfer function $\mathbf{G}(s)$ and reduced order $n \in \mathbb{N}^*$.
2: Make an initial choice for the shift points $\{\sigma_1^0, \ldots, \sigma_n^0\} \in \mathbb{C}$ initial interpolation points and tangential directions $\{\hat{\mathbf{b}}_{1,0}, \ldots, \hat{\mathbf{b}}_{n,0}\} \in \mathbb{C}^{n_u \times 1}$ and $\{\hat{\mathbf{c}}_{1,0}, \ldots, \hat{\mathbf{c}}_{n,0}\} \in \mathbb{C}^{n_y \times 1}$ closed by conjugation.
3: **while** not convergence **do**
4: **Build** $(\hat{E},\ \hat{A},\ \hat{B},\ \hat{C}, \tau)$ using the tangential data $\sigma_i^k, \hat{\mathbf{b}}_{i,k}^T, \hat{\mathbf{c}}_{i,k}$ (Derivative single delay Loewner framework, Theorem 6.4).
5: $k \leftarrow k + 1$.
6: Compute eigenvalue decomposition of $\hat{A}X = \hat{E}X\Sigma$, with $\Sigma = \mathbf{diag}(\hat{\lambda}_{1,k+1}, \ldots, \hat{\lambda}_{n,k+1})$.
7: Set $\sigma_i^{k+1} = -\mathbf{W}_0(\tau\hat{\lambda}_{i,k+1})$, new shift points.
8: Set $\left[\hat{\mathbf{b}}_{1,k+1}^T, \ldots, \hat{\mathbf{b}}_{n,k+1}^T\right] = (\hat{E}X)^{-1}\hat{B}$ and $\left[\hat{\mathbf{c}}_{1,k+1}, \ldots, \hat{\mathbf{c}}_{n,k+1}\right] = \hat{C}X$, new tangential directions.
9: **end while**
10: **Build** $\hat{\mathbf{H}}_d = (\hat{E},\ \hat{A},\ \hat{B},\ \hat{C}, \tau)$.

If the algorithm converges, the approximation model will satisfy the conditions given in Problem 6.10. The **dTF-IRKA** then allows to obtain good (in the sense of the $\mathcal{H}_2$ norm)

shift points and tangential directions for which the interpolation problem will lead to a good approximation model.

As a numerical remark, one should note that the Lambert function evaluated in the principal branch can sometimes associate a real number to a complex one. In this way, the shift points might not be a closed set (by conjugation) and the obtained single delay interpolation model will not have a real representation. To avoid this, one should enforce at each iteration the shift points to be closed by conjugation.

6.4.2 Numerical applications

This section is dedicated to the application of both methods proposed in both Sections 6.2 and 6.4, namely, the delay model interpolation and algorithm **dTF-IRKA**. We will emphasize the potential benefit and effectiveness of the proposed approach.

Example 1: rational interpolation

Let us consider a dynamical model governed by the following delay model $\mathbf{G} \in \mathcal{H}_2$ whose transfer function is given as

$$\mathbf{G}(s) = \frac{2s + 1.3e^{-s}}{s^2 + 1.3se^{-s} + 0.3e^{-2s}}. \tag{6.29}$$

First, model (6.29) (which is obviously of order 2) is approximated by a delay-free model of order $r = 2$ using the **TF-IRKA** (Figure 6.2, green dashed dotted curve). It is also interpolated using the delay Loewner framework with derivatives as stated in Theorem 6.4 whose delay is set to $\tau = 1$, at the shift points $s_1 = 0.1$ and $s_2 = 1$ (Figure 6.2, red dashed thick curve). Results are reported on Figure 6.2, and compared to the original model $\mathbf{G}(s)$ (solid blue line).

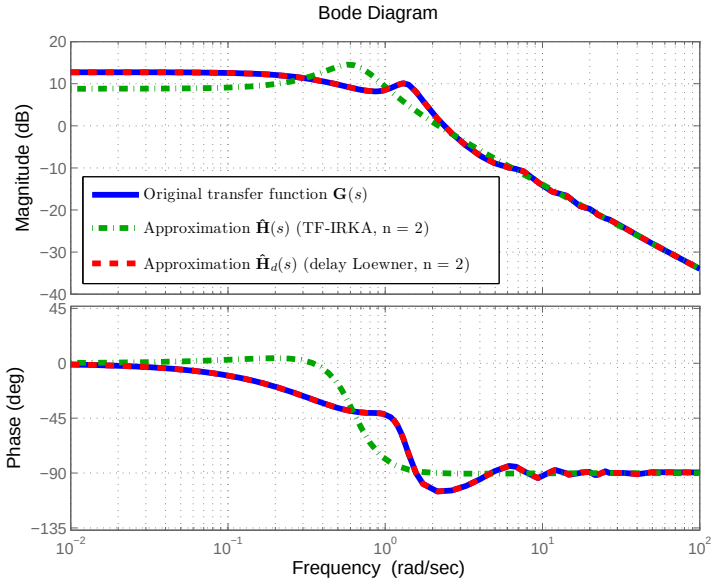

Figure 6.2: Bode diagram of original model (blue solid line), model of order $n = 2$ approximated with **TF-IRKA** (green dashed dotted curve) and delay interpolation model using Theorem 6.4 of order $n = 2$ (red dashed line).

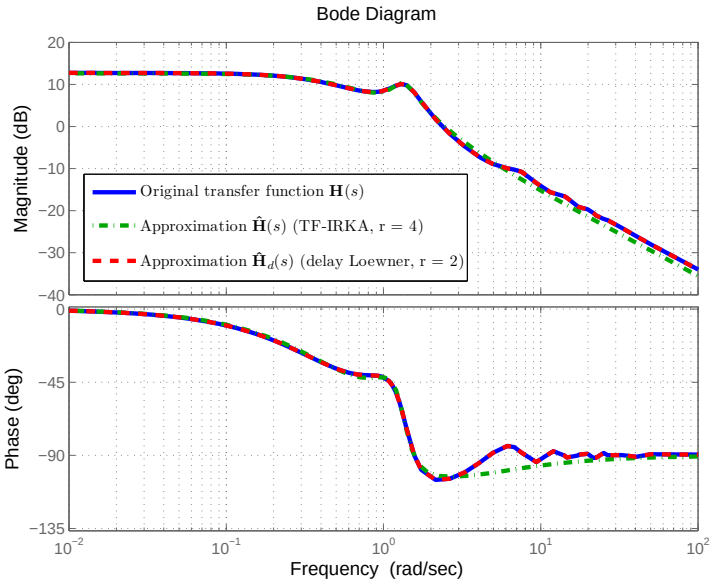

Figure 6.3: Bode diagram of original model (blue solid line), model of order $r = 4$ approximated with **TF-IRKA** (green dashed dotted curve) and delay interpolation model using Theorem 6.4 of order $r = 2$ (red dashed line).

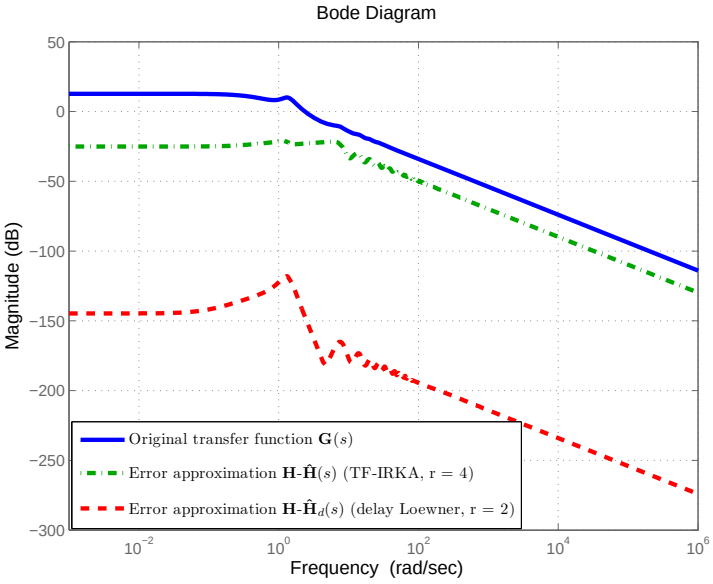

Figure 6.4: Bode diagram of original model (blue solid line), approximation error for delay-free model (**TF-IRKA**, n=2, green dashed dotted curve) and approximation error of the delay interpolation model (Theorem 6.4, $n = 2$,red dashed line).

Figure 6.2 shows that model defined in (6.29) is well interpolated by a delay model obtained by Theorem 6.4, for any interpolation points. Indeed, since the transfer function (6.29) has a realization of the form (6.1) of order 2 it can be reconstructed using the Theorem 6.4. Figure 6.3 shows quite similar results but where **TF-IRKA** has targeted an order $n = 4$.

This specific example clearly emphasizes the fact that, if the original model is a delay one, the counterpart of obtaining a good delay-free approximation (*e.g.*, using **TF-IRKA**) is to increase the approximation order (here the original model of order $n = 2$ must be approximated with an order $n = 4$ to well recover the frequency behavior). As illustrated in Figure 6.4, even with an order $n = 4$, the delay-free model cannot perfectly match the original infinite dimensional one, while the delay model (obtained by Theorem 6.4) provides perfect matching (subject to numerical machine precision errors). On the other hand, the proposed delay Loewner framework allows to find an exact realization.

Example 2: optimal approximation and method scalability

Let us now consider the SISO Los-Angeles Hospital model extracted from the COMPl_eib library [Leibfritz and Lipinski, 2003] whose order is $N = 48$, denoted $\mathbf{G}_{build} = C(sI_{48} - A)^{-1}B \in \mathcal{H}_2$. In order to fit the framework proposed in this paper, a delayed model is constructed by injecting an internal delay $\tau = 0.01$ to all states, *i.e.*, $\mathbf{G}_{delay} = C(sI_{48} - Ae^{-s\tau})^{-1}B$. This last transfer function is firstly interpolated on the basis of a realization of order $n = 10$ by applying the delay Loewner framework from Theorem 6.4 using ten real shift points logarithmically spaced from 0.1 to 1. Then, an approximation is obtained using the **dTF-IRKA** algorithm proposed in Section 6.4. Figure 6.5 compares the Bode plots of these models.

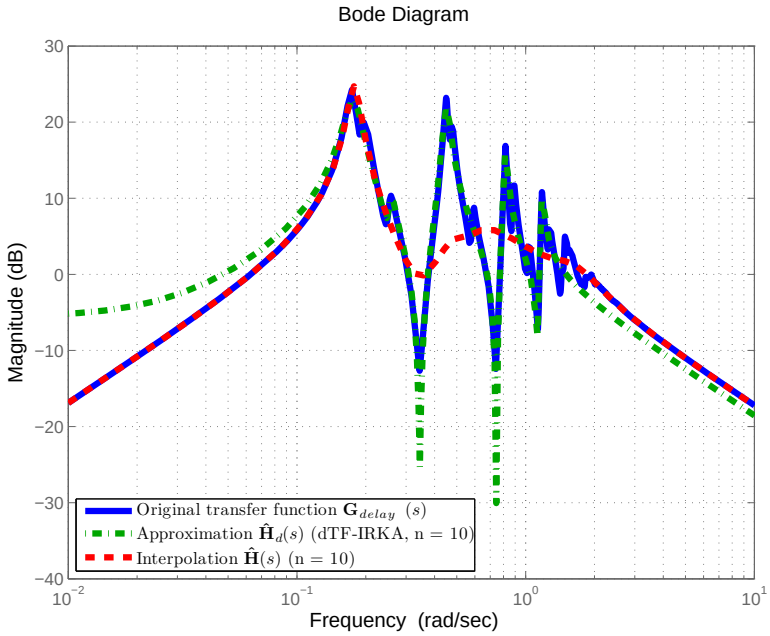

Figure 6.5: Bode diagram of original model (blue curve), delay Loewner interpolation model of order $r = 10$ (green dotted curve) and **dTF-IRKA** of order $r = 10$ (red dashed curve).

As clearly shown on Figure 6.5, the proposed **dTF-IRKA** allows once again to obtain shift points and tangential directions for which the interpolated delay model is much more accurate than the approximation using random shift points. This shows the scalability of the proposed approach for larger models.

Conclusion

In this chapter, the problem of interpolating and approximating any dynamical model (provided by its transfer function or its available evaluation at given points) by a single time-delay one is analyzed. Firstly, we present an extended framework which generalizes the Loewner one [Mayo and Antoulas, 2007] to the case where the interpolant is a single time-delay model. Then, as a second contribution, the finite dimensional inspired interpolation conditions are defined and lead to an infinite set of conditions. Finally, an algorithm, denoted **dTF-IRKA**, allowing to obtain a model which satisfies a finite number of those interpolation conditions is developed and successfully applied to some numerical examples.

This chapter is a first step towards data interpolation for more general structures. In addition, the recent paper [Schulze and Unger, 2015] proposes an extension to the results from this chapter to the case where the reduced order model has more general structure, *e.g.*,

$$\hat{\mathbf{H}}_d(s) = \hat{C}(s\hat{E} - \hat{A}_0 - \hat{A}_1 e^{-s\tau})^{-1}\hat{B}.$$

The main idea is to suppose that the matrix $\hat{A}_1$ is a linear combination of $\hat{E}$ and $\hat{A}_0$, *i.e.*, $\hat{A}_1 = \alpha\hat{E} + \beta\hat{A}_0$. Hence $\hat{\mathbf{H}}_d(s)$ can be written as

$$\begin{aligned} \hat{\mathbf{H}}_d(s) &= \hat{C}(s\hat{E} - \hat{A}_0 - (\alpha\hat{E} + \beta\hat{A}_0)e^{-s\tau})^{-1}\hat{B} \\ \hat{\mathbf{H}}_d(s) &= \hat{C}\left(\frac{s - \alpha e^{-s\tau}}{1 + \beta e^{-s\tau}}\hat{E} - \hat{A}_0\right)^{-1}\hat{B}\frac{1}{1 + \beta e^{-s\tau}} \\ &= \mathbf{G}(f_2(s))\tfrac{1}{1+\beta e^{-s\tau}} \end{aligned} \tag{6.30}$$

where $f_2(s) = \dfrac{s - \alpha e^{-s\tau}}{1 + \beta e^{-s\tau}}$ can be used as the new data transformation, which is an extension of $f(s) = se^{s\tau}$. Furthermore, this paper provides some new results on the moment matching problem for descriptor time-delay systems.

One weakness of the proposed method, is the fact that one should know in advance the delay value τ. Future works will investigate this issue by taking into consideration the delay as a decision variable in the $\mathcal{H}_2$ optimization problem. The extension to multiple delays case will also be addressed in future works. Moreover, next chapter derives the actual $\mathcal{H}_2$ optimality conditions of the model approximation problem when the reduced order model has a state-delay, which turn to be interpolatory conditions of series.

Chapter 7

$\mathcal{H}_2$ optimality conditions derivation for state-delay reduced models

In this chapter, we first attempt to generalize the $\mathcal{H}_2$ optimal interpolation conditions for more general reduced order model structure. To this aim, we first expose the necessary optimality conditions in the case where the reduced system is of dimension one and have a single state delay structure. This can be viewed as a first step towards the $\mathcal{H}_2$ optimal model approximation where the reduced system corresponds to an infinite dimensional one.

The chapter is organized as follows. Firstly, Section 7.1 introduces the $\mathcal{H}_2$ optimal approximation problem when the reduced order model is a single state-delay model of dimension one. In addition, it recalls the spectral decomposition of a time-delay system using the Lambert function and some convergence results. Secondly, Section 7.2.2 presents a spectra-based $\mathcal{H}_2$ inner product characterization for time-delay systems and some other useful identities. Finally, the $\mathcal{H}_2$ optimality conditions are derived and illustrated with an academic example in Section 7.3. These results were partially presented in [Pontes Duff et al., 2016a].

Contents

7.1 Problem statement

Chapter 6 has proposed a way to find single state-delay reduced order models based on the single-delay Loewner framework and the $\mathcal{H}_2$ inspired interpolation conditions. In this chapter, we are interested in finding the exact optimality conditions of the $\mathcal{H}_2$ model approximation problem when the reduced order model has similar structure as the one proposed in Chapter 6.

In order to obtain more general $\mathcal{H}_2$ optimality conditions, one could consider a reduced order model having the following transfer function

$$\hat{\mathbf{H}}_d(s) = \hat{\mathbf{C}}(s)\hat{\mathbf{K}}(s)^{-1}\hat{\mathbf{B}}(s).$$

This expression is referred in [Beattie and Gugercin, 2009a] as generalized coprime factorization. This is a more general than the finite dimensional state-space representation and a great amount of models can be represented by such coprime structure, *e.g.*, time-delay models and second order models. In addition, some interpolation-based model reduction techniques were already developed and reader might refer to [Beattie and Gugercin, 2009a] for structure preserving interpolation.

In this chapter, we consider the case where the reduced order system is described by a *single state-delay equation*, *i.e.*, the reduced order model $\hat{\mathbf{H}}_d$ is governed by the delay differential equation

$$\hat{\mathbf{H}}_d = \left\{ \begin{array}{rcl} \dot{x}(t) & = & \hat{\alpha}x(t-\tau) + \hat{\phi}u(t) \\ y(t) & = & x(t) \end{array} \right. .$$

Therefore, in this case $\hat{\mathbf{B}}(s) = \hat{\phi} \in \mathbb{R}$, $\hat{\mathbf{C}}(s) = 1 \in \mathbb{R}$ and $\hat{\mathbf{K}}(s) = s - \hat{\alpha}e^{-\tau s}$, with $\hat{\alpha} \in \mathbb{R}$ and $\tau \in \mathbb{R}^+$, and its transfer function is given by

$$\hat{\mathbf{H}}_d(s) = \frac{\hat{\phi}}{s - \hat{\alpha}e^{-s\tau}}. \tag{7.1}$$

In this chapter, we will consider that the delay τ is fixed. We call such model $\hat{\mathbf{H}}_d$ a *single reduced time-delay model of dimension one*. The word dimension refers here to the minimal number of DDEs necessary to represent the model. Indeed, even if $\hat{\mathbf{H}}_d$ is an infinite-dimensional model, it only depends on two parameters, namely, $\hat{\phi}$ and $\hat{\alpha}$.

Our objective, in this chapter is, given an original model $\mathbf{G} \in \mathcal{H}_2$, to study the problem of the $\mathcal{H}_2$ optimal approximation by a single reduced time-delay model of dimension one. This approximation problem can be stated as follows :

Problem 7.1 (Single state-delay $\mathcal{H}_2$ model approximation). *Given a stable strictly proper model $\mathbf{G}(s) \in \mathcal{H}_2$ and a delay $\tau \in \mathbb{R}^+$, find a model $\hat{\mathbf{H}}_d = \frac{\hat{\phi}}{s-\hat{\alpha}e^{-s\tau}} \in \mathcal{H}_2$ such that*

$$\|\mathbf{G} - \hat{\mathbf{H}}_d\|_{\mathcal{H}_2} = \min_{(\hat{\phi},\hat{\alpha})\in\mathbb{R}^2} \left\| \mathbf{G} - \frac{\hat{\phi}}{s - \hat{\alpha}e^{-s\tau}} \right\|_{\mathcal{H}_2} \tag{7.2}$$

Although the reduced order system is defined only by two real parameters, *i.e.*, $\hat{\phi}, \hat{\alpha} \in \mathbb{R}$, its infinite dimensional structure generates some very rich $\mathcal{H}_2$ optimality conditions as later illustrated by Theorem 7.8. Moreover, the results presented here can be seen as a first step towards the $\mathcal{H}_2$ optimality conditions for more general delay structured reduced models.

From now on, the model $\hat{\mathbf{H}}_d$ will be decomposed as

$$\hat{\mathbf{H}}_d(s) = \hat{\phi}\mathbf{P}_\tau(s),$$

where $\hat{\phi} \in \mathbb{R}$ is the gain part, and $\mathbf{P}_\tau(s) = \frac{1}{s-\hat{\alpha}e^{-s\tau}}$ is the transfer function containing the spectral information of $\hat{\mathbf{H}}_d$.

As presented in Chapter 4, the pole/residue decomposition of a reduced order model plays an important role in the characterization of the $\mathcal{H}_2$ inner product and, consequently, in the derivation

of the delay-free $\mathcal{H}_2$ optimality conditions (see Section 4.3 from Chapter 4). In this chapter we follow a very similar procedure. Firstly, the spectral decomposition of a single state-delay model will be derived using the Lambert function. Then, based on this spectral decomposition, a new $\mathcal{H}_2$ inner product characterization is derived for single state-delay models. In addition, some other useful theoretical identities are derived. Finally, based on the spectral $\mathcal{H}_2$ inner product characterization, the $\mathcal{H}_2$ optimality conditions from Problem 7.1 are asserted.

7.2 Spectral decomposition and $\mathcal{H}_2$ inner product formulation

7.2.1 Spectral decomposition of a single state-delay model

We recall that $\hat{\mathbf{H}}_d(s) = \frac{1}{s+e^{-s}}$. Hence, in order to evaluate the poles of the model $\mathbf{G}$, one should solve the following equation:

$$s + e^{-s} = 0 \Leftrightarrow se^{s} = -1, \text{ for } \quad s \in \mathbb{C}.$$

One can prove that this equation has an infinite number of zeros and that these values define the Lambert function at point -1. As shown in Chapter 6, a single state-delay model has infinite poles which can be characterized by means of the Lambert function.

The Lambert function has many applications for the analysis and control of time-delay systems (see [Yi, 2009]). Thanks to the Lambert function, one is able to determine the poles and the infinite partial fraction decomposition of $\hat{\mathbf{H}}_d$. This is stated in the following proposition.

Proposition 7.2. *Let $\hat{\mathbf{H}}_d = \frac{\hat{\phi}}{s-\hat{\alpha}e^{-s\tau}}$. Then, the model $\hat{\mathbf{H}}_d$ has infinite poles which can be computed using the Lambert function as follows :*

$$\lambda_k = \frac{1}{\tau}\mathbf{W}_k(\tau\hat{\alpha}), \quad \text{for} \quad k \in \mathbb{Z}, \tag{7.3}$$

and $\hat{\mathbf{H}}_d$ is in $\mathcal{H}_2$ iff $0 < -\hat{\alpha} < \frac{\pi}{2\tau}$. Moreover, if $\hat{\mathbf{H}}_d = \hat{\phi}\mathbf{P}_\tau$, the infinite partial fraction decomposition of $\mathbf{P}_\tau = \frac{1}{s-\hat{\alpha}e^{-s\tau}}$ is given by

$$\mathbf{P}_\tau(s) \quad = \quad \sum_{k=-\infty}^{\infty} \phi_k \frac{1}{s-\lambda_k}, \tag{7.4}$$

where

$$\phi_k = \frac{1}{1+\tau\lambda_k}.$$

Proof. Since $\mathbf{H}_d(s) \to 0$ as $s \to 0$ and $|\mathbf{H}_d(i\omega)|^2 = O\left(\frac{1}{\omega^2}\right)$, it is clear that $\mathbf{H}_d \in \mathcal{H}_2$ if and only if $\mathbf{H}_d$ has no poles in the closed right-half plane. From Example 6.2.4 of [Partington, 2004], this happens if $0 < -\hat{\alpha} < \frac{\pi}{2\tau}$.

In addition, the poles of $\hat{\mathbf{H}}_d(s)$ are the zeros of $\mathbf{K}(s) = s - \hat{\alpha}e^{-s\tau}$. Let us assume that λ_k is the pole on the k-th complex branch. Then, applying the Lambert function definition, we have

$$\lambda_k e^{\lambda_k\tau} = \hat{\alpha} \Leftrightarrow \lambda_k = \frac{1}{\tau}\mathbf{W}_k(\tau\hat{\alpha}).$$

See [Cepeda-Gomez and Michiels, 2015] for more details. In order to obtain the partial fraction decomposition of $\mathbf{P}_\tau(s) = \frac{1}{s-\hat{\alpha}e^{-s\tau}}$, one should apply the residues formula. Since all the poles of $\mathbf{P}_\tau(s)$ are simple, one obtains

$$\begin{aligned}
\phi_k &= \lim_{s\to\lambda_k} \mathbf{P}_\tau(s)(s-\lambda_k) \\
&= \frac{1}{\lim\limits_{s\to\lambda_k} \frac{s-\hat{\alpha}e^{-s\tau}}{s-\lambda_k}} \\
&= \frac{1}{(s-\hat{\alpha}e^{-s\tau})'|_{s=\lambda_k}} \\
&= \frac{1}{1+\tau\hat{\alpha}e^{-\lambda_k\tau}} \\
&\underbrace{=}_{\lambda_k=\hat{\alpha}e^{-\lambda_k\tau}} \frac{1}{1+\tau\lambda_k}.
\end{aligned}$$

See [Pontes Duff et al., 2015c] for additional details. □

One should note that $\hat{\mathbf{H}}_d$ has an infinite number of poles and that is why it represents an infinite dimensional model. We now present the infinite partial fraction decomposition of $\mathbf{P}_\tau^2$, which will be useful for the $\mathcal{H}_2$ optimality conditions.

Proposition 7.3. *Let* $\mathbf{P}_\tau = \frac{1}{s-\hat{\alpha}e^{-s\tau}} \in \mathcal{H}_2$. *Then*

$$\mathbf{P}_\tau^2(s) = \sum_{k=-\infty}^{\infty} \psi_k \frac{1}{(s-\lambda_k)^2} + \rho_k \frac{1}{s-\lambda_k} \tag{7.5}$$

where

$$\psi_k = \frac{1}{(1+\tau\lambda_k)^2} \quad \text{and} \quad \rho_k = \frac{2\tau^2\lambda_k}{(1+\tau\lambda_k)^3}.$$

Proof. A similar procedure is followed here such as the one in Proposition 7.2. First, notice that all the poles of $\mathbf{P}_\tau^2$ have a multiplicity two. That is why the square term is needed as well. To obtain the coefficients one should use the residues formula taking into account that all the poles have multiplicity two. Thus :

$$\left.\begin{aligned}
\psi_k &= \lim_{s\to\lambda_k} \mathbf{P}_\tau^2(s)(s-\lambda_k)^2 \\
&= \frac{1}{\left((s-\hat{\alpha}e^{-s\tau})'|_{s=\lambda_k}\right)^2} \\
&= \frac{1}{(1+\tau\lambda_k)^2}
\end{aligned}\right\} \quad \text{and} \quad \left\{\begin{aligned}
\rho_k &= \lim_{s\to\lambda_k} \frac{d}{ds}\left(\mathbf{P}_\tau^2(s)(s-\lambda_k)^2\right) \\
&= \frac{2\tau^2\lambda_k}{(1+\tau\lambda_k)^3}
\end{aligned}\right.$$

□

Both Propositions 7.2 and 7.3 will be useful to obtain the $\mathcal{H}_2$ optimality conditions for Problem 7.1. Now, let us derive a spectral characterization of the $\mathcal{H}_2$ inner product for single state-delay models.

7.2.2 Spectral $\mathcal{H}_2$ inner product of a single state-delay model

Once we have the spectral decomposition of a single state-delay model, the next step consists in characterizing the $\mathcal{H}_2$ inner product. The following proposition allows to express the $\mathcal{H}_2$ inner product using the spectral decomposition of $\hat{\mathbf{H}}_d$.

Proposition 7.4. ***Spectral $\mathcal{H}_2$ inner product :*** *Let $\mathbf{F} \in \mathcal{H}_2$ be a real stable system and $\mathbf{P}_\tau = \frac{1}{s-\hat{\alpha}e^{-s\tau}} \in \mathcal{H}_2$. Then :*

$$\langle \mathbf{F}, \mathbf{P}_\tau \rangle_{\mathcal{H}_2} = \sum_{k=-\infty}^{\infty} \phi_k \mathbf{F}(-\lambda_k), \tag{7.6}$$

where $\phi_k = \frac{1}{1+\tau\lambda_k}$ and $\lambda_k = \frac{1}{\tau}\mathbf{W}_k(\tau\hat{\alpha})$ for $k \in \mathbb{Z}$.

Proof. One can straightforward writes :

$$\begin{aligned} \langle \mathbf{F}, \mathbf{P}_\tau \rangle_{\mathcal{H}_2} &= \langle \mathbf{F}, \sum \phi_k \tfrac{1}{s-\lambda_k} \rangle_{\mathcal{H}_2} \\ &= \sum \phi_k \langle \mathbf{F}, \tfrac{1}{s-\lambda_k} \rangle_{\mathcal{H}_2} \\ &= \sum \phi_k \mathbf{F}(-\lambda_k) \end{aligned}$$

where we have applied Lemma 4.2 from Chapter 4 in the last step. □

In a similar way, we obtain an expression of the $\mathcal{H}_2$ inner product computation using the partial fraction decomposition of $\mathbf{P}_\tau^2(s)$. This point is the result presented in Proposition 7.5.

Proposition 7.5. ***Spectral $\mathcal{H}_2$ inner product 2 :*** *Let $\mathbf{F} \in \mathcal{H}_2$ and $\mathbf{P}_\tau = \frac{1}{s-\hat{\alpha}e^{-s\tau}}$. Then*

$$\langle \mathbf{F}, \mathbf{P}_\tau^{\mathbf{2}} \rangle_{\mathcal{H}_2} = \sum_{k=-\infty}^{\infty} \rho_k \mathbf{F}(-\lambda_k) - \psi_k \mathbf{F}'(-\lambda_k), \tag{7.7}$$

where $\psi_k = \dfrac{1}{(1+\tau\lambda_k)^2}$, $\rho_k = \dfrac{2\tau^2\lambda_k}{(1+\tau\lambda_k)^3}$ and $\lambda_k = \frac{1}{\tau}\mathbf{W}_k(\tau\hat{\alpha})$ for $k \in \mathbb{Z}$. .

Proof. Similar to Proposition 7.4. Thus

$$\begin{aligned} \langle \mathbf{F}, \mathbf{P}_\tau^2 \rangle_{\mathcal{H}_2} &= \langle \mathbf{F}, \sum \rho_k \tfrac{1}{s-\lambda_k} \rangle_{\mathcal{H}_2} + \langle \mathbf{F}, \sum \psi_k \tfrac{1}{(s-\lambda_k)^2} \rangle_{\mathcal{H}_2}, \\ &= \sum \rho_k \langle \mathbf{F}, \tfrac{1}{s-\lambda_k} \rangle_{\mathcal{H}_2} + \sum \psi_k \langle \mathbf{F}, \tfrac{1}{(s-\lambda_k)^2} \rangle_{\mathcal{H}_2}. \end{aligned}$$

where, in the last step, we apply the Lemma 4.2 ($\mathcal{H}_2$ inner product for first order system) to the first member and Lemma 4.8 ($\mathcal{H}_2$ inner product for higher order system) to the second member. □

In addition, the following result will be useful to demonstrate the main theorem in Section 7.3.

Proposition 7.6. *Let* $\mathbf{P}_\tau = \frac{1}{s-\hat{\alpha}e^{-s\tau}} \in \mathcal{H}_2$. *Then :*

$$\begin{cases} \mathbf{P}_\tau{}'(s) &= \dfrac{-1-\tau\hat{\alpha}e^{-s\tau}}{(s-\hat{\alpha}e^{-s\tau})^2} \\ \dfrac{\partial \mathbf{P}_\tau}{\partial \hat{\alpha}}(s) &= \dfrac{e^{-s\tau}}{(s-\hat{\alpha}e^{-s\tau})^2} \end{cases}$$

which, by association, leads to :

$$\frac{\partial \mathbf{P}_\tau}{\partial \hat{\alpha}}(s) = -\frac{1}{\hat{\alpha}\tau}\Big(\mathbf{P}_\tau{}'(s) + \mathbf{P}_\tau{}^2(s)\Big). \tag{7.8}$$

We are now ready to state the main result of this chapter, namely the $\mathcal{H}_2$ optimality conditions for single reduced time-delay models.

7.3 $\mathcal{H}_2$ optimality conditions for single state-delay models

7.3.1 Derivation of the $\mathcal{H}_2$ optimality conditions

In this section, we state and prove an extension of the interpolation conditions from Theorem 4.14 formulated in Chapter 4 in the case the reduced order model is a single time-delay model of dimension one, *i.e.*, $\hat{\mathbf{H}}_d(s) = \frac{\hat{\phi}}{s-\hat{\alpha}e^{-s\tau}}$. In order to do so, we have followed some equivalent steps as those presented in Chapter 4 to prove delay-free $\mathcal{H}_2$ optimality conditions, namely, (i) we derive the pole residue decomposition of the reduced order model; (ii) we characterize the $\mathcal{H}_2$ inner product for the reduced order models. The next steps are : (a) to characterize the $\mathcal{H}_2$ mismatch error as a function of the reduced order model and (b) compute the gradient of this error to derive the optimality conditions. Both of these steps will be presented in this section.

Firstly, let us write the $\mathcal{H}_2$ mismatch error $\mathcal{E}$ between the full model $\mathbf{G}$ and the reduced model $\hat{\mathbf{H}}_d$ as follows :

$$\mathcal{E}(\hat{\phi},\hat{\alpha}) = \|\mathbf{G}-\hat{\mathbf{H}}_\mathbf{d}\|^2_{\mathcal{H}_2} = \langle \mathbf{G}-\hat{\mathbf{H}}_d, \mathbf{G}-\hat{\mathbf{H}}_\mathbf{d}\rangle_{\mathcal{H}_2}.$$

Here we follow the same reasoning developed in Subsection 4.3.4 from Chapter 4, *i.e.*, the optimality result with respect to parameters presented in Proposition 4.16. Indeed, the results present here are an application of Proposition 4.16.

Let $\Theta = \{\hat{\phi},\hat{\alpha}\}$ be a variable parameter vector defining $\hat{\mathbf{H}}_d$. Let us take the derivative of $\mathcal{E}$ with respect to Θ. Since only $\hat{\mathbf{H}}_d$ depends on Θ, the derivative of the $\mathcal{H}_2$ mismatch error with respect to Θ is given by :

$$\frac{\partial \mathcal{E}}{\partial \Theta} = -2\langle \mathbf{G}-\hat{\mathbf{H}}_d, \frac{\partial \hat{\mathbf{H}}_\mathbf{d}}{\partial \Theta}\rangle_{\mathcal{H}_2}. \tag{7.9}$$

Notice that in order to formally obtain equation (7.9), one needs to assess that the inner-product commutes with the partial derivative. This is made possible thanks to the Lebesgue dominated convergence theorem (see [Pontes Duff et al., 2016a] for more details). Hence, by using the propositions from Section 2, the following results can be stated.

Proposition 7.7. *The partial derivative of the* $\mathcal{H}_2$ *error* $\mathcal{E}$ *with respect to the parameters*

are given analytically by :

$$\begin{cases} \dfrac{\partial \mathcal{E}}{\partial \hat{\phi}} &= -2\langle \mathbf{G} - \hat{\mathbf{H}}_d, \mathbf{P}_\tau \rangle_{\mathcal{H}_2} \\ \dfrac{\partial \mathcal{E}}{\partial \hat{\alpha}} &= \dfrac{2\hat{\phi}}{\hat{\alpha}\tau} \langle \mathbf{G} - \hat{\mathbf{H}}_d, {\mathbf{P}_\tau}' + {\mathbf{P}_\tau}^2 \rangle_{\mathcal{H}_2} \end{cases}$$

Proof. One has to apply the formula (7.9) for each parameter and use $\frac{\partial \hat{\mathbf{H}}_d}{\partial \hat{\phi}} = \mathbf{P}_\tau$ and the formula (7.8). For the second equation, we use the fact that $\frac{\partial}{\partial \hat{\phi}}$ commutes with the $\mathcal{H}_2$ inner-product (see [Pontes Duff et al., 2016a] for more details). □

Finally, by setting those partial derivatives to zero, one obtains the $\mathcal{H}_2$ optimality conditions for Problem 7.1. Moreover, by applying the inner product computation result (7.6), one obtains the generalization of the interpolation conditions from Theorem 4.14.

Theorem 7.8 ($\mathcal{H}_2$ **optimality conditions for single state delay reduced order models).** *Let* $\hat{\mathbf{H}}_d = \frac{\hat{\phi}}{s - \hat{\alpha} e^{-s\tau}} \in \mathcal{H}_2$ *and* $\mathbf{G} \in \mathcal{H}_2$*. Let us suppose also that* $\mathbf{G}' \in \mathcal{H}_2$*. If* $\hat{\mathbf{H}}_d$ *is the best* $\mathcal{H}_2$ *approximation of* $\mathbf{G}$*, then :*

$$\sum_{k=-\infty}^{\infty} \mathbf{G}(-\lambda_k)\phi_k = \sum_{k=-\infty}^{\infty} \hat{\mathbf{H}}_\mathbf{d}(-\lambda_k)\phi_k, \tag{7.10}$$

$$\sum_{k=-\infty}^{\infty} \mathbf{G}'(-\lambda_k)(\phi_k - \psi_k) + \sum_{k=-\infty}^{\infty} \mathbf{G}(-\lambda_k)\rho_k = \sum_{k=-\infty}^{\infty} \hat{\mathbf{H}}'_d(-\lambda_k)(\phi_k - \psi_k) + \sum_{k=-\infty}^{\infty} \hat{\mathbf{H}}_d(-\lambda_k)\rho_k, \tag{7.11}$$

where $\lambda_k = \frac{1}{\tau}\mathbf{W}_k(\tau\hat{\alpha})$*, for* $k \in \mathbb{Z}$*, are the poles of* $\hat{\mathbf{H}}_d$*,* $\phi_k = \dfrac{1}{1+\tau\lambda_k}$*,* $\psi_k = \dfrac{1}{(1+\tau\lambda_k)^2}$ *and* $\rho_k = \dfrac{2\tau^2\lambda_k}{(1+\tau\lambda_k)^3}$.

Proof. Equation (7.10) is obtained by setting $\frac{\partial \mathcal{E}}{\partial \hat{\phi}} = 0$ and using the $\mathcal{H}_2$ inner product computation formula (7.6). Equation (7.11) is obtained by setting $\frac{\partial \mathcal{E}}{\partial \hat{\alpha}} = 0$ which leads to

$$\langle \mathbf{G}, {\mathbf{P}_\tau}' + {\mathbf{P}_\tau}^2 \rangle_{\mathcal{H}_2} = \langle \hat{\mathbf{H}}_d, {\mathbf{P}_\tau}' + {\mathbf{P}_\tau}^2 \rangle_{\mathcal{H}_2}.$$

Hence, using the Hermitian property from the $\mathcal{H}_2$ inner product (see Chapter 4, Proposition 4.7),

$$\langle \mathbf{G}, {\mathbf{P}_\tau}' \rangle_{\mathcal{H}_2} = \langle \mathbf{G}', \mathbf{P}_\tau \rangle_{\mathcal{H}_2},$$

and the result is obtained using the $\mathcal{H}_2$ inner product computation formula (7.6) and (7.7). □

Indeed, Theorem 7.8 generalizes Theorem 4.14 for single state-delay reduced-order models. Due to the infinite dimensional nature of time-delay systems, the $\mathcal{H}_2$ optimality conditions here are no longer interpolation conditions and become interpolation of series depending on both $\mathbf{G}$ and $\hat{\mathbf{H}}_d$ and the spectral decomposition of $\hat{\mathbf{H}}_d$ and $\mathbf{P}_\tau^2$. If we choose $\tau = 0$, *i.e.*, $\hat{\mathbf{H}}_d(s) = \frac{1}{s-\alpha}$, we get back the result from Theorem 4.14, because the sum are over the poles and in this case $\hat{\mathbf{H}}_d$ has only one pole. The reader should remark that the conditions presented in Theorem 7.8 are necessary optimality conditions.

7.3.2 Numerical application

In this section we find one local optimum to Problem 7.1 when $\mathbf{G}$ is a very simple model. Then, we verify numerically that it satisfies the conditions (7.10) and (7.11). Let us consider Problem 7.1 applied to the following very simple model :

$$\mathbf{G}(s) = \frac{10}{s^2 + 11s + 10}.$$

$\mathbf{G}$ and $\mathbf{G}'$ are clearly elements of $\mathcal{H}_2$. Let us fix $\tau = 1$. Then, Problem 7.1 consists in finding $\hat{\phi} \in \mathbb{R}$ and $\hat{\alpha} \in (-\pi/2, 0)$ which minimizes :

$$\mathcal{E}(\hat{\phi}, \hat{\alpha}) = \|\mathbf{G} - \hat{\mathbf{H}}_d\|^2_{\mathcal{H}_2} = \|\mathbf{E}(\hat{\phi}, \hat{\alpha})\|^2_{\mathcal{H}_2}$$

where $\hat{\mathbf{H}}_d(s) = \dfrac{\hat{\phi}}{s - \hat{\alpha}e^{-s}}$.

To find a local minimum the MATLAB function `fminunc` has been used. The criterion was computed solving delay Lyapunov equations (see [Jarlebring et al., 2011]). The delay state space realization of $\mathbf{E}(\hat{\phi}, \hat{\alpha})$ used here was

$$\mathbf{E}(\hat{\phi}, \hat{\alpha}) := \left\{ \begin{array}{rcl} \dot{\mathbf{x}}(t) & = & \mathbf{A}\mathbf{x}(t) + \mathbf{A}_\tau \mathbf{x}(t-\tau) + \mathbf{B}u(t) \\ y(t) & = & \mathbf{C}\mathbf{x}(t) \end{array} \right. ,$$

where

$$\mathbf{A} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -10 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_\tau = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hat{\alpha} \end{bmatrix},$$

$B = \begin{bmatrix} 10/9 \\ -10/9 \\ -\hat{\phi} \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$. The optimal parameters obtained are :

$$\hat{\alpha}^* \approx -0.5371 \quad \hat{\phi}^* \approx 0.4986. \tag{7.12}$$

In Figure 7.1, the error $\mathcal{H}_2$ is represented as function of $\hat{\phi}$ and $\hat{\alpha}$. It shows that $(\hat{\alpha}^*, \hat{\phi}^*)$ is effectively a local minimizer.

Now, let us verify that $\mathbf{G}$ and $\hat{\mathbf{H}}_d$ satisfy the conditions stated in Theorem 7.8. Since this result involves infinite sums, we then rely on a truncation scheme to evaluate them. Let us use the set :

$$\begin{array}{rcl} \mathbf{D_N} & = & \{\text{first } \mathbf{N} \text{ dominant poles of } \mathbf{P}_\tau(s)\} \\ & = & \{\lambda_k = \mathbf{W_k}(-\alpha), k = -N, -N+1, \ldots, N-2, N-1\} \end{array} ,$$

to compute an approximation of the infinite sum. Notice that we have chosen the poles λ_k to go from $-\mathbf{N}$ to $\mathbf{N-1}$ so that $\mathbf{D_N}$ is closed under complex conjugation, *i.e.*, if $\lambda \in \mathbf{D_N}$, $\overline{\lambda} \in \mathbf{D_N}$. In this way, the model constructed by modal truncation is real. Then one should have :

$$S_{1,\mathbf{G},N} = \sum_{k=-N}^{N-1} \mathbf{G}(-\lambda_k)\phi_k \approx \sum_{k=-N}^{N-1} \hat{\mathbf{H}}_d(-\lambda_k)\phi_k = S_{1,\hat{\mathbf{H}}_d,N},$$

and

$$S_{2,\mathbf{G},N} = \sum_{k=-N}^{N-1} \mathbf{G}'(-\lambda_k)(\phi_k - \psi_k) + \sum_{k=-N}^{N-1} \mathbf{G}(-\lambda_k)\rho_k,$$

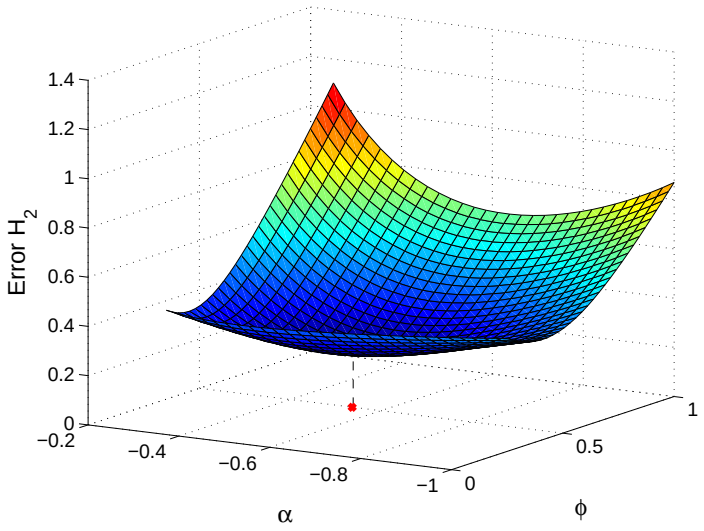

Figure 7.1: Error $\mathcal{H}_2$ as function of $\hat{\phi}$ and $\hat{\alpha}$ (colored surfaced) & optimal parameters (red dot)

$$S_{2,\hat{\mathbf{H}}_d,N} = \sum_{k=-N}^{N-1} \hat{\mathbf{H}}_d'(-\lambda_k)(\phi_k - \psi_k) + \sum_{k=-N}^{N-1} \hat{\mathbf{H}}_d(-\lambda_k)\rho_k,$$

and

$$S_{2,\mathbf{G},N} \approx S_{2,\hat{\mathbf{H}}_d,N}.$$

The following table shows the truncation results for different number of poles :

N	$S_{1,\mathbf{G},N}$	$S_{1,\hat{\mathbf{H}}_d,N}$	$S_{2,\mathbf{G},N}$	$S_{2,\hat{\mathbf{H}}_d,N}$
2	0.80890	0.80326	0.15280	0.15361
6	0.81620	0.81234	0.15302	0.15310
10	0.81656	0.81410	0.15306	0.15308
200	0.81667	0.81655	0.15307	0.15307

Figure 7.2: Truncation evaluation of the conditions from Theorem 7.8 for $N = 2, 6, 10, 200$ poles.

This approximation shows that the truncation conditions are close to the optimality conditions from Theorem 7.8.

Conclusion

In this chapter, the $\mathcal{H}_2$ optimality conditions for single reduced time-delay models of dimension one were derived as an extension of the interpolation results from the delay-free case (see Theorem 4.14). In order to prove the main theorem, some results about the spectral decomposition of $\hat{\mathbf{H}}_d$ and the $\mathcal{H}_2$ inner product computation were exposed. In addition, an academic example were presented in order to illustrate these conditions. These results have a great potential to be

extended to more complex structured reduced models and an algorithm needs to be developed in order to find reduced order models satisfying these new optimality conditions.

Part III

Stability charts of time-delay systems and model approximation

Chapter 8

Model approximation framework for evaluating time-delay systems' stability

So far, our central object of study was the $\mathcal{H}_2$ approximation problem when the reduced-order model has a time-delay structure, and only stable dynamical systems were considered. Part II was dedicated to study this problem, especially for two different structures, namely : (i) input and output delays and (ii) single state delay. In addition, until now, all models were considered to be in $\mathcal{H}_2$, *i.e.*, stable with bounded impulse response energy.

In this part of the manuscript, corresponding to Chapters 8 and 9, we follow a completely different philosophy from the content presented until now. We consider models which represent LTI systems, not necessarily stable, having multiple state delays. This means that their transfer functions are not necessarily elements of $\mathcal{H}_2$ and the results presented so far are no longer directly applicable. Our main focus here is to analyze the stability of Time-Delay Systems (TDS) using interpolation based model approximation methods.

In this chapter, in particular, we consider TDS as being elements of the Hilbert space $\mathcal{L}_2(i\mathbb{R})$ (defined in Chapter 2). Our main tool here is the orthogonal decomposition $\mathcal{L}_2(i\mathbb{R}) = \mathcal{H}_2(\mathbb{C}^-) \bigoplus \mathcal{H}_2(\mathbb{C}^+)$. Based on this, in Section 8.2, we give a Hardy space characterization for systems' stability and we derive some theoretical results linking stability and $\mathcal{L}_2(i\mathbb{R})$ model approximation. The main contribution, stated on Theorem 8.6 in Section 8.2, shows that, given a TDS represented by $\mathbf{G}_d$, if one find an unstable approximation $\hat{\mathbf{H}}$ of a $\mathbf{G}_d$ for which the mismatch error $\mathcal{L}_2(i\mathbb{R})$ norm $\|\hat{\mathbf{H}} - \mathbf{G}_d\|_{\mathcal{L}_2(i\mathbb{R})}$ is small enough, then $\mathbf{G}_d$ is also unstable. Finally, in Section 8.3, the algorithm **TF-IRKA** (see [Beattie and Gugercin, 2012]), presented in Chapter 4, is applied to produce such approximations. The framework presented here is exploited in Chapter 9 to produce stability charts estimation of a TDS.

Contents

8.1 Introduction

8.1.1 Context

The so-called LTI time-delay dynamical systems is a broad class of systems that can model a wide range of phenomena (see [Gu et al., 2003] for some concrete examples). Delays might cause instability and their effect should be taken into account. Because of its infinite dimensional behavior, standard techniques based on finite dimensional state-space representation are not directly applicable and dedicated developments are necessary.

The stability analysis of any time-delay system, for example, cannot be simply established solving a matrix eigenvalue problem, as in the finite dimensional case. Indeed, TDS can be seen as systems governed by an abstract differential equation[1] in a Banach space and the eigenvalue problem becomes itself a spectral problem of a linear infinite dimensional operator (see [Curtain and Zwart, 2012; Partington, 2004]).

Henceforth, we considered discrete[2] retarded LTI time-delay system (shortly, LTI TDS), only. Mathematically, the family of SISO dynamical systems $\mathbf{G}_d$ governed by the following retarded discrete LTI TDS

$$\mathbf{G}_d := \left\{ \begin{array}{rcl} E\dot{\mathbf{x}}(t) & = & A_0\mathbf{x}(t) + \sum_{k=1}^{n_d} A_k\mathbf{x}(t-\tau_k) + \mathbf{b}u(t) \\ \mathbf{y}(t) & = & \mathbf{c}\mathbf{x}(t) \end{array} \right. , \tag{8.1}$$

is considered, where $\mathbf{b}, \mathbf{c}^T \in \mathbb{R}^N$, E, A_0, $A_k \in \mathbb{R}^{N\times N}$, and $\tau_k \in \mathbb{R}^+$ for $k = 1, \ldots, n_d$ (N denotes the dimension of $\mathbf{x}(t)$). The initial condition of (8.1) is $\mathbf{x}(\theta) = \phi(\theta)$, $\theta \in [-\tau, 0]$ where $\tau = \max\{\tau_1, \ldots, \tau_{n_d}\}$. Notice that the state of (8.1) at time t is completely determined by the function $\mathbf{x}_t(\theta)$, with $\theta \in [-\tau, 0]$. Since it does not play a role in stability analysis, the feedthrough term D is considered to be zero. We recall that the transfer function associated with $\mathbf{G}_d$, between the input $u(t)$ and the output $y(t)$, is given by :

$$\mathbf{G}_d(s, \Delta) = \mathbf{c}\mathcal{K}(s, \Delta)^{-1}\mathbf{b}, \tag{8.2}$$

where

$$\mathcal{K}(s, \Delta) = sE - A_0 - \sum_{k=1}^{n_d} A_k e^{-s\tau_k}, \tag{8.3}$$

and $\Delta = \mathbf{diag}(\tau_1, \ldots, \tau_{n_d}) \in \mathbb{R}^{n_d\times n_d}$. We define the characteristic quasi-polynomial $p(s, \Delta)$ of (8.2) as

$$\begin{array}{rcl} p(s, \Delta) & = & \det\left(\mathcal{K}(s, \Delta)\right) \\ & = & \det(sE - A_0 - \sum_{k=1}^{n_d} A_k e^{-s\tau_k}). \end{array} \tag{8.4}$$

[1] the system is governed by $\dot{\mathbf{x}}_t = \mathcal{A}\mathbf{x}_t + \mathcal{B}u(t)$, where $\mathbf{x}_t$ is a function and $\mathcal{A}$, $\mathcal{B}$ are bounded or unbounded operators over these functions.

[2] discrete here means a finite number of delays in contrast to distributed delay systems.

Notice that the highest order in the s variable of the characteristic quasi-polynomial $p(s,\Delta)$ associated to (8.3) is n. In addition, $p(s,\Delta)$ does not have any term of the form $s^n e^{-s\gamma}$, otherwise $\mathbf{G}_d$ would be a neutral system. For LTI TDS, finding the roots of $p(s,\Delta)$ is intrinsically related to determining the stability of model $\mathbf{G}_d$. However, because of the exponential terms, the characteristic quasi-polynomial $p(s,\Delta)$ has an infinite number of roots, and the problem of assessing the stability of $\mathbf{G}_d$ might be hard in practice.

Among the various related tools to investigate stability of a TDS, the following list is highlighted :

- **Zeros of a quasi-polynomial:** as stated before, the zeros of the characteristic quasi-polynomial $p(s,\Delta)$ plays an important role in the stability analysis of any TDS. Indeed, for retarded TDS, if there are no roots in the closed right half-plane, then the system is stable. Historically, the asymptotic location of these roots was firtly studied in [Pontryagin, 1955]. In the context of time-delay systems, the first systematic and complete reference was [Bellman and Cooke, 1963]. See also [Hale and Lunel, 2013; Gu et al., 2003] for further details.

- **Razumikhin theory and Lyapunov-Krasovskii theory :** this category congregates all the time-delay techniques which generalizes the Lyapunov based methods. Most strategies are based on the generalized Lyapunov methods, involving some functionals instead of a classical positive definite functions. As an example, if the Lyapunov-Krasovskii functional candidate
$$V(\mathbf{x}_t) = \mathbf{x}_t(0)^T P \mathbf{x}_t(0) + \int_{-\tau}^{0} \mathbf{x}_t(\theta) S \mathbf{x}_t(\theta) d\theta,$$
where $P, S \in \mathbb{R}^{n\times n}$, and $P > 0$, $S \geq 0$, satisfies some conditions, the TDS
$$\dot{\mathbf{x}}(t) = A_0 \mathbf{x}(t) + A_1 \mathbf{x}(t-\tau),$$
is asymptotically stable. The historical references are [Krasovskii and Brenner, 1963] and [Razumikhin, 1956], and the reader might refer to [Datko, 1978; Niculescu, 2001a] for more details.

- **Spectral discretization methods:** this category relies on the approximation of the infinite dimensional eigenvalue problem into a finite dimensional one. Several numerical approaches for computing characteristic roots exist and the most commonly known are based on (i) the discretization of the solution operator associated to the DDEs (for example [Breda, 2006] and [Engelborghs and Roose, 2002]), or, (ii) the discretization of the infinitesimal generator of the solution operators (see for example [Breda et al., 2005] and [Wu and Michiels, 2012]).

The reader might refer to the monographs [Gu et al., 2003; Michiels and Niculescu, 2014] for a general overview. This chapter and Chapter 9 aim at proposing a methodology based on interpolatory model approximation to address the problem of estimating the stability of a LTI TDS. In particular, this chapter provides a model approximation framework enabling to evaluate if a given time-delay system is unstable. The philosophy of this chapter is presented as follows.

Philosophy of the proposed method

The purposes of this chapter is to address the LTI TDS stability problem through a novel angle: *the model approximation in the interpolatory framework* one. More precisely, given a model of

the form (8.1) or (8.2), the objective is to find a finite dimensional n-th order model $\hat{\mathbf{H}}$ of the form:

$$\hat{\mathbf{H}} := \left\{ \begin{array}{rcl} \hat{E}\dot{\hat{\mathbf{x}}}(t) & = & \hat{A}\hat{\mathbf{x}}(t) + \hat{\mathbf{b}}\mathbf{u}(t) \\ \hat{\mathbf{y}}(t) & = & \hat{\mathbf{c}}\hat{\mathbf{x}}(t) \end{array} \right. \tag{8.5}$$

with transfer function $\hat{\mathbf{H}}(s) = \hat{\mathbf{c}}(s\hat{E} - \hat{A})^{-1}\hat{\mathbf{b}} \in \mathcal{L}_2(i\mathbb{R})$, where $\hat{\mathbf{b}}, \hat{\mathbf{c}} \in \mathbb{R}^n$ and $\hat{E}, \hat{A} \in \mathbb{R}^{n\times n}$, such that (8.5) well reproduces the (in)stability property of (8.1) or (8.2)[3]. Here, n denotes the order of the finite dimensional approximation model. If such an approximation is achieved, clearly, the exact eigenvalues computation of the simple model $\hat{\mathbf{H}}$ provides the stability property. In addition, we will provide some *a posteriori* results which enable to certificate that if the approximation (8.5) is unstable and close enough to (8.1), then (8.1) will be also unstable.

The chapter is organized as follows :

(i) In this section, we recall some interesting nice spectral properties of LTI TDS. Although these results are well know in the literature related to LTI TDS, we recall them for unfamiliar reader.

(ii) In Section 8.2, we consider LTI TDS as being elements of $\mathcal{L}_2(i\mathbb{R})$. Firstly, we characterize stability using subsets of $\mathcal{L}_2(i\mathbb{R})$. The main result of this section is the Theorem 8.6. It states that if $\mathbf{G}_d$ is an LTI TDS, $\hat{\mathbf{H}}$ is an unstable approximation of $\mathbf{G}_d$ and the $\mathcal{L}_2(i\mathbb{R})$ norm of the approximation error $\boldsymbol{\mathcal{E}} = \mathbf{G}_d - \hat{\mathbf{H}}$ is small enough, then $\mathbf{G}_d$ is also unstable. Hence, if one is able to find $\hat{\mathbf{H}}$, an unstable good enough $\mathcal{L}_2(i\mathbb{R})$ approximation of $\mathbf{G}_d$, then $\mathbf{G}_d$ will be also unstable.

(iii) Then, Section 8.3 is dedicated to the construction of finite dimensional reduced order models that can be used to detect instability exploiting the theoretical results presented in Section 8.2. Hence, we recall some $\mathcal{L}_2(i\mathbb{R})$ model approximation results. Since they are not applicable for solving the stability problem, we propose some relaxed conditions inspired of the $\mathcal{H}_2$ model approximation, that are achieved using the algorithm **TF-IRKA**.

8.1.2 Time-delay system properties

As exposed before, the spectrum of an LTI TDS is determined by the zeros of its characteristic quasi-polynomial $p(s, \Delta)$. Even if $p(s, \Delta)$ is a transcendental function and has infinite number of zeros, it has some nice qualitative properties. Some of these properties, which are useful in this chapter, are recalled without demonstration and the reader can refer to [Bellman and Cooke, 1963] and [Michiels and Niculescu, 2014] for further information.

Property 8.1 (Properties of LTI TDS). *A retarded LTI discrete time-delay system (LTI TDS) have the following properties:*

(i) ***Asymptotic stability of LTI TDS :*** *for an LTI TDS as in (8.1), a necessary and sufficient condition for the system in (8.1) to be asymptotically stable is to have no pole in the closed right half-plane.*

(ii) ***Finite number of unstable poles:*** *a LTI TDS has at most a finite number of characteristic roots in the right-half plane.*

[3]Note that the paradigm presented here is different to the traditional model approximation one. Here the (in)stability is the main interest instead of the input/output behavior. This is also a justification for the limitation to the **SISO** systems case.

(iii) ***Continuity of the spectral abcissa:*** *the spectral abcissa function :*

$$\begin{array}{rcl} \alpha & : \mathbb{R}^{n_d \times n_d} \times \mathbb{R}^{n \times n \times n_d} & \rightarrow \quad \mathbb{R} \\ \alpha & (\Delta, A_0, A_1, \ldots, A_{n_d}) & := \quad \sup\{\mathrm{Re}(\lambda), \quad p(\lambda, \Delta) = 0\}, \end{array} \tag{8.6}$$

is continuous.

(iv) ***Loss/acquisition of stability:*** *if the matrices $A_0, A_1, \ldots, A_{n_d}$ and the delays Δ are varied, then the loss or acquisition of stability of an LTI TDS as in (8.1) is associated with characteristic roots on the imaginary axis.*

Therefore, for an LTI TDS to be stable, it is enough to show that its characteristic quasi-polynomial $p(s, \Delta)$ has no root in the right half plane. Reader should note that this condition is no longer sufficient for other classes of TDS, such as neutral systems (see [Partington and Bonnet, 2004] for more details). Moreover, even if an LTI TDS has an infinite number of poles, only finitely many will be unstable and its poles move continuously with respect to small perturbations.

It is worth noting that the spectral abcissa is also continuous when $\Delta = 0$, *i.e*, the delays are set to 0. In this case, the system $\mathbf{G}_d$ is a finite dimensional one and its poles can be easily computed. If the delays Δ pass from $\Delta = 0$ to $\|\Delta\| \leq \varepsilon$, for ε small, the number of poles of $\mathbf{G}_d$ goes from finite to infinite. However, as ε tends to 0, those infinite new poles will satisfy $\mathrm{Re}(s) \rightarrow -\infty$. This observation is relevant for the stability charts evaluation. Indeed, the stability of the system when $\Delta = 0$ (delay free) is easily computed by solving a finite dimensional eigenvalue problem. Furthermore, the system remains stable with respect to the parameters $(\Delta, A_0, A_1, \ldots, A_{n_d})$ until a pole crosses the imaginary axis (see item (iv) from Property 8.1).

In what follows, we will consider an LTI TDS $\mathbf{G}_d$ as being an element of the space $\mathcal{L}_2(i\mathbb{R})$. This is a Hilbert space equipped with the $\mathcal{L}_2(i\mathbb{R})$ inner product (introduced in Chapter 2, Section 2.1), and can be decomposed as

$$\mathcal{L}_2(i\mathbb{R}) = \mathcal{H}_2(\mathbb{C}^-) \bigoplus \mathcal{H}_2(\mathbb{C}^+).$$

Moreover,

$$\mathcal{H}_2(\mathbb{C}^-)^\perp = \mathcal{H}_2(\mathbb{C}^+).$$

Hence, given $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$, then there exist $\mathbf{G}_d^s \in \mathcal{H}_2(\mathbb{C}^+)$ and $\mathbf{G}_d^a \in \mathcal{H}_2(\mathbb{C}^-)$[4] , such that

$$\mathbf{G}_d = \mathbf{G}_d^s + \mathbf{G}_d^a, \quad \langle \mathbf{G}_d^s, \mathbf{G}_d^a \rangle_{\mathcal{L}_2(i\mathbb{R})} = 0.$$

Therefore, the pythagorean property is satisfied

$$\|\mathbf{G}_d\|^2_{\mathcal{L}_2(i\mathbb{R})} = \|\mathbf{G}_d^s\|^2_{\mathcal{H}_2(\mathbb{C}^+)} + \|\mathbf{G}_d^a\|^2_{\mathcal{H}_2(\mathbb{C}^-)}. \tag{8.7}$$

This decomposition plays a very important role in this chapter and it will be used to characterize the stability of any LTI TDS.

8.2 $\mathcal{L}_2(i\mathbb{R})$ topology for unstable systems

In this section, we consider an LTI TDS to be an element of $\mathcal{L}_2(i\mathbb{R})$. Our main goal is to show that if $\mathbf{G}_d$ is unstable and $\hat{\mathbf{H}}$ is a "good enough" approximation of $\mathbf{G}_d$, then it is also unstable. Let us start by showing why an LTI TDS can be viewed as an element of $\mathcal{L}_2(i\mathbb{R})$ and how we characterize stability through this space.

[4] the subindex s and a comes from stable and antistable.

8.2.1 $\mathcal{L}_2(i\mathbb{R})$ characterization of stability

The following result shows that the transfer function of $\mathbf{G}_d$ is an element of $\mathcal{L}_2(i\mathbb{R})$.

Proposition 8.2 (LTI TDS as element of $\mathcal{L}_2(i\mathbb{R})$). *Let $\mathbf{G}_d$ be a LTI TDS having no pole over the imaginary axis. Then it satisfies*

$$\int_{-\infty}^{\infty} |\mathbf{G}_d(i\omega)|^2 d\omega < \infty.$$

Hence, $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$.

Proof. The transfer function of any LTI TDS given by (8.2) is strictly proper and has the following asymptotic behavior[5] :

$$|\mathbf{G}_d(i\omega)|^2 = \left(\left|\frac{n(i\omega)}{d(i\omega)}\right|\right)^2 = O\left(\frac{1}{\omega^2}\right)$$

Hence, if $|\mathbf{G}_d(i\omega)|$ is continuous over $\mathbb{R}$, which means that it has no pole over the imaginary axis, then

$$\int_{-\infty}^{\infty} |\mathbf{G}_d(i\omega)|^2 d\omega < \infty,$$

which implies that $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$. □

The following result shows a $\mathcal{L}_2(i\mathbb{R})$ characterization of the stability:

Proposition 8.3 ($\mathcal{L}_2(i\mathbb{R})$ stability characterization). *Let $\mathbf{G}_d$ be an LTI TDS having no pole on the imaginary axis. Then, there exist*

$$\mathbf{G}_d^s \in \mathcal{H}_2(\mathbb{C}^+) \quad \text{and} \quad \mathbf{G}_d^a \in \mathcal{H}_2(\mathbb{C}^-),$$

where $\mathbf{G}_d^a$ has a finite number of poles, such that

$$\mathbf{G}_d = \mathbf{G}_d^s + \mathbf{G}_d^a, \text{ and } \langle \mathbf{G}_d^s, \mathbf{G}_d^a \rangle_{\mathcal{L}_2(i\mathbb{R})} = 0.$$

Moreover, the system $\mathbf{G}_d$ is stable if and only if

$$\text{proj}_{\mathcal{H}_2(\mathbb{C}^-)} \mathbf{G}_d = \mathbf{G}_d^a = 0.$$

In other words,[a]

1. *$\mathbf{G}_d$ is **stable** if and only if $\mathbf{G}_d \in \mathcal{H}_2(\mathbb{C}^+)$.*
2. *$\mathbf{G}_d$ is **unstable** if and only if $\mathbf{G}_d \in (\mathcal{H}_2(\mathbb{C}^+))^c = \mathcal{L}_2(i\mathbb{R}) \backslash \mathcal{H}_2(\mathbb{C}^+)$.*

[a] $\mathbf{G}_d \in (\mathcal{H}_2(\mathbb{C}^+))^c$ denotes the complement set of $\mathcal{H}_2(\mathbb{C}^+)$ with respect to $\mathcal{L}_2(i\mathbb{R})$.

Proof. The decomposition result follows from $\mathcal{L}_2(i\mathbb{R}) = \mathcal{H}_2(\mathbb{C}^-) \bigoplus \mathcal{H}_2(\mathbb{C}^+)$ (see Proposition 2.3 in Chapter 2). In addition, $\mathbf{G}_d^a$ has a finite number of poles because of item (ii) of Property 8.1. Finally, a causal LTI TDS is stable if all of its poles lie in the left-half plane, which is equivalent to $\mathbf{G}_d^a = 0$. □

[5] We say that $f = O(g)$ if there exists a positive real number M and a real number x_0 such that $|f(x)| \leq M|g(x)|$, if $x \geq x_0$.

Hence, Proposition 8.2 provides a Hardy space characterization for LTI TDS stability. Since $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$, it is stable if and only if its projection onto $\mathcal{H}_2(\mathbb{C}^-)$ is null. Moreover, the set $\mathcal{H}_2(\mathbb{C}^+)$ contains the stable LTI TDS and the set $\mathcal{L}_2(i\mathbb{R})\backslash\mathcal{H}_2(\mathbb{C}^+)$ contains the unstable LTI TDS. This characterization of the notion of stability motivates the use of the $\mathcal{L}_2(i\mathbb{R})$ tools, and this is the subject of the following subsection.

8.2.2 Topological and approximation results in $\mathcal{L}_2(i\mathbb{R})$

In what follows, we provide some results connecting $\mathcal{L}_2(i\mathbb{R})$-topology, $\mathcal{L}_2(i\mathbb{R})$ approximation and LTI TDS stability. The first one concerns the topology of the set of stable systems :

Proposition 8.4 (**Unstable $\mathcal{L}_2(i\mathbb{R})$ approximation of stable model).** *For every stable system $\mathbf{G}_d \in \mathcal{H}_2(\mathbb{C}^+)$, there exists a sequence of unstable systems $\hat{\mathbf{H}}_k \in \mathcal{L}_2(i\mathbb{R})\backslash\mathcal{H}_2(\mathbb{C}^+)$, $k \in \mathbb{N}^*$, such that*

$$\|\mathbf{G}_d - \hat{\mathbf{H}}_k\|_{\mathcal{L}_2(i\mathbb{R})} \to 0, \quad \text{when } k \to \infty \tag{8.8}$$

In other words, the set of stable systems is not an open set of $\mathcal{L}_2(i\mathbb{R})$.

Proof. Given $\mathbf{G}_d \in \mathcal{H}_2(\mathbb{C}^+)$, let $\mathbf{h} \in \mathcal{H}_2(\mathbb{C}^-)$ be an element such that $\|\mathbf{h}\|_{\mathcal{L}_2(i\mathbb{R})} = 1$. The system $\mathbf{H}_k = \mathbf{G}_d + \frac{1}{k}\mathbf{h} \in \mathcal{L}_2(i\mathbb{R})\backslash\mathcal{H}_2(\mathbb{C}^+)$ and $\|\mathbf{G}_d - \hat{\mathbf{H}}_k\|_{\mathcal{L}_2(i\mathbb{R})} = \frac{1}{k}\|\mathbf{h}\|_{\mathcal{L}_2(i\mathbb{R})} \to 0$ when $k \to \infty$. □

Proposition 8.4 tells that it is possible to approximate[6] a stable LTI TDS by a sequence of unstable systems. In other words, given $\mathbf{G}_d$ a stable LTI TDS, an approximation $\hat{\mathbf{H}}$ might be unstable, even if the $\mathcal{L}_2(i\mathbb{R})$ approximation error $\|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})}$ is small enough. This is not the case for unstable systems as stated in the following proposition.

Proposition 8.5 ($\mathcal{L}_2(i\mathbb{R})$ **topology of unstable systems).** *The set of unstable systems $\mathcal{L}_2(i\mathbb{R})\backslash\mathcal{H}_2(\mathbb{C}^+)$ is an open set of $\mathcal{L}_2(i\mathbb{R})$. In order words, given an unstable system $\mathbf{H} \in \mathcal{L}_2(i\mathbb{R})\backslash\mathcal{H}_2(\mathbb{C}^+)$, there exists an $\varepsilon > 0$ such that the ball $B_\varepsilon(\mathbf{H}) = \{\mathbf{G} \in \mathcal{L}_2(i\mathbb{R}) | \|\mathbf{G} - \mathbf{H}\|_{\mathcal{L}_2(i\mathbb{R})} < \varepsilon\} \subset \mathcal{L}_2(i\mathbb{R})\backslash\mathcal{H}_2(\mathbb{C}^+)$.*

Proof. Since $\mathcal{H}_2(\mathbb{C}^+)$ is a $\mathcal{L}_2(i\mathbb{R})$ closed set, its complement $(\mathcal{H}_2(\mathbb{C}^+))^c = \mathcal{L}_2(i\mathbb{R})\backslash\mathcal{H}_2(\mathbb{C}^+)$ is open. □

Proposition 8.5 states that unstable systems form an open set in the sense of the $\mathcal{L}_2(i\mathbb{R})$ norm. This means that, given $\mathbf{G}_d$, if there exists a unstable system $\hat{\mathbf{H}}$ closed enough to $\mathbf{G}_d$, then $\mathbf{G}_d$ will be unstable as well. Moreover, the following theorem provides an instability guaranty certificate.

Theorem 8.6 (**Instability certificate for an LTI TDS).** *Let $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$ and let us consider an approximation $\hat{\mathbf{H}} \in \mathcal{L}_2(i\mathbb{R})$. Let $\hat{\mathbf{H}} = \hat{\mathbf{H}}^s + \hat{\mathbf{H}}^a$ be the decomposition of $\hat{\mathbf{H}}$ in $\mathcal{L}_2(i\mathbb{R}) = \mathcal{H}_2(\mathbb{C}^-) \bigoplus \mathcal{H}_2(\mathbb{C}^+)$, where $\hat{\mathbf{H}}^s \in \mathcal{H}_2(\mathbb{C}^+)$ and $\hat{\mathbf{H}}^a \in \mathcal{H}_2(\mathbb{C}^-)$. Then, if:*

$$\|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})} < \|\hat{\mathbf{H}}^a\|_{\mathcal{H}_2(\mathbb{C}^-)} \tag{8.9}$$

then $\mathbf{G}_d$ is unstable.

[6] in the sense of the $\mathcal{L}_2(i\mathbb{R})$ norm

Proof. We prove the contrapositive. Let us suppose that $\mathbf{G}_d$ is stable, *i.e.*, $\mathbf{G}_d \in \mathcal{H}_2(\mathbb{C}^+)$. Then,

$$\begin{array}{ccc} \|\mathbf{G}_d - \hat{\mathbf{H}}\|^2_{\mathcal{L}_2(i\mathbb{R})} & = & \|\mathbf{G}_d - \hat{\mathbf{H}}^s - \hat{\mathbf{H}}^a\|^2_{\mathcal{L}_2(i\mathbb{R})} \\ & \underbrace{=}_{\text{pythagorean property}} & \|\mathbf{G}_d - \hat{\mathbf{H}}^s\|^2_{\mathcal{H}_2(\mathbb{C}^+)} + \|\hat{\mathbf{H}}^a\|^2_{\mathcal{H}_2(\mathbb{C}^-)} \\ & \geq & \|\hat{\mathbf{H}}^a\|^2_{\mathcal{H}_2(\mathbb{C}^-)}. \end{array}$$

Hence, if $\mathbf{G}_d$ is stable, then $\|\mathbf{G}_d - \hat{\mathbf{H}}\|^2_{\mathcal{L}_2(i\mathbb{R})} \geq \|\hat{\mathbf{H}}^a\|^2_{\mathcal{H}_2(\mathbb{C}^-)}$. Therefore , a sufficient condition to $\mathbf{G}_d$ to be unstable is $\|\mathbf{G}_d - \hat{\mathbf{H}}\|^2_{\mathcal{L}_2(i\mathbb{R})} < \|\hat{\mathbf{H}}^a\|^2_{\mathcal{H}_2(\mathbb{C}^-)}$. □

Theorem 8.6 provides a sufficient condition to check if $\mathbf{G}_d$ is unstable. It is based on the following geometric idea expressed by Figure 8.1. We recall that proposition 8.5 states that unstable systems form an open set. In other words, if $\hat{\mathbf{H}}$ is unstable, there exists a radius $R > 0$ such that the open ball

$$B_R(\hat{\mathbf{H}}) = \{\mathbf{G} \in \mathcal{L}_2(i\mathbb{R}) | \|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})} < R\},$$

contains only unstable systems. Theorem 8.6 enables to compute this radius R of instability. Therefore, if $\hat{\mathbf{H}}$ is unstable, then every system $\mathbf{G} \in \mathcal{L}_2(i\mathbb{R})$ satisfying

$$\|\boldsymbol{\mathcal{E}}\|_{\mathcal{L}_2(i\mathbb{R})} = \|\mathbf{G} - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})} < \|\hat{\mathbf{H}}^a\|_{\mathcal{H}_2(\mathbb{C}^-)} = R$$

will be also unstable. This result is geometrically sketched in Figure 8.1.

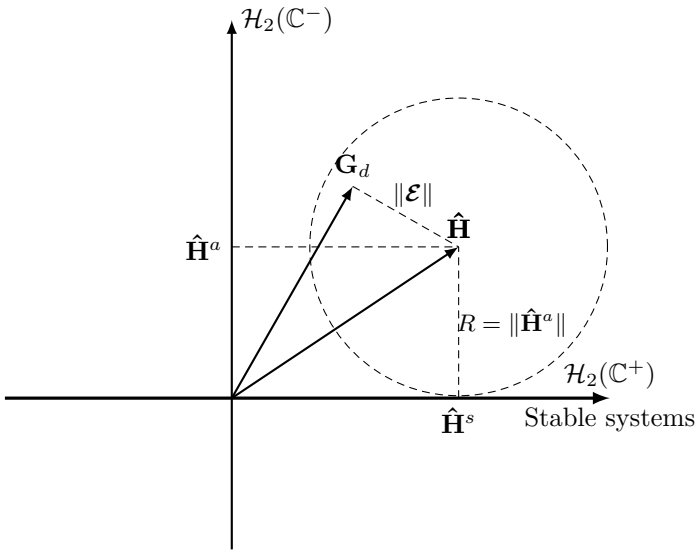

Figure 8.1: Intuition behind Theorem 8.6.

Indeed, the smallest perturbation leading $\hat{\mathbf{H}}$ to be stable is $-\hat{\mathbf{H}}^a$. Reader should notice that, since the set of stable systems is not open, we are not able to derive an equivalent theorem for stable systems.

Let us summarize what was presented so far in this section. Assume that, for a given $\mathbf{G}_d$ LTI TDS, one has obtained an approximation $\hat{\mathbf{H}}$ in the sense of the $\mathcal{L}_2(i\mathbb{R})$ norm. Then, two scenarios can occur :

(i) If $\hat{\mathbf{H}}$ is stable, *a priori* nothing can be said about the stability of $\mathbf{G}_d$ (see Proposition 8.4).

(ii) If $\hat{\mathbf{H}}$ is unstable, and if it is a good enough $\mathcal{L}_2(i\mathbb{R})$ approximation of $\mathbf{G}_d$, then $\mathbf{G}_d$ is also unstable (see Proposition 8.5).

Moreover, Theorem 8.6 provides a sufficient condition for $\mathbf{G}_d$ to be unstable. Hence, if one obtains an unstable approximation $\hat{\mathbf{H}} = \hat{\mathbf{H}}^s + \hat{\mathbf{H}}^a$, where $\hat{\mathbf{H}}^s \in \mathcal{H}_2(\mathbb{C}^+)$ and $\hat{\mathbf{H}}^a \in \mathcal{H}_2(\mathbb{C}^-)$ and

$$\|\mathbf{G}_d - \hat{\mathbf{H}}\|^2_{\mathcal{L}_2(i\mathbb{R})} = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\mathbf{G}_d(i\omega) - \hat{\mathbf{H}}(i\omega)|^2 d\omega < \|\hat{\mathbf{H}}^a\|^2_{\mathcal{H}_2(\mathbb{C}^-)}, \tag{8.10}$$

then $\mathbf{G}_d$ is also unstable. Furthermore, if $\hat{\mathbf{H}}$ is a finite dimensional system, so does $\hat{\mathbf{H}}^a \in \mathcal{H}_2(\mathbb{C}^-)$. Since $\hat{\mathbf{H}}^a(-s)$ is an element of $\mathcal{H}_2(\mathbb{C}^+) = \mathcal{H}_2$, the norm $\|\hat{\mathbf{H}}^a\|_{\mathcal{L}_2(i\mathbb{R})} = \|\hat{\mathbf{H}}^a(-s)\|_{\mathcal{H}_2}$ can be either computed using grammians or pole/residue decomposition (see Chapters 2 and 4).

8.2.3 Model approximation based method for evaluating systems' instability

Theorem 8.6 leads to a procedure to detect instability of a given LTI TDS. It is described as follows :

Method 8.7 (Model approximation based method to detect instability). *Given any LTI TDS $\mathbf{G}_d$, then the following steps allow to determine if it is unstable:*

1. *Guess a finite dimension unstable approximation $\hat{\mathbf{H}} = \hat{\mathbf{H}}^s + \hat{\mathbf{H}}^a$, where $\hat{\mathbf{H}}^s \in \mathcal{H}_2(\mathbb{C}^+)$ and $\hat{\mathbf{H}}^a \in \mathcal{H}_2(\mathbb{C}^-)$.*
2. *Verify, by numerical integration, if the integral inequality (8.9) is satisfied.*

Method 8.7 provides a way to asses that a given LTI TDS is unstable. Likewise the Lyapunov methods, where one needs to guess a Lyapunov function to prove stability, here, one needs to guess an unstable approximation $\hat{\mathbf{H}}$ and check if it is closed enough to $\mathbf{G}_d$. In addition, Step 2 gives a guaranty certificate of instability by means of the estimation of the $\mathcal{L}_2(i\mathbb{R})$ norm of the approximation error $\boldsymbol{\mathcal{E}}(s) = \mathbf{G}_d(s) - \hat{\mathbf{H}}(s)$. This can be done by computing numerically the following integral :

$$\|\mathbf{G}_d(s) - \hat{\mathbf{H}}(s)\|^2_{\mathcal{L}_2(i\mathbb{R})} = \|\boldsymbol{\mathcal{E}}\|^2_{\mathcal{L}_2(i\mathbb{R})} = \int_{-\infty}^{\infty} |\boldsymbol{\mathcal{E}}(i\omega)|^2 d\omega.$$

Reader should note the numerical integration of $\|\boldsymbol{\mathcal{E}}\|_{\mathcal{L}_2(i\mathbb{R})}$ might be a difficult task if $|\boldsymbol{\mathcal{E}}(i\omega)|$ has an intense oscillatory behavior. In this case, the integration quadrature should be computed with a large number of sample points and this can be very time-demanding. In addition, a sufficient condition for inequality (8.9) be verified is to show that

$$|\boldsymbol{\mathcal{E}}(i\omega)| \leq |\hat{\mathbf{H}}^a(i\omega)|, \text{ for all } \omega \in \mathbb{R},$$

i.e., the Bode magnitude diagram of $\boldsymbol{\mathcal{E}}$ is bounded by the one from $\hat{\mathbf{H}}^a$. This method is exemplified by the following study-case:

Example 8.8 (Stability of a single state-delay TDS). *Let us consider* $\mathbf{G}_d$ *an LTI TDS whose transfer function is given by*

$$\mathbf{G}_d(s) = \frac{1}{s + 2e^{-s}}.$$

Let $\hat{\mathbf{H}}$ *be a finite dimension approximation*[a] *of* $\mathbf{G}_d$ *whose transfer is given by:*

$$\hat{\mathbf{H}}(s) = \frac{0.5616s + 1.244}{s^2 - 0.3456s + 2.831}.$$

The poles of $\hat{\mathbf{H}}$ *are*

$$\lambda_1 = 0.1728 + 1.6737i \quad \lambda_2 = 0.1728 - 1.6737i,$$

and $\hat{\mathbf{H}}$ *is obviously unstable. Notice that* $\hat{\mathbf{H}} \in \mathcal{H}_2(\mathbb{C}^-)$ *and hence,* $\hat{\mathbf{H}} = \hat{\mathbf{H}}^s + \hat{\mathbf{H}}^a$*, where*

$$\hat{\mathbf{H}}^s = 0 \in \mathcal{H}_2(\mathbb{C}^-) \quad \text{and} \quad \hat{\mathbf{H}}^a = \hat{\mathbf{H}} \in \mathcal{H}_2(\mathbb{C}^-)$$

Figure 8.2 provides the magnitude bode plot of $\mathbf{G}_d(s)$*,* $\hat{\mathbf{H}}(s)$ *and the mismatch approximation error* $\boldsymbol{\mathcal{E}}(s) = \mathbf{G}_d(s) - \hat{\mathbf{H}}(s)$*.*

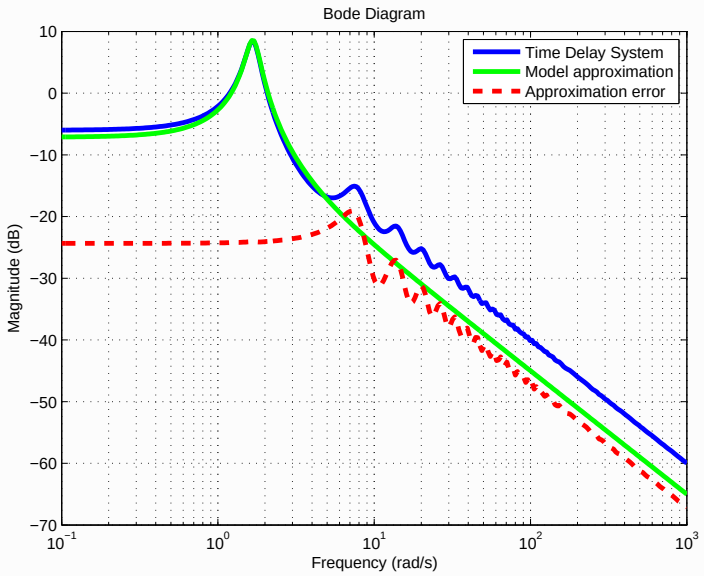

Figure 8.2: Bode magnitude diagram of $\mathbf{G}_d$ (blue solid line), unstable approximation $\hat{\mathbf{H}}$ (green solid line) and mismatch error $\boldsymbol{\mathcal{E}}$ (red dashed line).

$\hat{\mathbf{H}}(s)$ *seems to be a quite good approximation of* $\mathbf{G}_d(s)$ *in the sense of the* $\mathcal{L}_2(i\mathbb{R})$ *norm. Let us compute the following norm by gramians*

$$\|\hat{\mathbf{H}}^a\|_{\mathcal{H}_2(\mathbb{C}^-)} = \|\hat{\mathbf{H}}(-s)\|_{\mathcal{H}_2(\mathbb{C}^+)} \approx 1.1169.$$

Then, $\mathcal{L}_2(i\mathbb{R})$ norm of the mismatch error was computed by numerical integration :

$$\|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})} = \|\boldsymbol{\mathcal{E}}\|_{\mathcal{L}_2(i\mathbb{R})} = 0.1581.$$

Therefore,

$$\|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})} < \|\hat{\mathbf{H}}^a\|_{\mathcal{L}_2(i\mathbb{R})}$$

and, by Theorem 8.6, the system $\mathbf{G}_d$ is also unstable. To sum up,

(i) A LTI TDS $\mathbf{G}_d$ was given.

(ii) An unstable finite dimensional approximation $\hat{\mathbf{H}}$ was guessed.

(iii) $\hat{\mathbf{H}}$ is a good approximation of $\mathbf{G}_d$ in the sense of the $\mathcal{L}_2(i\mathbb{R})$ norm.

(iv) In addition, the inequality (8.9) was numerically verified, proving that $\mathbf{G}_d$ is also unstable.

[a] This approximation is a modal truncation of order 2 obtained via the Lambert function.

In this section, given $\mathbf{G}_d$, an LTI TDS, we derived an approach to asses its instability based on a good enough $\mathcal{L}_2(i\mathbb{R})$ approximation. Up to now, the bottleneck of Method 8.7 is to guess a good approximation of $\mathbf{G}_d$. In what follows, we will present a heuristic procedure to built such that approximation.

8.3 Model approximation for unstable systems by interpolation

8.3.1 Optimal $\mathcal{L}_2$ approximation problem

As it was stated before, given any LTI TDS $\mathbf{G}_d$, if one is able to construct $\hat{\mathbf{H}}$, a finite dimensional unstable system which is a good enough $\mathcal{L}_2(i\mathbb{R})$ approximation of $\mathbf{G}_d$, then $\mathbf{G}_d$ is also unstable. Hence, the aim of this section is to provide some heuristics to find such approximation. As in Chapter 4, where the $\mathcal{H}_2$ optimal model approximation was studied, one might consider here the $\mathcal{L}_2(i\mathbb{R})$ model approximation problem. This problem is formally stated as:

Problem 8.9 (Optimal $\mathcal{L}_2$ approximation problem). *Given any $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$ an LTI TDS and an order n, find an n-th finite dimensional LTI system $\hat{\mathbf{H}}^* \in \mathcal{L}_2(i\mathbb{R})$ such that*

$$\hat{\mathbf{H}}^* := \underset{\hat{\mathbf{H}} \in \mathcal{L}_2(i\mathbb{R}), \dim(\hat{\mathbf{H}}) \leq n}{\boldsymbol{arg\ min}} \|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})}. \tag{8.11}$$

Problem 8.9 is the natural extension of Problem 4.1 (see Chapter 4) to the case where unstable systems are considered. Hence, following the steps from Chapter 4, one should derive the $\mathcal{L}_2(i\mathbb{R})$ optimality conditions of the problem 8.9. This was the object of [Magruder et al., 2010] and the conditions are presented as follows:

Proposition 8.10 ($\mathcal{L}_2(i\mathbb{R})$ necessary optimality conditions). *Given $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$ and its decomposition $\mathbf{G}_d = \mathbf{G}_d^s + \mathbf{G}_d^a$ where $\mathbf{G}_d^s \in \mathcal{H}(\mathbb{C}^+)$ and $\mathbf{G}_d^a \in \mathcal{H}(\mathbb{C}^-)$. Let $\hat{\mathbf{H}}$ be the local minimizer of order n of Problem 8.9, whose poles are all simple $\{\hat{\lambda}_1, \ldots, \hat{\lambda}_k\} \in \mathbb{C}^-$ and*

$\{\hat{\lambda}_{k+1}, \ldots, \hat{\lambda}_n\} \in \mathbb{C}^+$, and its decomposition is given by $\hat{\mathbf{H}} = \hat{\mathbf{H}}^s + \hat{\mathbf{H}}^a$ where $\hat{\mathbf{H}}^s \in \mathcal{H}(\mathbb{C}^+)$ and $\hat{\mathbf{H}}^a \in \mathcal{H}(\mathbb{C}^-)$. Then, the following Hermitien interpolation conditions are satisfied:

$$\begin{aligned} \mathbf{G}_d^s(-\hat{\lambda}_i) &= \hat{\mathbf{H}}^s(-\hat{\lambda}_i), \quad \left.\frac{d\mathbf{G}_d^s}{ds}\right|_{s=-\hat{\lambda}_i} = \left.\frac{d\hat{\mathbf{H}}^s}{ds}\right|_{s=-\hat{\lambda}_i} \quad \text{for } i = 1, \ldots, k \\ &\text{and} \\ \mathbf{G}_d^a(-\hat{\lambda}_i) &= \hat{\mathbf{H}}^a(-\hat{\lambda}_i), \quad \left.\frac{d\mathbf{G}_d^a}{ds}\right|_{s=-\hat{\lambda}_i} = \left.\frac{d\hat{\mathbf{H}}^a}{ds}\right|_{s=-\hat{\lambda}_i} \quad \text{for } i = k+1, \ldots, n \end{aligned} \quad . \tag{8.12}$$

Proof. See [Magruder et al., 2010]. □

Proposition 8.10 states the $\mathcal{L}_2(i\mathbb{R})$ necessary optimality conditions related to Problem 8.9. As in the $\mathcal{H}_2$ case (see Chapter 4), they are Hermite interpolation conditions. However, in the $\mathcal{L}_2(i\mathbb{R})$ case, $\mathbf{G}_d^s$ (stable part of $\mathbf{G}_d$) interpolates $\hat{\mathbf{H}}^s$ (stable part of $\hat{\mathbf{H}}$) and $\mathbf{G}_d^a$ (unstable part of $\mathbf{G}_d$) interpolates $\hat{\mathbf{H}}^a$ (unstable part of $\hat{\mathbf{H}}$). Hence, in order to use the $\mathcal{L}_2(i\mathbb{R})$ optimality conditions for building a reduced order model, one needs to know *a priori* the decomposition of $\mathbf{G}_d = \mathbf{G}_d^s + \mathbf{G}_d^a$, where $\mathbf{G}_d^s \in \mathcal{H}(\mathbb{C}^+)$ and $\mathbf{G}_d^a \in \mathcal{H}(\mathbb{C}^-)$. Hence, in this case, we already know if system $\mathbf{G}_d$ is stable (just verify if $\mathbf{G}_d^a = 0$). Therefore, these conditions cannot be used to construct the approximation required in Method 8.7 and a further heuristic need to be developed.

8.3.2 Interpolatory based heuristic for $\mathcal{L}_2(i\mathbb{R})$ model approximation

The main heuristic we will employ to construct a good enough $\mathcal{L}_2(i\mathbb{R})$ approximation is inspired by the $\mathcal{H}_2$ optimality conditions (see Chapter 4, Section 4.3). Hence, given $\mathbf{G}_d$, an LTI TDS, we will construct an approximation satisfying the following Hermite interpolation conditions:

Problem 8.11 ($\mathcal{L}_2(i\mathbb{R})$ **interpolation problem).** *Given $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$ an LTI TDS and an order n, construct a finite dimensional approximation $\hat{\mathbf{H}} \in \mathcal{L}_2(i\mathbb{R})$ satisfying*

$$\mathbf{G}_d(-\hat{\lambda}_k) = \hat{\mathbf{H}}(-\hat{\lambda}_k), \quad \text{and} \quad \mathbf{G}_d'(-\hat{\lambda}_k) = \hat{\mathbf{H}}'(-\hat{\lambda}_k) \tag{8.13}$$

where $\{\hat{\lambda}_1, \ldots, \hat{\lambda}_n\} \in \mathbb{C}$ are the poles of the approximation $\hat{\mathbf{H}}$.

As exposed in Chapter 4, a reduced-order model can be constructed using interpolation-based iterative methods (*e.g.* **IRKA** or **TF-IRKA**). In our case, **TF-IRKA** is the most adapted algorithm, because $\mathbf{G}_d$ is an LTI TDS and, hence, is represented by a irrational transfer function.

Reader should note that conditions (8.13) correspond to the optimality conditions of the $\mathcal{H}_2$ model approximation problem (see Chapter 4, Corollary 4.15):

$$\text{Find } \hat{\mathbf{H}}^* \in \mathcal{H}_2 \text{ s.t. } \hat{\mathbf{H}}^* := \underset{\hat{\mathbf{H}} \in \mathcal{H}_2 \, \dim(\hat{\mathbf{H}}) \leq n}{\textbf{arg min}} \|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{H}_2}. \tag{8.14}$$

Hence, if $\mathbf{G}_d$ is stable and we construct an approximation $\hat{\mathbf{H}}$ satisfying the conditions (8.13), then it might be a local optimum. In additional, numerical experiences shows that :

(i) For orders n not too high, **TF-IRKA** generally converges, even if the system $\mathbf{G}_d$ is unstable. Indeed, the following monograph [Sinani, 2015] shows numerically that **TF-IRKA** converges for a great amount of unstable systems.

(ii) If the original system $\mathbf{G}_d$ is stable and if the algorithm **TF-IRKA** converges, the reduced order model $\hat{\mathbf{H}}$ is generally also stable.

Henceforth, we apply **TF-IRKA** in order to find the approximation needed in the Method 8.7. Let us apply this procedure in the following example :

Example 8.12 (**Stability of a single state-delay TDS).** *Let us consider* $\mathbf{G}_d$ *an LTI TDS whose transfer function is given by*

$$\mathbf{G}_d(s) = \frac{1}{s^2 + 0.1s + e^{-3s}s + e^{-2s}}.$$

Let $\hat{\mathbf{H}}$ *be a finite dimension approximation of order* $r = 8$ *of* $\mathbf{G}_d$ *obtained by* ***TF-IRKA****. Model* $\hat{\mathbf{H}}$ *is unstable and Figure 8.3 provides the magnitude Bode plot of* $\mathbf{G}_d(s)$, $\hat{\mathbf{H}}(s)$, $\hat{\mathbf{H}}^a(s)$ *the unstable part of* $\hat{\mathbf{H}}(s)$*, and the approximation error* $\boldsymbol{\mathcal{E}}(s) = \mathbf{G}_d(s) - \hat{\mathbf{H}}(s)$.

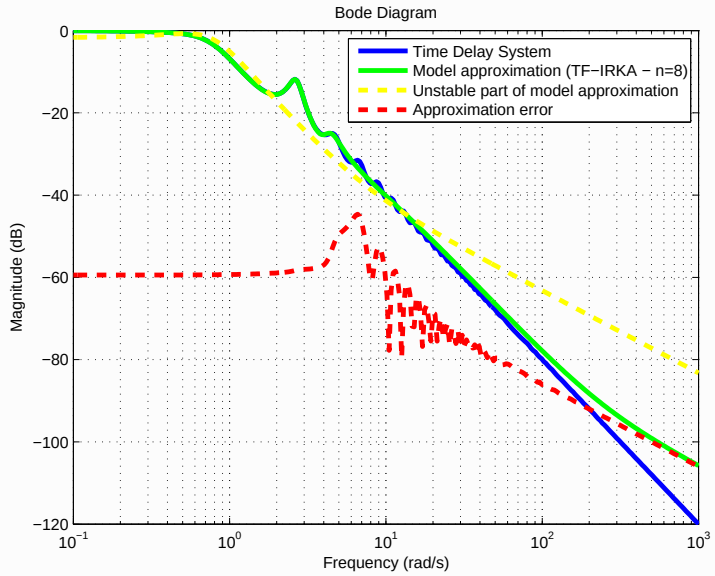

Figure 8.3: Bode magnitude diagram of $\mathbf{G}_d$ (blue solid line), unstable approximation $\hat{\mathbf{H}}$ (green solid line), unstable part of the approximation $\hat{\mathbf{H}}^a$ (green solid line) and mismatch error $\boldsymbol{\mathcal{E}}$ (red dashed line).

$\hat{\mathbf{H}}(s)$ *seems to be a good approximation of* $\mathbf{G}_d(s)$ *in the sense of the* $\mathcal{L}_2(i\mathbb{R})$ *norm. In addition, let us compute the following norm*

$$\|\hat{\mathbf{H}}^a\|_{\mathcal{L}_2(i\mathbb{R})} \approx 0.5032,$$

by grammians. Moreover, the $\mathcal{L}_2(i\mathbb{R})$ *norm of the mismatch error was computed by numerical integration :*

$$\|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})} = \|\boldsymbol{\mathcal{E}}\|_{\mathcal{L}_2(i\mathbb{R})} = 0.0038.$$

Obviously, the magnitude of $|\hat{\mathbf{H}}^a(i\omega)|$ *majored the error* $|\mathcal{E}(i\omega)|$ *as shown numerically in Figure 8.3 (which is enough to prove instability). Therefore,*

$$\|\boldsymbol{\mathcal{E}}\|_{\mathcal{L}_2(i\mathbb{R})} = \|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})} < \|\hat{\mathbf{H}}^a\|_{\mathcal{L}_2(i\mathbb{R})}$$

and, by Theorem 8.6, the system $\mathbf{G}_d$ *is also unstable.*

To sum up, the algorithm **TF-IRKA** provides reduced order models that are good enough $\mathcal{L}_2(i\mathbb{R})$ approximation of the original. These approximations, allied with Theorem 8.6, can be used to prove if a given system $\mathbf{G}_d$ is unstable. Up to now, the choice of the order n of the approximation is still arbitrary and this will be discussed in the next chapter.

Based in what was presented up to now, in the next chapter, we will propose an interpolation-based procedure in order to estimate the stability charts of an LTI TDS using **TF-IRKA**.

Conclusion

In this chapter, some theoretical results have been exposed linking stability and model approximation. The main results exposed can be summarized as :

(1) A Hardy space characterization LTI TDS stability was developed in Proposition 8.3.

(2) Given $\mathbf{G}_d$ an LTI TDS and $\hat{\mathbf{H}}$ a good enough $\mathcal{L}_2(i\mathbb{R})$ unstable approximation, Theorem 8.6 exhibits a sufficient condition to be unstable. This condition involves $\mathcal{L}_2(i\mathbb{R})$ norm computation and can be evaluated by numerical integration.

(3) Finally, Section 8.3 shows that, given $\mathbf{G}_d$ an LTI TDS, one can use **TF-IRKA** to produce a good-enough approximation to $\mathbf{G}_d$. If this approximation is unstable and the condition (8.9) is verified, then $\mathbf{G}_d$ is also unstable. The choice of the approximation order n is not evident and will be discussed in the next chapter.

The results presented here enables to prove if a given LTI TDS is unstable. As stated before, one cannot provide similar results to prove if a system is stable. In the next chapter, we propose a weaker notion of stability in order to derive similar results. In addition , inspired on the results presented here, we will use **TF-IRKA** to estimate stability charts of any LTI TDS.

Chapter 9

Stability chart approximation from interpolatory methods

In this chapter, the ideas presented in Chapter 8 will be exploited in order to estimate the stability charts of any system $\mathbf{G}_d$. This chapter is divided in two sections. The first one, is dedicated to the introduction of the problem and to the development of some theoretical results, such as (i) a weaker notion of stability, enabling to derive a certificate of weaker stability result, and (ii) some properties of the algorithm **TF-IRKA**. The second section is dedicated to the construction of a procedure to estimate the stability charts of a given LTI TDS. It is exemplified with some numerical problems demonstrating that the proposed method provides a good estimation of the stability charts, even in a large scale setting.

Contents

9.1 Problem statement and some theoretical results

In the latter chapter, a model approximation based framework for evaluating the stability of **TDS** was presented along with a numerical procedure (Method 8.7). This enables to prove that a given LTI TDS $\mathbf{G}_d$ is unstable. Based on this idea, this chapter aims at providing a general methodology to estimate the stability charts of any LTI TDS using model approximation methods. The chapter is organized as follows :

(i) In this section, the stability chart problem of an LTI TDS is defined in Problem 9.1. In addition, a weaker notion of stability is also defined, enabling to derive a guaranty certificate result.

(ii) In Section 9.2, we propose a brute-force procedure in order to estimate the stability charts of any LTI TDS. In addition, we take benefit from the continuity property of the shifts (see Remark 9.10) in order to improve the convergence speed of the algorithm. Then, the procedure is validated in a large scale example.

(iii) Finally, in Section 9.3 a procedure enabling to vary the order from the ROM is introduced and validated in some numerical examples.

9.1.1 Problem statement

This chapter aims at using interpolation-based model approximation to tackle the following problem :

Problem 9.1 (Stability charts of an LTI TDS). *Let $\mathbf{G}_d(\mathbf{p})$ be an LTI TDS depending continuously on the vector of parameters $\mathbf{p} \in \mathcal{D}$[a]. Determine, for each $\mathbf{p}^* \in \mathcal{D}$ fixed, the stability of system $\mathbf{G}_d(\mathbf{p}^*)$.*

[a] The parameters considered are delays, gains, ...

This problem can be exemplified as follows.

Example 9.2 (Stability charts example). *Let us consider the second order open-loop system $\mathbf{G}(s) = \frac{1}{s^2+9}$ in feedback with the delayed controller $\mathbf{C}(s) = ke^{-\tau s}$. This closed-loop system, denoted as $\mathbf{G}_d(k, \tau)$, was considered in [Sipahi et al., 2011] and [Abdallah et al., 1993], and its transfer function is given by*

$$\mathbf{G}_d(s, \tau, k) = \frac{ke^{-\tau s}}{s^2 + 9 - ke^{-\tau s}}. \tag{9.1}$$

The problem we are interested in is to derive the stability chart *as a function of the parameters $\mathbf{p} = (\tau, k)$, couple of delay and gain values i.e. the conditions for $\tau \in [0, 10]s$ and $k \in [0, 4]s$ for which $\mathbf{G}_d(\tau, k)$ is stable or not. This specific problem has been theoretically solved in [Abdallah et al., 1993] using a deep analysis of its characteristic equation leading to the exact stability chart of $\mathbf{G}_d$ as a function of the (τ, k) couple. As a first illustration of the proposed approach, Figure 9.1 illustrates the stability regions based on the evaluation of the stability of approximated models $\hat{\mathbf{H}}$, obtained by interpolatory methods over a finite grid set of frozen (τ, k) values (here 100×100 grid points), using the procedure detailed in the following sections.*

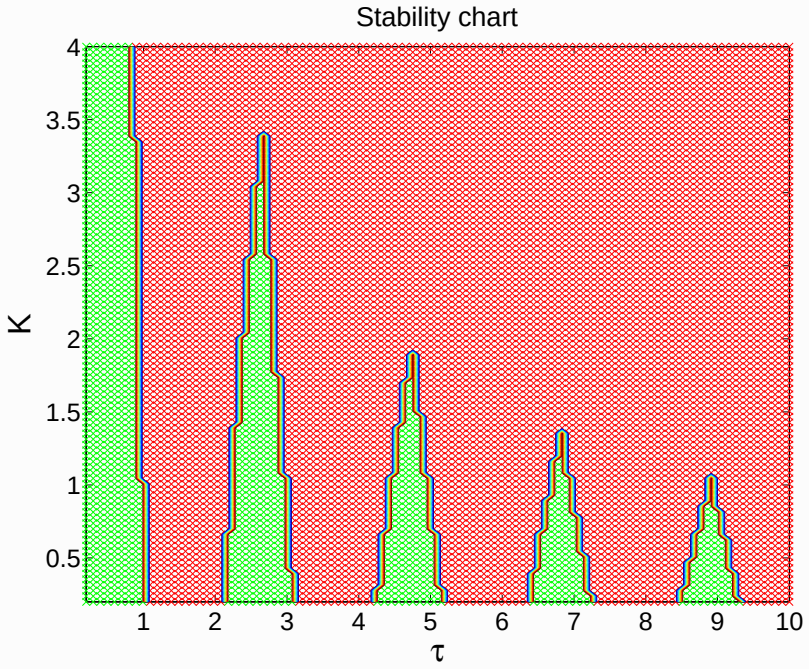

Figure 9.1: Example 1 - $\mathbf{G}_d$ is a **TDS** of order 2. The red points correspond to the unstable cases, while the green points correspond to stable systems.

The red and green regions correspond to the parameters (τ, k) for which the system $\mathbf{G}_d$ is unstable and stable respectively. The data corresponding to Figure 9.1 were obtained using the Method 9.12 in approximatively 190 seconds. By comparing with [Sipahi et al., 2011] results, it is clear that the proposed algorithm well catches the (in)stability property while being quite fast.

Before detailing the proposed contribution, we stress out that the main objective of the proposed methodology is not to provide an exact solution to the stability chart problem (as it is well solved by dedicated techniques) but rather a procedure to obtain an approximation of it in shorter time. As more clearly stated in what follows, we are interested in providing an algorithm to approximate (quite accurately, but still approximate) the stability regions of any LTI TDS of the form (8.1) or (8.2). The ground motivation for such *approximation result* is that most of the exact techniques which are based on algebraic or optimization techniques (such as bifurcation [Engelborghs et al., 2001], linear matrix inequalities [Briat, 2015; Seuret and Gouaisbaut, 2015]) reveal a computational weakness when the complexity of the aforementioned LTI TDS increases (*i.e.* its state dimension n or number of delays n_d). As a matter of consequence, in practice, these methods became untractable for (very) large-scale systems.

9.1.2 Preliminary results

As it was exposed in chapter 8, one way to prove that $\mathbf{G}_d$ is unstable, is to show an unstable model approximation $\hat{\mathbf{H}}$ which is close enough to $\mathbf{G}_d$. In addition, Theorem 8.6 provides a way to certify the instability of $\mathbf{G}_d$. Since the set of stable systems $\mathcal{H}_2$ is not open, we are not able to derive similar results. In what follows, we consider a weaker notion of stability for which we

are able to derive a similar certificate theorem as Theorem 8.6.

Weaker notion of stability

As previously stated, the set of stable systems is not an open set[1] and, as a consequence, no certificate of stability guaranty can be derived. One way of dealing with that is to consider a weaker stability notion, allowing the systems to be "just a little bit" unstable. This can be mathematically described by the following definition.

Definition 9.3 (α**-stable system).** *Let $\mathbf{G}_d$ be an LTI TDS and $\alpha > 0$ be a small positive real number. Then, if $\mathbf{G}_d = \mathbf{G}_d^s + \mathbf{G}_d^a$, where $\mathbf{G}_d^s \in \mathcal{H}_2(\mathbb{C}^+)$ and $\mathbf{G}_d^a \in \mathcal{H}_2(\mathbb{C}^-)$, we say that the system $\mathbf{G}_d$ is* α-stable *if and only if*

$$\|\mathbf{G}_d^a\|_{\mathcal{H}_2(\mathbb{C}^-)} < \alpha.$$

Hence, given $\alpha > 0$, the set of α-stable systems consists of systems that are just a little bit unstable, *i.e.*, the $\mathcal{L}_2(i\mathbb{R})$ norm of their unstable part is bounded by α. Reader should note that the notion of α-stability employed here differs from the notion of α-exponentially stability[2]. Moreover, the set of α-stable systems is open as stated as follows :

Proposition 9.4 (**Topology of α-stable systems).** *Let α be a positive real number. Then, the set of α-stable systems,*

$$\alpha\mathcal{S} = \{\mathbf{G} \in \mathcal{L}_2(i\mathbb{R}), \|\mathbf{G}^a\|_{\mathcal{H}_2(\mathbb{C}^-)} < \alpha\}$$

is an open set of $\mathcal{L}_2(i\mathbb{R})$.

Proof. The projector function :

$$\begin{array}{rcl} \text{proj}_{\mathcal{H}_2(\mathbb{C}^-)} : \mathcal{L}_2(i\mathbb{R}) & \to & \mathcal{H}_2(\mathbb{C}^-) \\ \mathbf{G} & \mapsto & \text{proj}_{\mathcal{H}_2(\mathbb{C}^-)}(\mathbf{G}) = \mathbf{G}^a \end{array},$$

and the $\mathcal{L}_2(i\mathbb{R})$ norm

$$\begin{array}{rcl} \|.\|_{\mathcal{L}_2(i\mathbb{R})} : \mathcal{L}_2(i\mathbb{R}) & \to & \mathbb{R} \\ \mathbf{G} & \mapsto & \|\mathbf{G}\|_{\mathcal{L}_2(i\mathbb{R})} \end{array},$$

are continuous functions (with respect to the $\mathcal{L}_2(i\mathbb{R})$ norm). Hence, their composition, $f = \|\text{proj}_{\mathcal{H}_2(\mathbb{C}^-)(.)}\|$ is also continuous. The set of α-stable systems is the inverse image of a continuous function of an open set,

$$\alpha\mathcal{S} = f^{-1}\big((-\infty, \alpha)\big),$$

and it is open as well. □

Moreover, a similar certificate result as Theorem 8.6 can be stated for α-stable systems as follows.

[1] with respect to the $\mathcal{L}_2(i\mathbb{R})$ norm.

[2] a continuous LTI system is said to be α-exponentially stable if and only if the system has eigenvalues with real parts strictly less than α.

Theorem 9.5 (Certificate α-stability). *Let $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$ be an LTI TDS and $\hat{\mathbf{H}}$ be a stable approximation of $\mathbf{G}_d$. Then, if*

$$\|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})} < \alpha, \tag{9.2}$$

then $\mathbf{G}_d$ is α-stable.

Proof. Since $\hat{\mathbf{H}}$ is stable, then $\hat{\mathbf{H}} \in \mathcal{H}_2(\mathbb{C}^+)$. Hence,

$$\|\mathbf{G}_d^a\|_{\mathcal{H}_2(\mathbb{C}^-)} \leq \|\mathbf{G}_d^a\|_{\mathcal{H}_2(\mathbb{C}^-)} + \|\mathbf{G}_d^s - \hat{\mathbf{H}}^s\|_{\mathcal{H}_2(\mathbb{C}^+)} = \|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2(i\mathbb{R})} < \alpha$$

which proves the Theorem. □

Hence, given a $\alpha > 0$, Theorem 9.5 provides a sufficient condition to verify if $\mathbf{G}_d$ is α-stable, enabling us to compute numerically an α-stability certificate using a similar procedure as in 8.7. Reader should note that, by means of Theorem 9.5, we are not able to provide a stability certificate, but an α-stability one. That is because the set of stable system is not open while the set of α-stable systems is. The following proposition states a connection between α-stability and stability notions.

Proposition 9.6. *Let $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$ be an LTI TDS. If $\mathbf{G}_d$ is an α-stable system, for every $\alpha > 0$, then is is stable.*

Proof. The result is straightforward. The norm of the unstable part of $\mathbf{G}_d$ should satisfy

$$\|\mathbf{G}_d^a\|_{\mathcal{L}_2(i\mathbb{R})} < \alpha, \text{ for every } \alpha > 0.$$

Hence, $\mathbf{G}_d^a = 0$. Therefore, $\mathbf{G}_d \in \mathcal{H}_2$ and thus stable. □

In conclusion, the notion of α-stability was created to remedy the fact the set of stable systems is not open. Indeed, the set of α-stable systems is open and congregates the systems $\mathbf{G}$ whose unstable $\mathbf{G}^a$ have small norm. Moreover, Theorem 9.5 provides a sufficient condition for a system $\mathbf{G}$ to be α-stable. In addition, the same procedure sketched in Chapter 8 (Method 8.7) can be employed in order to detect a system is α-stable.

In what follows we recall some properties of the algorithm **TF-IRKA** and we justify its use for the estimation of stability charts, even if the certificates are not computed.

TF-IRKA and stability

As stated before, given $\mathbf{G}_d \in \mathcal{L}_2(i\mathbb{R})$, the algorithm **TF-IRKA** will be used to construct a model approximation $\hat{\mathbf{H}}$. This algorithm was designed initially to find reduced order models for $\mathcal{H}_2$ systems only. In practice, even for unstable systems, this algorithm produces ROMs that are good approximations. In what follows, we provide some arguments justifying (but not giving mathematical proofs) why this algorithm can be used as well to estimate the stability of any LTI TDS, even if the certificates from Theorems 8.6 and 9.5 are not computed.

Firstly, we recall that **TF-IRKA** is an iterative algorithm. In [Krajewski et al., 1995], the authors show that saddle points and local maxima of the $\mathcal{H}_2$ minimization problem are know to be repellent. Hence, if $\mathbf{G}_d$ is stable and the algorithm converges, it should converge to a local minimizer. The following proposition shows that if $\mathbf{G}_d$ is stable, so does the $\mathcal{L}_2(i\mathbb{R})$ local minimizer (from Problem 8.9).

Proposition 9.7 (stable $\mathcal{L}_2(i\mathbb{R})$ local minimizers). *Given $\mathbf{G}_d \in \mathcal{H}_2$, if $\hat{\mathbf{H}} \in \mathcal{L}_2(i\mathbb{R})$ is a local minimizer of the $\mathcal{L}_2$ approximation Problem 8.9, then $\hat{\mathbf{H}} \in \mathcal{H}_2$.*

Proof. Let $\hat{\mathbf{H}} \in \mathcal{L}_2(i\mathbb{R})$ be the local minimizer of (8.9). Since $\mathbf{G}_d \in \mathcal{H}_2(\mathbb{C}^+)$, its unstable part is zero, *i.e.*, $\mathbf{G}_d^a = 0$. Hence, the approximation error is

$$\|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2}^2 = \|\mathbf{G}_d - \hat{\mathbf{H}}^s\|_{\mathcal{L}_2}^2 + \|\mathbf{0} - \hat{\mathbf{H}}^a\|_{\mathcal{L}_2}^2 \tag{9.3}$$

Thus, $\hat{\mathbf{H}}^a = 0$, otherwise $\hat{\mathbf{H}}$ is not a local minimizer. Indeed, $\tilde{\mathbf{H}}_\varepsilon = \hat{\mathbf{H}}^s + (1-\varepsilon)\hat{\mathbf{H}}^a$ is a system such that $\|\mathbf{G}_d - \tilde{\mathbf{H}}_\varepsilon\|_{\mathcal{L}_2}^2 < \|\mathbf{G}_d - \hat{\mathbf{H}}\|_{\mathcal{L}_2}^2$, for every $\varepsilon > 0$. □

In other words, Proposition 9.7 states that, given $\mathbf{G}_d$ a stable LTI TDS, then if one finds a $\mathcal{L}_2(i\mathbb{R})$ local minimizer $\hat{\mathbf{H}}$ of Problem 8.9, then $\hat{\mathbf{H}}$ is also stable. Hence, if $\mathbf{G}_d$ is stable and **TF-IRKA** converges to a $\mathcal{L}_2(i\mathbb{R})$ local minimizer, then it should be stable as well. Based on this result we propose the following procedure to obtain the stability charts of an LTI TDS.

Given $\mathbf{G}_d$ an LTI TDS and let $\hat{\mathbf{H}}$ be an approximation of order n obtained after the convergence of **TF-IRKA**. Let us list here the possible scenarios:

(i) Unstable approximation : if $\hat{\mathbf{H}}$ is unstable and is a good enough $\mathcal{L}_2(i\mathbb{R})$ approximation of $\mathbf{G}_d$, then $\mathbf{G}_d$ is also unstable. In addition, since the algorithm converges in general to local minima, it is very likely for $\mathbf{G}_d$ to be unstable. Moreover, a certificate of guaranty is given Theorem by 8.6.

(ii) Stable approximation : if $\hat{\mathbf{H}}$ is stable and is a good enough $\mathcal{L}_2(i\mathbb{R})$ approximation of $\mathbf{G}_d$, then $\mathbf{G}_d$ might be also stable. In this case no guaranty of stability is given and we can only certify α stability using Theorem 9.5. However, it might happen to find a stable approximation $\hat{\mathbf{H}}$ even if the original system $\mathbf{G}_d$ is unstable. In this case, the approximation is probably not good enough to capture the unstable behavior of $\mathbf{G}_d$ and one should increase the order n of the approximation in order to capture this unstable behavior.

To sum up, given a system $\mathbf{G}_d$, if we *a priori* know an order n, such that the n-th order approximation $\hat{\mathbf{H}}$ obtained by **TF-IRKA** is very closed to $\mathbf{G}_d$, it is very likely that $\mathbf{G}_d$ and $\hat{\mathbf{H}}$ are either both stable or both unstable. In what follows we will use this idea to construct an algorithm to estimate the stability charts of an LTI TDS.

9.2 Procedure for stability estimation of an LTI TDS

9.2.1 Brute-force numerical procedure

Let $\mathbf{G}_d(\mathbf{p})$ be an LTI TDS as in (8.1) depending on the parameters values (delays, gains, etc.) $\mathbf{p} \in \mathcal{D}$, whose variations impact the system stability. The procedure to find the stability chart related to $\mathbf{G}_d(\mathbf{p})$ with respect of the parameters $\mathbf{p} \in \mathcal{D}$ is described in what follows

Method 9.8 (Brutal force procedure). *Let $\mathbf{G}_d(\mathbf{p})$ be an LTI TDS depending continuously on the parameters values $\mathbf{p} \in \mathcal{D}$. Follow the steps :*

(1) Discretize the parameter domain $\mathcal{D}$ into $\hat{\mathcal{D}}$.

(2) Choose a reduced order n for which reduced order models of order n might be a good approximation of the systems $\mathbf{G}_d(s, \mathbf{p}^i)$ using **TF-IRKA**.

(3) For each $\mathbf{p}^i \in \hat{\mathcal{D}}$, find a finite reduced order model $\hat{\mathbf{H}}$ and evaluate the stability of $\hat{\mathbf{H}}$.

(4) Use the stability of $\hat{\mathbf{H}}$ to estimate that of $\mathbf{G}_d(s, \mathbf{p}^i)$.

(5) (Optional) Use Theorems 8.6 and 9.5 to provide a guaranty certificate of instability and α-stability.

Let us consider the following example illustrates the proposed brute-force procedure:

Example 9.9 (**Example with certificates**). *Let us consider the system*

$$\mathbf{G}_d(\tau, \gamma) = \left\{ \begin{array}{rcl} \dot{\mathbf{x}}(t) & = & -\mathbf{x}(t-\tau) - \mathbf{x}(t-\gamma) + u(t) \\ y(t) & = & \mathbf{x}(t) \end{array} \right. , \tag{9.4}$$

with transfer function

$$\mathbf{G}_d(s, \tau, \gamma) = \frac{1}{s + e^{-\tau s} + 2e^{-\gamma s}}, \tag{9.5}$$

*where parameters values are the delay pair $\mathbf{p} = (\tau, \gamma) \in [0, 3] \times [0, 3] = \mathcal{D}$. We grid the domain $\mathcal{D}$ using a 40×40 uniformly spaced points to construct the discrete domain $\hat{\mathcal{D}}$. For each point in this grid, we compute a reduced order model of order $n = 2$ using **TF-IRKA**. The stability region evaluated using the brute-force procedure is plotted in Figure 9.2.*

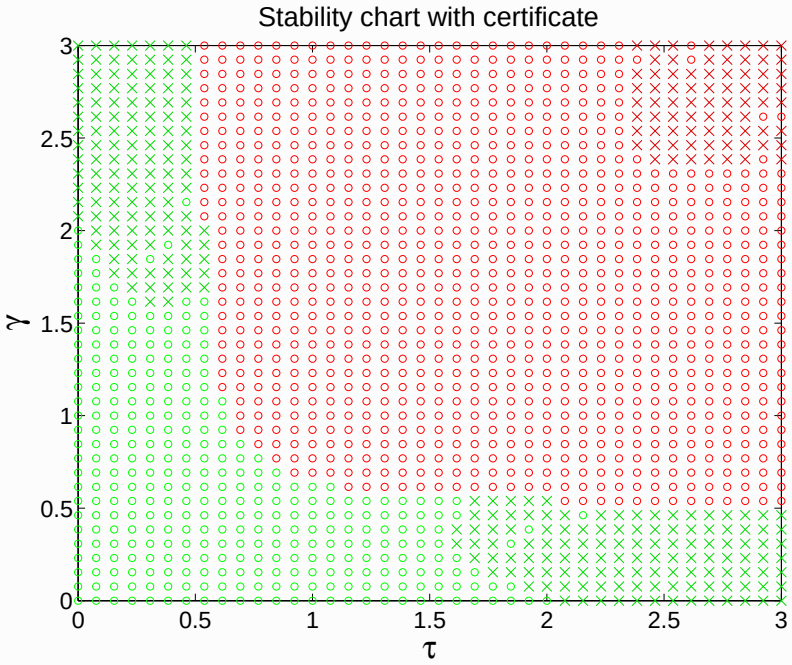

Figure 9.2: Example 2: System governed by (9.4). Red circles (approximation unstable with certificate), red crosses (approximation unstable but no certificate), green circles (approximation stable with certificate of 0.2 stability) and green crosses (approximation stable but no certificate of 0.2 stability). All points interpolated with a model of order $r = 2$ using **TF-IRKA**.

On Figure 9.2, the following holds :

- ***Red circles :*** *correspond to the case where the approximation obtained by* ***TF-IRKA*** *is unstable and a guaranty of instability of the original model was produced using Theorem 8.6.*
- ***Red crosses :*** *correspond to the case where the approximation obtained by* ***TF-IRKA*** *is unstable but we were not able to provide a certificate of instability. Hence, in this case, in order to obtain a guaranty certificate, it is necessary to obtained a better approximation by increasing the order of the reduced order model. This can be done by increasing the order of the reduced order model. Indeed, if $r = 4$, the approximations obtained by* ***TF-IRKA*** *are good enough to produce the guaranty certificate.*
- ***Green circles :*** *correspond to the case where the approximation obtained by* ***TF-IRKA*** *is stable and a guaranty of α-stability of the original model was produced using Theorem 9.5, where $\alpha = 0.2$.*
- ***Green crosses :*** *correspond to the case where the approximation obtained by* ***TF-IRKA*** *is stable but, for $\alpha = 0.2$, we were not able to provide a certificate of α-stability. This can be done by increasing the order of the reduced order model. Indeed, if $r = 4$, the approximations obtained by* ***TF-IRKA*** *are good enough to produce the guaranty certificate.*

In Example 9.9, we were able to prove numerically that the brute-force approach sketches gives a good estimation of the stability charts of the proposed LTI TDS. The following considerations should be notice :

- For a great majority of points from $\mathcal{D}$ we were able to certify instability and α-stability, for $\alpha = 0.2$, (corresponding to the red and green circles in Figure 9.2). However, for some points (the green and red crosses), the approximation obtained was not faithful enough to have those certifications. In those cases, one needs to increase the order of the approximation. In Section 9.3 we propose a procedure to adapt the order n of this approximation.
- The bottle-neck of this method is the computation of the guaranty certificates. Indeed, to produce these guaranty certificates, one needs, for each system $\mathbf{G}_d(s, \mathbf{p}^i)$, $\mathbf{p}^i \in \hat{\mathcal{D}}$, to evaluate the $\mathcal{L}_2(i\mathbb{R})$ norm of the approximation error, *i.e.*,

$$\|\mathbf{G}_d(s, \mathbf{p}^i) - \hat{\mathbf{H}}\|^2_{\mathcal{L}_2(i\mathbb{R})} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\mathbf{G}_d(i\omega, \mathbf{p}^i) - \hat{\mathbf{H}}(i\omega)|^2 d\omega$$

  by numerical integration, which is a very expensive computation, specially for large-scale systems.

Since the certificate computation is expensive, from now on we will not compute them anymore. Instead, we suppose that we know the order r for which the approximation $\hat{\mathbf{H}}$ is closed enough to $\mathbf{G}_d$ and we will use the stability of $\hat{\mathbf{H}}$ to estimate that of $\mathbf{G}_d$.

9.2.2 Continuity of interpolation points

The following remark is very important for a fast implementation of the procedure :

Remark 9.10 (**Continuity of interpolation points for small perturbation**). *Let* $\mathbf{G}_d(\mathbf{p}^*) \in \mathcal{L}_2(i\mathbb{R})$ *be an LTI TDS and the perturbed system* $\mathbf{G}(\mathbf{p}^* + \boldsymbol{\varepsilon}) \in \mathcal{L}_2(i\mathbb{R})$, *where* $\|\boldsymbol{\varepsilon}\| \ll 1$ *is a small perturbation.*

- ➤ *If the* ***TF-IRKA*** *applied to* $\mathbf{G}_d(\mathbf{p}^*) = \mathbf{G}_d$ *converges, it produces an approximation* $\hat{\mathbf{H}}$ *of order* r, *satisfying the Hermite interpolation conditions:*

$$\mathbf{G}_d(-\hat{\lambda}_k) = \hat{\mathbf{H}}(-\hat{\lambda}_k), \quad \text{and} \quad \mathbf{G}_d'(-\hat{\lambda}_k) = \hat{\mathbf{H}}'(-\hat{\lambda}_k), \ \text{for } k = 1, \ldots, r,$$

  where $\{\hat{\lambda}_1, \ldots, \hat{\lambda}_r\} \in \mathbb{C}$ *are the poles of* $\hat{\mathbf{H}}$.

- ➤ *Then, if* ***TF-IRKA*** *is applied to the perturbed system* $\mathbf{G}_d(\mathbf{p}^* + \boldsymbol{\varepsilon}) \in \mathcal{L}_2(i\mathbb{R})$ *initialized by using minus the poles of* $\hat{\mathbf{H}}$, i.e.,

$$\sigma_1, \ldots, \sigma_r = \{-\hat{\lambda}_1, \ldots, -\hat{\lambda}_r\},$$

  the convergence is very fast and the poles of the approximation obtained $\hat{\mathbf{H}}_{\text{perturbated}}$ *are a small perturbation of the poles of* $\hat{\mathbf{H}}$.

Therefore, the poles of the approximation obtained by ***TF-IRKA*** *follow somewhere a continuity property with respect to small perturbations on the original model* $\mathbf{G}_d$.

Remark 9.10 can be taken in advantage to increase the convergence speed of the proposed method. Hence, if $\mathbf{p}^i, \mathbf{p}^i + \boldsymbol{\varepsilon} \in \hat{\mathcal{D}}$, where $\boldsymbol{\varepsilon} \ll 1$, are parameter vectors, one should :

(i) Compute the approximation $\hat{\mathbf{H}}$ for $\mathbf{G}_d(\mathbf{p}^i)$ with **TF-IRKA**.

(ii) Initialize the algorithm with the optimal shifts points $(\hat{\sigma}_1, \ldots, \hat{\sigma}_n)$[3] when computing the approximation of $\mathbf{G}_d(\mathbf{p}^i + \boldsymbol{\varepsilon})$, so the convergence is faster.

This is illustrated in the following example

Example 9.11 (**Continuity of shifts example**). *Let us consider* $\mathbf{G}_d(s, \tau, \gamma)$ *the transfer function*

$$\mathbf{G}_d(s, \tau_1, \tau_2) = \frac{1}{s + e^{-s\tau} + e^{-s\gamma}}$$

where $\mathbf{p} = (\tau, \gamma)$. *If initially we consider* $\tau = 0.3$ *and* $\gamma = 0.5$, *the* 4*-th order approximation* $\hat{\mathbf{H}}_1$ *obtained by* ***TF-IRKA***, *which converges in* **10** *iterations. In addition, the poles of* $\hat{\mathbf{H}}_1$ *are*

$$\hat{\lambda}_{1,2} = -7.6856 \pm 8.1767i \quad \hat{\lambda}_{3,4} = -1.1859 \pm 2.8611i.$$

Now, let we consider $\tau = 0.31$ *and* $\gamma = 0.51$, *corresponding to a small perturbation on the parameters. Then the* 4*-th order approximation* $\hat{\mathbf{H}}_\Delta$ *obtained by* ***TF-IRKA*** *initialized with the shifts* $\{\sigma_1, \ldots, \sigma_4\} = \{-\hat{\lambda}_1, \ldots, -\hat{\lambda}_4\}$ *converges in* **2** *iterations. In addition, the*

[3] which are given by $(\hat{\sigma}_1, \ldots, \hat{\sigma}_n) = -\text{poles}(\hat{\mathbf{H}})$

following poles of $\hat{\mathbf{H}}_\Delta$ *are*

$$\tilde{\lambda}_{1,2} = -7.6274 \pm 8.2262i \quad \tilde{\lambda}_{3,4} = -1.1202 \pm 2.8380i.$$

Let us apply the brute-force procedure for estimating the stability charts of a machining chatter systems (see [Sipahi et al., 2011; Seuret and Gouaisbaut, 2015] for more details). This system is described by the following LTI TDS :

$$\mathbf{G}_d(k,\tau) = \left\{ \begin{array}{rcl} \dot{\mathbf{x}}(t) & = & A_0\mathbf{x}(t) + A_1\mathbf{x}(t-\tau) + \mathbf{b}u(t) \\ y(t) & = & \mathbf{c}^T\mathbf{x}(t) \end{array} \right. , \tag{9.6}$$

where,

$$A_0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -10-k & 10 & 0 & 0 \\ 5 & -15 & 0 & -0.25 \end{bmatrix}, \; A_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & k & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and

$$\mathbf{b} = [0\ 0\ 0\ 1]^T \quad \text{and} \quad \mathbf{c}^T = [1\ 0\ 0\ 0].$$

The problem we are interested in is to derive the *stability chart* as a function of the parameters $\mathbf{p} = (k, \tau)$, couple of gain and delay values, where $k \in [0, 10]$ and $\tau \in [0\ 6]$s. This specific problem has been theoretically solved in [Seuret and Gouaisbaut, 2015] using a LMI approach. Here, by following the brute-force procedure :

(i) We consider 100 equidistant grid points for the delay τ and 100 logarithmically spaced grid points for the controller gain k.

(ii) For every pair (k_i, τ_j) we compute an approximation of order $n = 6$ using **TF-IRKA**.

(iii) Every **TF-IRKA** was initialized with the poles of a neighbor system.

(iv) If the algorithm converges and the approximation is stable, we plot a green cross.

(v) If the algorithm converges and the approximation is unstable, we plot a red cross.

(vi) If the algorithm does not converge, nothing is plotted.

The results are exposed in the Figure 9.3. By comparing with the results in [Seuret and Gouaisbaut, 2015], it is clear that the proposed algorithm well catches the (in)stability property. In addition, for this specific system, the exact stability can be derived using the stability margins and the Nyquist method. The blue curve was obtained applying this method and it separates the regions where the system looses stability.

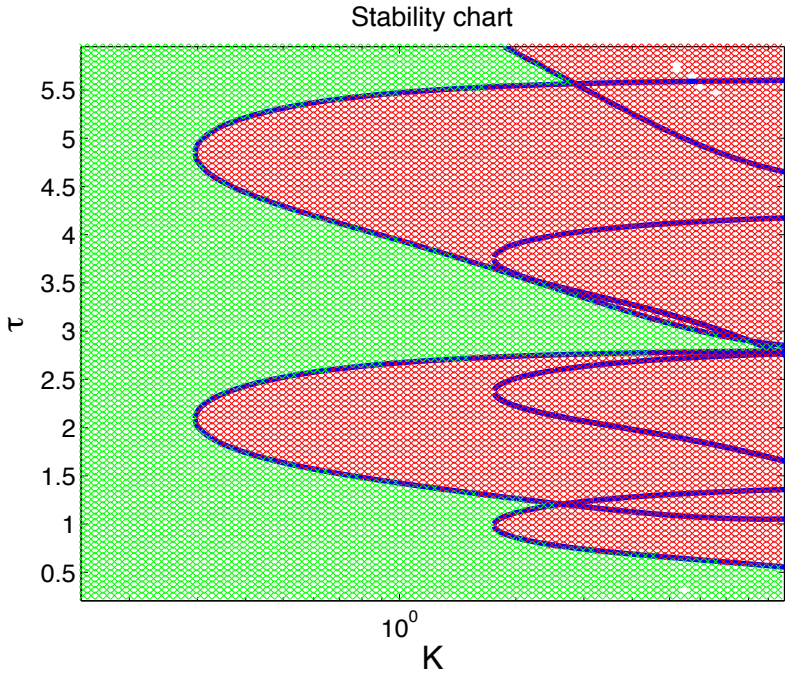

Figure 9.3: Stability charts estimation of $\mathbf{G}_d(k,\tau)$, machining chatter systems governed by (9.6) using the brutal-force method. For each point in the graph, a 6-th order approximation $\hat{\mathbf{H}}$ was obtained by **TF-IRKA**. Green crosses represent stable approximations. Red crosses represent unstable approximations. The blue line corresponds where obtained using the Nyquist method.

Let us list some important information about this simulation :

- The order $r = 6$ was chosen heuristically because it provides good approximation for a given (k,τ).
- 10000 approximations where computed in total.
- **TF-IRKA** needs on average 20 iterations to converge for this system for any given parameter (k_i,τ_j).
- However, the algorithm is initialized with minus the poles of a neighbor system, it converges on average in 2 or 3 iterations.
- The total time needed was 120 seconds. Hence, the mean time per model approximation is $\frac{130}{100^2} = 0.013$s. Reader should notice that this was only possible because of the smart initialization of **TF-IRKA** (based on the Remark 9.10).
- The white points correspond to the parameters where **TF-IRKA** did not converge. However, if one choose another order $r > 6$ it does converge and the approximation obtained has the coherent stability.

Notice that the order 6 was heuristically determined because 6-th approximation are closed enough to a given $\mathbf{G}_d(k_i,\tau_j)$. The same procedure was applied to approximations of orders 2, 4 and 8 and the results follows on Figures 9.4, 9.5 and 9.6, respectively :

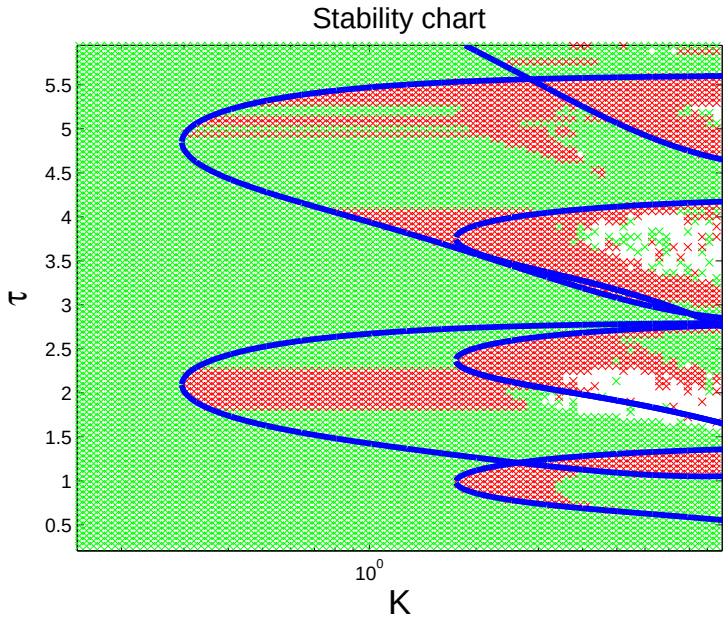

Figure 9.4: Stability charts estimation of $\mathbf{G}_d(k,\tau)$, machining chatter systems governed by (9.6) when the approximations have order 2.

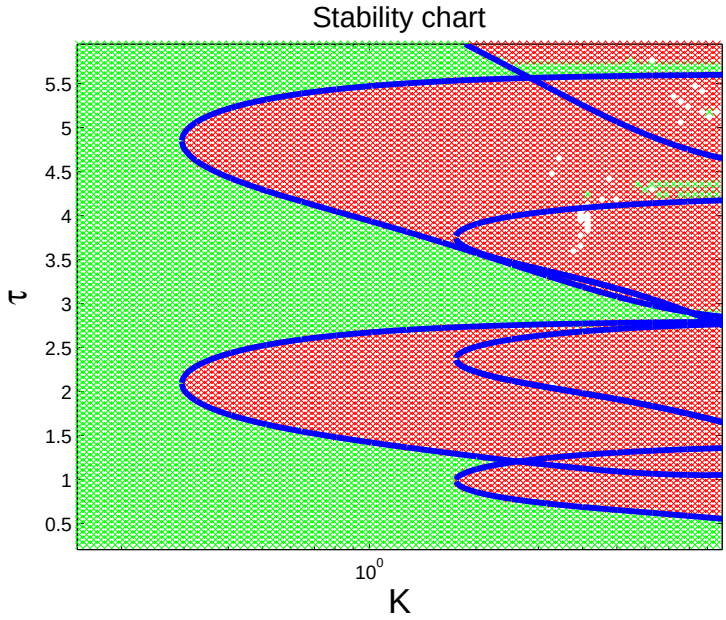

Figure 9.5: Stability charts estimation of $\mathbf{G}_d(k,\tau)$, machining chatter systems governed by (9.6) when the approximations have order 4.

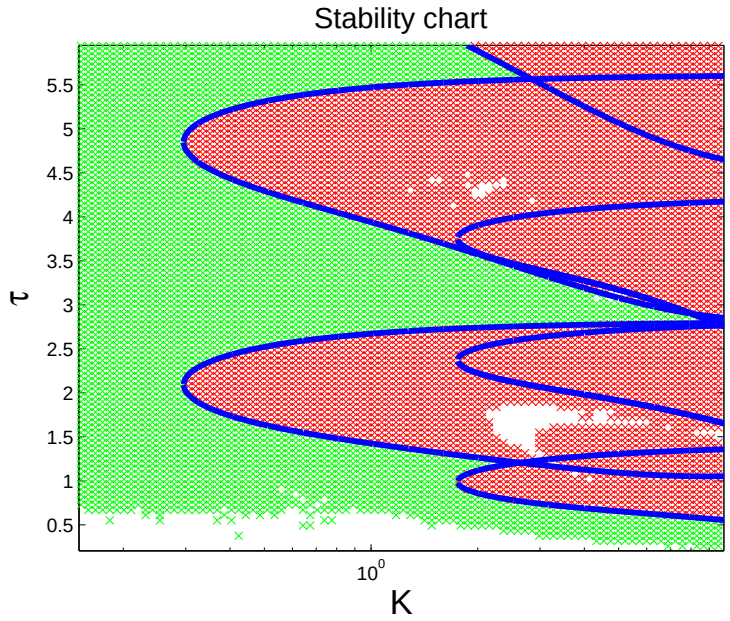

Figure 9.6: Stability charts estimation of $\mathbf{G}_d(k,\tau)$, machining chatter systems governed by (9.6) when the approximations have order 8.

One should note that an order $r = 2$ or 4 is not enough to estimate the stability of all $\mathbf{G}_d(k,\tau)$ and that is why some of the approximations are stable while the original system is unstable. On the other hand, for the order 8 the algorithm does not converge for a great amount of points. That is probably due to the fact that the 8-th approximation model $\hat{\mathbf{H}}$ is overparametrized. Practically this happens if the approximation order is too high to catch the input/output behavior, and, as a consequence, the shift points $(\sigma_1, \ldots, \sigma_n)$ never converge. Hence, a smaller order should be considered.

Hence, as long as the approximation order n is well chosen, the proposed brute-force procedure enables to capture almost exactly the stability charts of the system (9.6). In addition, based on the Remark 9.10, the approximations can be computed very quickly (2 or 3 iterations, instead of 50). In what follows we apply the same procedure in order to estimate the stability of a large scale time-delay system.

9.2.3 Application to large-scale LTI TDS

Let us now consider the clamped beam model from COMP$l_e ib$ library [Leibfritz and Lipinski, 2003] whose order is $n = 348$, denoted $\mathbf{G}_{BEAM}$ (presented as well in Chapter 1). For this system, a PID controller has been designed without taking into account any delay. The transfer function of this PID is given by

$$\mathbf{PID}(s) = \frac{9.791s^2 + 0.04095s + 0.0.07712}{s^2 + 0.0628s}.$$

Therefore the closed-loop system including the controller, denoted $\mathbf{H}_{PID}$, in series has dimension $n = 350$. We now add a constant delay $\tau \in [0, 10]$ between the output of the system and the input of the controller, and a scaling parameter $k \in [0, 1.5]$ on the gains of the PID controller. The resulting transfer function is then given by

$$\mathbf{G}_d(s, k, \tau) = \frac{k\mathbf{PID}(s)\mathbf{G}_{BEAM}(s)}{1 + k\mathbf{PID}(s)\mathbf{G}_{BEAM}(s)e^{-\tau s}}. \tag{9.7}$$

As before, we consider 60 equidistant grid points for the delay and for the controller gain k. Then, we compute $n = 14$-th order approximations using **TF-IRKA**. The order 14 was chosen because it provides approximation which are close enough to $\mathbf{G}_d$. Once again, the algorithm is initialized using the shift points derived from a neighbor system, in order to take advantage of the continuity of shifts property (see Remark 9.10). The results are depicted in Figure 9.7. As the system has one single delay, the Nyquist method can also be applied to construct an exact result, which is shown by the blue curve.

Reader should note that the brute-force procedure captures almost exactly the stability chart, with exception only for the points where the algorithm did not converge. Let us list the details from this simulation :

- The order $r = 14$ was chosen heuristically because it provides good approximation for a given (k, τ).
- $60 \times 60 = 3600$ approximations were computed in total.
- **TF-IRKA** needs 30 to 50 iterations to converge for this system for any given parameter (k_i, τ_j) couple.
- However, if we start the algorithm with the using the poles of a neighbor system, it converges in 2 or 3 iterations.
- The total time needed was 4500 seconds, *i.e.*, 1 hour and 15 minutes. Hence, the mean time per model approximation is 1.25s. Reader should notice that this was only possible because of the smart initialization of **TF-IRKA** (based on the Remark 9.10).
- The white points correspond to the parameters where **TF-IRKA** did not converge.

The example presented shows that the proposed procedure enables us to capture the stability charts of a large-scale LTI TDS. Hence, it can be used as an alternative to LMI-based and exact methods, which are not suitable in the large-scale context.

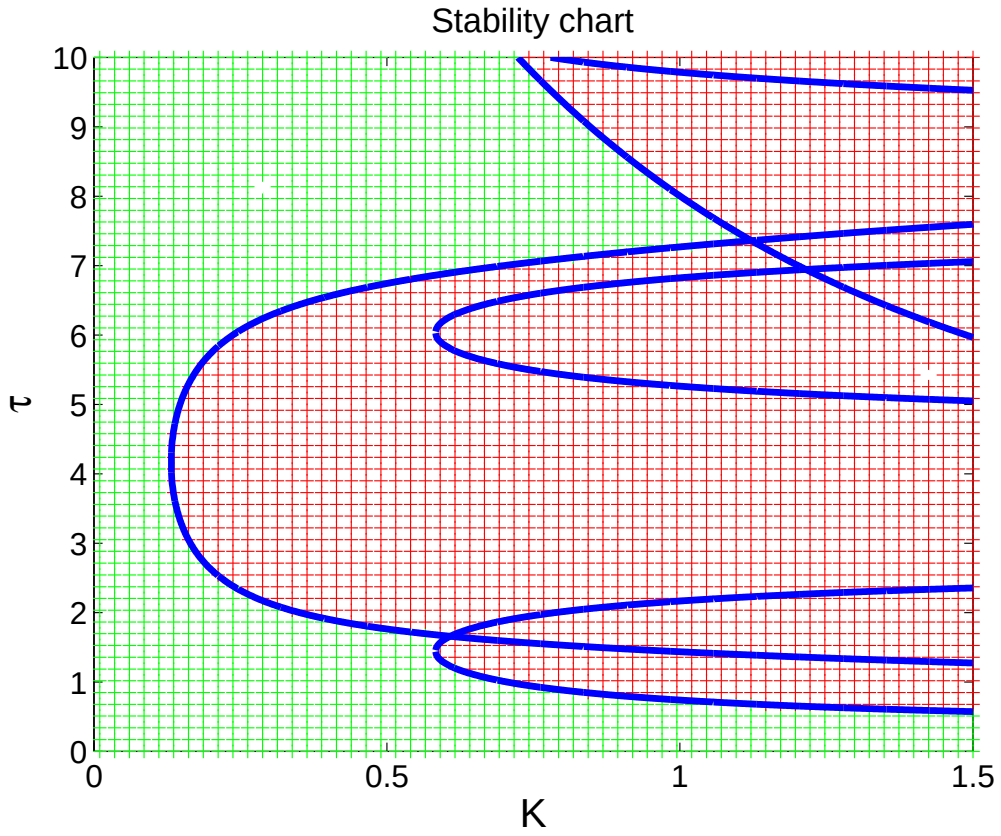

Figure 9.7: Stability charts estimation of $\mathbf{G}_d(k, \tau)$, delayed Beam model governed by (9.7). Green crosses correspond to stable approximations. Red crosses correspond to unstable approximation. The blue line corresponds to the stability frontier obtained by the Nyquist method.

In what follows, we propose an heuristic extension enabling to vary the order n from the approximation.

9.3 Heuristic procedure for stability charts of an LTI TDS

9.3.1 Varying the reduced order

As long as the order n is well chosen, it is very likely that if **TF-IRKA** converges, then it will preserve the stability. However, if n is too high the algorithm might not converge. In what follows, we propose a procedure to estimate the stability charts by varying the order n. We start with n high, so that the algorithm will capture the stability for the worst case scenarios. Then, we progressively decrease the order until we have an estimation of stability for every point in the domain. This procedure is sketched as follows.

Method 9.12 (**Model approximation based method for stability chart estimation**). *Given* $\mathbf{G}_d(\mathbf{p})$ *an LTI TDS, where the parameter* $\mathbf{p} \in \mathcal{D}$*, then follow the steps:*

(1) Discretize the parameter domain $\mathcal{D}$ *into* $\hat{\mathcal{D}}$*.*

(2) Choose a worst case order $k = r^{(0)}$ *such that*

(3) ***Repeat***

(i) For each $\mathbf{p}^i \in \hat{\mathcal{D}}$ *for which no stability estimation have been assigned, apply* ***TF-IRKA*** *for an order* k*.*

(ii) If ***TF-IRKA*** *converges, assign the stability associated to* $\mathbf{p}^i$ *as being the one from the approximation obtained.*

(iii) $k = k - 1$ *(the approximation order is decreased).*

Until a stability has been assigned for every $\mathbf{p}^i \in \hat{\mathcal{D}}$*.*

In order to justify this decreasing sequence of approximating order mechanism, let us remind the following points:

- If the **TF-IRKA** does not converge, it might be due to the fact that the approximation model $\hat{\mathbf{H}}$ is over-parametrized. In that case, the iterative shift σ_i selection might never stop and oscillate. Practically this happens if the approximation order is too high to catch the input/output behavior. Therefore, the order k is probably too high and is thus decreased.

- On the other hand, if **TF-IRKA** has converged, an appropriate approximation has been found and should well reproduce the input/output behavior, also in the $\mathcal{L}_2(i\mathbb{R})$-norm sense. Therefore, it is very likely that this approximation has the same stability as the original model.

In what follows, some numerical examples illustrate the efficiency and the accuracy of the approach.

9.3.2 Numerical Examples

Example 1: stabilizing feedback delay model

Let us consider the $\mathbf{G}_d(\tau, k)$ system of order 2 presented on Example 9.2. For this example let us first try to estimate the stability using order 6. Figure 9.8 shows the stability chart estimation for all parameters $\mathbf{p}^i \in \hat{\mathcal{D}}$ for which the algorithm has converged.

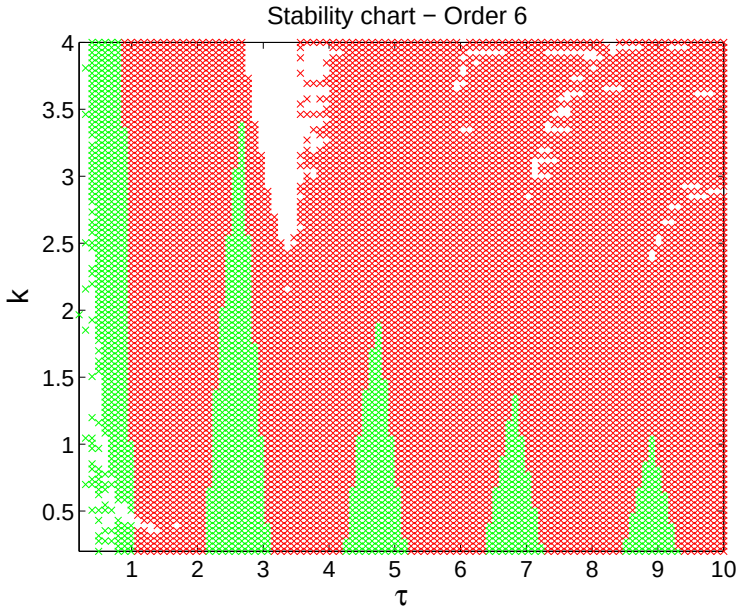

Figure 9.8: Stability charts estimation when the approximation has order 6.

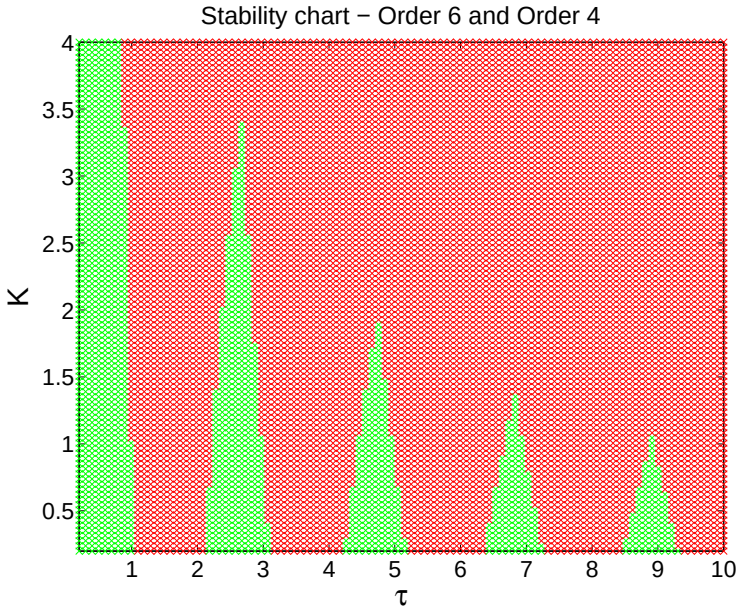

Figure 9.9: Stability charts estimation when the approximation has order 6 and 4.

Reader should notice that the white regions correspond to systems for which **TF-IRKA** was not able to converge for an order 6. Then, for the remaining parameters, the same brute force procedure is applied, now for 5-th order approximations. For none of this parameters the

algorithm had converged as well. Finally, **TF-IRKA** was applied for 4-th order approximations and enabling to assign an estimation for all the remaining parameters. Figure 9.9 shows the stability chart estimation (which is the same one presented on Example 9.2).

It is worth mentioning that the total time of simulation was around 190s.

Example 2: congestion system of a high speed network

Let us consider a congestion system of a high speed network, denoted $\mathbf{G}_{hsn}(\tau_1, \tau_2)$, originally derived from [Izmailov, 1996] and analyzed in [Niculescu, 2002] and [Sipahi et al., 2011], which behavior is driven by the following DDEs:

$$\mathbf{G}_{hsn}(\tau_1, \tau_2) = \begin{cases} \dot{\mathbf{x}}(t) = A_0 x(t) + A_1 x(t-\tau_1) + A_2 x(t-\tau_1-\tau_2) + \mathbf{b}u(t) \\ y(t) = \mathbf{c}^T x(t), \end{cases} \tag{9.8}$$

where $\mathbf{b} = [1\ 0]^T$, $\mathbf{c} = [0\ 1]$ and $(a = 2,\ b = -1.75)$

$$A_0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, A_1 = \begin{bmatrix} 0 & -a \\ 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & -b \\ 0 & 0 \end{bmatrix}.$$

We are interested in deriving the *stability chart* as a function of the couple $\{\tau_1, \tau_2\}$ delay values, *i.e.* , the conditions for $\tau_1 \in [0\ 1.4]$s and $\tau_2 \in [0\ 1.8]$s for which $\mathbf{G}_{hsn}(\tau_1, \tau_2)$ is stable or not. This specific problem has been theoretically solved in [Niculescu, 2002] using a deep analysis of its characteristic equation leading to the exact stability chart of $\mathbf{G}_{hsn}(\tau_1, \tau_2)$ as a function of $\{\tau_1, \tau_2\}$.

Hence, let us use the heuristic procedure 9.12. Firstly, let us consider a finite grid set of frozen $(\tau_1, \tau_2) = \mathbf{p}^i \in \hat{\mathcal{D}}$ values (here 100×100 grid points). Let us first try to estimate the stability using order 6. Figure 9.10 shows the stability chart estimation for all parameters $\mathbf{p}^i \in \hat{\mathcal{D}}$ for which the algorithm has converged.

Once again, the white regions correspond to systems for which **TF-IRKA** was not able to converge for an order 6. Then, for the remaining parameters, the same brute force procedure is applied, now for 5-th order approximations and the results are reported in Figure 9.11. Then, for the remaining parameters, the same brute force procedure is applied, now for 4-th order approximations. For none of this parameters the algorithm had converged as well. Finally, **TF-IRKA** was applied for 3-th order approximations and enabling to assign an estimation for all the remaining parameters, which is reported in Figure 9.12.

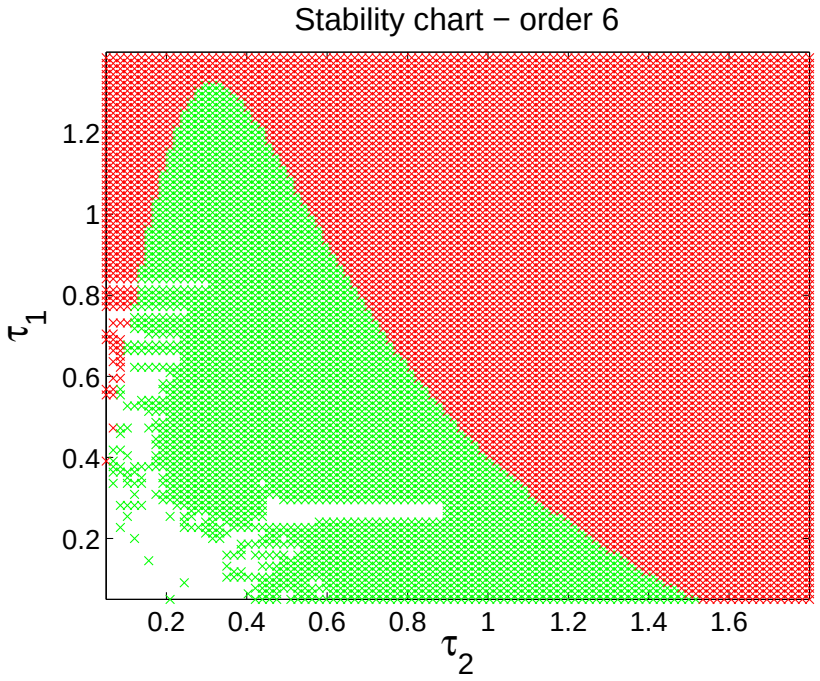

Figure 9.10: Stability charts estimation when the approximation has order 6.

Figure 9.11: Stability charts estimation when the approximation has orders 6 and 5.

Figure 9.12: Stability charts estimation when the approximation has orders 6, 5 and 3.

By comparing with the results in [Niculescu, 2002], it is clear that the proposed algorithm well catches the (in)stability property while being quite fast (result obtained in approximately 210 seconds). Although it is clear that Method 9.12 has neither formal guarantee of finding the stability proof it does provide a very good approximation of the stability charts.

Conclusion

In this chapter, the main contributions proposed are twofold: (i) a new paradigm to treat the stability of LTI TDS using interpolatory-based model approximation and (ii) the Method 9.12 enabling to estimate the stability charts of any LTI TDS. The other contribution correspond to the use of the continuity of shifts (Remark 9.10) in order to speed up the procedure. Obviously the approach is not restricted to delayed models but still required a smooth parameter variations (in the eigenspace). All the numerical results presented in this chapter are promising and the gain obtained in term of computational time involved appears spectacular in comparison with other tools, *e.g.*, LMI-based or bifurcation techniques. In addition, successful results have been shown in the large scale setting. Still, unlike these latter, no formal guarantee on the results obtained by the proposed heuristic can be established. For the case of unstable and α-stable systems, a certificate of guaranty can be provided *a posteriori*, but it is computationally expensive. Finally, this approach can be used as an initialization for more formal methods when applicable or provide an answer when the system to be analyzed is too complex.

Part IV

Conclusion

Chapter 10

Discussion and perspectives

Although throughout the study the majority of the results have been discussed and assessed, this chapter offers a general overview of the main contributions of this thesis highlighting their advantages and limitations.

On the topic of model approximation, this work sheds new light on the approximation of linear time invariant systems by reduced order models with richer structures, *e.g.*, input and output delays and state delays. The originality of this work comes from the combination of results from two different fields : time-delay systems and model approximation. Although it is a model approximation thesis, we believe that many results developed here can be useful in other fields, *e.g.*, analysis of stability using model approximation and the spectral $\mathcal{H}_2$ inner product formulation. In what follows, for each different topic developed in this thesis, we summarize the major contributions as well as the possible future research directions.

Model reduction for input/output-delay structure

Contribution

One of the main contributions is the study of the $\mathcal{H}_2$ model approximation problem when the reduced order model has input and output delays, which is the subject of Chapter 5. This problem is specially interesting when the original model has an intrinsic delay behavior, *e.g.*, transport equations. We believe that the following topics can be considered as the major contributions :

(i) **Spectral formulation of the $\mathcal{H}_2$ inner product in the presence of input/output delays :** This result corresponds to Theorem 5.5 for SISO systems and in Theorem 5.22 for MIMO systems. One of the main difference from the delay free case, is the non-symmetric expression of these expressions (see Example 5.6). This non-symmetry in the expression of the $\mathcal{H}_2$-inner product in the presence of input/output delays is the crucial point in the derivation of the optimality conditions.

(ii) **First-order $\mathcal{H}_2$ optimality conditions for input/output-delay structure :** This result corresponds to Theorem 5.10 for SISO systems and Theorem 5.20 for MIMO ones. As in the delay-free case, for a fixed input/output delay structure, they are also interpolation conditions. However, the main difference is that instead of interpolating the full order model $\mathbf{G}$, a new model $\tilde{\mathbf{G}}$ is here considered. The construction of the new $\tilde{\mathbf{G}}$ requires the pole/residue decomposition of $\mathbf{G}$ which can be expensive for very large scale systems.

(iii) **IO-dIRKA :** This algorithm is developed in Section 5.5 enabling to construct a reduced order model $\hat{\mathbf{H}}_d(s) = \hat{\boldsymbol{\Delta}}_o(s)\hat{\mathbf{H}}(s)\hat{\boldsymbol{\Delta}}_i(s)$. It corresponds to an iterative algorithm in which each iteration can be decomposed in two steps:

(a) for fixed delay structures $\hat{\boldsymbol{\Delta}}_o(s)$ and $\hat{\boldsymbol{\Delta}}_i(s)$, a reduced order model $\hat{\mathbf{H}}$ is obtained satisfying the interpolation conditions on $\tilde{\mathbf{G}}$.

(b) for a fixed reduced order model $\hat{\mathbf{H}}$, the optimal values for the delay structures $\hat{\boldsymbol{\Delta}}_o(s)$ and $\hat{\boldsymbol{\Delta}}_i(s)$ are determined solving an optimization problem.

If this procedure converges, it permits to construct a reduced order model having input/output structure satisfying the $\mathcal{H}_2$ optimality conditions.

Future research directions : general input/output structures

Instead of considering input/output delays, one might consider the model approximation problem where the reduced order model is represented by

$$\hat{\mathbf{H}}_{\text{iso}} = \hat{\mathbf{H}}\boldsymbol{\sigma}, \tag{10.1}$$

where $\hat{\mathbf{H}} = (E, A, B, C)$ is a finite dimensional reduced order model and $\boldsymbol{\sigma} \in \mathcal{H}_\infty$ represents a general input structure.

The case where $\boldsymbol{\sigma}$ is an $\mathcal{H}_2$ isometry, *i.e.*,

$$\|\hat{\mathbf{H}}\boldsymbol{\sigma}\|_{\mathcal{H}_2} = \|\hat{\mathbf{H}}\|_{\mathcal{H}_2},$$

is an example of possible extension. Notice that, in addition to input delays, others structures satisfy this isometric condition, *e.g.*, the Laguerre shifts. An example of such a structure is given by $\boldsymbol{\sigma}(s) = \dfrac{s-\tau}{s+\tau}$, for a given $\tau > 0$. Hence, the model approximation problem when the reduced order model has an isometric structure can be considered as a generalization for input/output-delay structures.

In addition, other input structures, not fitting in the isometric case, can also be interesting. As example, the reduced order models having more than one delay for the same input. The following system

$$\hat{\mathbf{H}}_d := \left\{ \begin{array}{rcl} \hat{E}\dot{\hat{\mathbf{x}}}(t) & = & \hat{A}\hat{\mathbf{x}}(t) + \hat{\mathbf{b}}(\mathbf{u}(t) + \mathbf{u}(t-\tau)) \\ \hat{\mathbf{y}}(t) & = & \hat{\mathbf{c}}\hat{\mathbf{x}}(t) \end{array} \right. \tag{10.2}$$

illustrates one of this cases.

Data driven model approximation for single state-delay structure

Contribution

In Chapter 6, a data-driven framework for single state-delay systems, *i.e.*, models denoted by $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ whose transfer functions are represented by

$$\hat{\mathbf{H}}_d(s) = \hat{C}(\hat{E}s - \hat{A}e^{-s\tau})^{-1}\hat{B}.$$

The main contributions are summarized as follows :

(i) **State-delay transformation :** Lemma 6.2 states that a single state-delay model $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ represented by a delay-free model $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ by means of the transformation

$$\hat{\mathbf{H}}_d(s) = \hat{\mathbf{H}}\big(f(s)\big)e^{s\tau}$$

where $\hat{\mathbf{H}}(s)$ is the transfer function of the delay-free model $\hat{\mathbf{H}} = (\hat{E}, \hat{A}, \hat{B}, \hat{C})$ and

$$f(s) = se^{s\tau}.$$

(ii) **Single state-delay Loewner framework :** Theorems 6.3 and 6.4 provide a method to construct a model $\hat{\mathbf{H}}_d = (\hat{E}, \hat{A}, \hat{B}, \hat{C}, \tau)$ whose transfer function $\hat{\mathbf{H}}_d(s) = \hat{C}(s\hat{E} - \hat{A}e^{-s\tau})^{-1}\hat{B}$ interpolates some tangential interpolation data.

(iii) **dTF-IRKA :** This algorithm is developed in Section 6.4 enabling to construct a reduced order single state-delay model satisfying some interpolation conditions inspired in the finite dimensional case.

Those results interpolation results were generalized for systems having a more general delay structure, *i.e.*,

$$\hat{\mathbf{H}}_d(s) = \hat{C}(\hat{E}s - \hat{A}_0 - \hat{A}_1e^{-s\tau})^{-1}\hat{B}$$

in [Schulze and Unger, 2015], assuming that matrix $\hat{A}_1$ is a linear combination between $\hat{E}$ and $\hat{A}_1$.

Future research direction: delay optimization and more general structures

One weakness of the proposed method, is the fact that one should know in advance the delay value τ. Future works will investigate this issue by taking into consideration the delays as decision variables in the $\mathcal{H}_2$ optimization problem. The extension to more general structures, *e.g.*, the multiple delays case, should also be addressed.

$\mathcal{H}_2$ optimality conditions for single state-delay models

Contribution

Chapter 7 introduces the $\mathcal{H}_2$ optimal approximation problem when the reduced order model is a single state-delay model of dimension one, *i.e.*, when the reduced order model $\hat{\mathbf{H}}_d$ has the following representation :

$$\hat{\mathbf{H}}_d(s) = \frac{\hat{\phi}}{s - \hat{\alpha}e^{-s\tau}}.$$

(i) **Spectral $\mathcal{H}_2$ inner product for single state-delay models :** Based on the spectral decomposition of a single state-delay model, Propositions 7.4 and 7.5 provide a spectral characterization of the $\mathcal{H}_2$ inner product. Due to the infinite dimensional nature of the reduced order models considered here, those expressions are no longer sums, but infinite series.

(ii) **$\mathcal{H}_2$ optimality conditions for single state delay reduced order models :** The main contribution of this chapter is Theorem 7.8. It generalizes the $\mathcal{H}_2$ optimality conditions to the case where the ROM is a single state-delay model. In this case, instead of simple

interpolation, those conditions are infinite series interpolation. Therefore, interpolation based methods are no longer applicable and further numerical developments need to be considered.

To sum up, Theorem 7.8 generalizes Theorem 4.14 for single state-delay reduced-order models. Due to the infinite dimensional nature of time-delay systems, the $\mathcal{H}_2$ optimality conditions here are not interpolation conditions any more and become interpolation of series. Hence, optimum $\mathcal{H}_2$ reduced order models cannot be constructed by interpolation methods and new tools enabling interpolation of sums and series might be a future research direction.

Future generalization directions: Other structures

Following similar arguments, we could consider other simple structures, *e.g.*,

(i) **Second order structures :** Let us consider the simple second order structure defined as follows :

$$\mathbf{h}_{2nd}(s) = \frac{1}{s^2 + \beta_1 s + \beta_2}.$$

(ii) **More general delay structures :** Let us consider the simple second order structure defined as follows :

$$\mathbf{h}_{d}(s) = \frac{1}{s + \alpha_1 + \alpha_2 e^{-s\tau}}.$$

In addition, the reduced order mode given by a linear combination of those proposed structures, *i.e.*,

$$\hat{\mathbf{H}}_{Str2nd} = \sum_{k=1}^{n} \hat{\phi}_k \mathbf{h}_{2nd}^k(s),$$

or

$$\hat{\mathbf{H}}_{StrDelay}(s) = \sum_{k=1}^{n} \hat{\phi}_k \mathbf{h}_{d}^k(s).$$

could be also considered. Then the $\mathcal{H}_2$ approximation problem for such structured systems could be derived following the same ideas presented on Chapter 7. As a consequence, these conditions should be also interpolation of sums (for the second order structured case) or interpolation of series (for the state delay case) and further results on interpolation of sums/series should be also derived.

Model approximation based framework for evaluating the stability of a time-delay system

A model approximation framework is developed in order to evaluate the stability of a time-delay system. Systems are considered to be elements of $\mathcal{L}_2(i\mathbb{R})$ (instead of $\mathcal{H}_2$). The set of stable and unstable systems are characterized as subsets of $\mathcal{L}_2(i\mathbb{R})$, leading to some topological and approximation results. Thereby, a certificate of instability is derived based on the numerical estimation of the $\mathcal{L}_2(i\mathbb{R})$ norm. Equipped with those results, one can estimate the stability of any time-delay systems by finding a model approximation with is good enough in the sense of $\mathcal{L}_2(i\mathbb{R})$. This results are presented in Chapters 8 and 9.

Contribution

The main contributions are summarized as follows :

(i) **Topology of stability:** Based on the Hardy space characterization of stability, some interesting properties linking model approximations and stability are derived in Section 8.2 from Chapter 8 (for unstable and stable systems) and Section 9.1 from Chapter 9 (concerning α-stable systems). The main contribution is the awareness that the set of unstable and α-stable systems are open sets of $\mathcal{L}_2(i\mathbb{R})$. Thereby, (i) if an unstable approximation is close enough to a given system, it is also unstable and (ii) if an stable approximation is close enough to a given system, it is α-stable.

(ii) **Certificate of instability guaranty:** Theorems 8.6 (Chapter 8) and 9.5 (Chapter 9) give sufficient conditions to a system to be unstable and α-stable, respectively, based on it distance from an unstable approximation.

(iii) **Method to evaluate instability/α-stability:** In Chapter 8, the approximation-based Method 8.7 is proposed in order the evaluate if a given system is unstable. The same concept can be used to evaluate α-stability and the approximation can be constructed with **TF-IRKA**.

(iv) **Brute-force procedure :** the procedure 9.8 enables to estimate the stability charts of any **LTI TDS**.

(v) **Continuity of shifts:** as stated in Remark 9.10, the shift points behave continuously when the original model is perturbed. This is used to speed up the brute force procedure. The results are applied in some numerical examples, including a large-scale **LTI TDS**.

(vi) **Heuristic procedure:** In Section 9.3, a procedure enabling to vary the order n is sketched and applied to some numerical examples.

The overall approach have been validated on many different TDS, including a large-scale TDS, providing promising perspectives for the TDS stability estimation in the large-scale setting. Future works should include comparisons with some other literature methods, *e.g.*, LMI-based or bifurcation techniques.

Future research directions

Based on the results presented, a boundary-search heuristic algorithm can be developed so the number of model approximations required to establish the stability charts estimation could be drastically reduced. This work is on progress and the following paper on this topic has been submitted [Poussot-Vassal et al., 2016].

Bibliography

Abdallah, G., Dorato, P., Benitez-Read, J., and Byrne, R. (1993). Delayed positive feedback can stabilize oscillatory systems. In *Proceedings of the IEE American Control Conference*, page 3106–3107. 152

Alevisakis, G. and Seborg, D. (1973). An extension of the smith predictor method to multivariable linear systems containing time delays. *International Journal of Control*, 17(3):541–551. 27

Anderson, B. and Antoulas, A. (1990). Rational interpolation and state-variable realizations. *Linear Algebra and its Applications*, 137:479–509. 46

Anić, B., Beattie, C., Gugercin, S., and Antoulas, A. C. (2013). Interpolatory weighted-$\mathcal{H}_2$ model reduction. *Automatica*, 49(5):1275–1280. 33, 35

Antoulas, A. and Anderson, B. (1986). On the scalar rational interpolation problem. *IMA Journal of Mathematical Control and Information*, 3(2-3):61–88. 46

Antoulas, A. and Astolfi, A. (2002). $\mathcal{H}_\infty$-norm approximation. *MTNS Problem Book, Open Problems on the Mathematical Theory of Networks and Systems, V. Blondel and A. Megretski Eds*, pages 267–270. 57

Antoulas, A. C. (2005). *Approximation of large-scale dynamical systems*, volume 6. SIAM, Philadelphia. 15, 26, 33, 35, 39, 40, 42, 43, 60, 68

Antoulas, A. C., Beattie, C. A., and Gugercin, S. (2010). Interpolatory model reduction of large-scale dynamical systems. In *Efficient modeling and control of large-scale systems*, pages 3–58. Springer. 43

Antoulas, A. C. and Gosea, I. V. (2015). Data-driven model reduction for weakly nonlinear systems: A summary. In *Proceedings of the 8th Vienna International Conference on Mathematical Modelling*, volume 48, pages 3–4. 50

Antoulas, A. C., Sorensen, D. C., and Gugercin, S. (2001). A survey of model reduction methods for large-scale systems. *Contemporary mathematics*, 280:193–220. 35

Astolfi, A. (2010). Model reduction by moment matching for linear and nonlinear systems. *IEEE Transactions on Automatic Control*, 55(10):2321–2336. 45

Bai, Z. (2002). Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. *Applied numerical mathematics*, 43(1):9–44. 43

Bai, Z. and Skoogh, D. (2006). A projection method for model reduction of bilinear dynamical systems. *Linear Algebra and its Applications*, 415(2):406–425. 45

Balluchi, A., Bicchi, A., Mazzi, E., Vincentelli, A. L. S., and Serra, G. (2006). Hybrid modelling and control of the common rail injection system. In *Proceedings of the International Workshop on Hybrid Systems: Computation and Control*, pages 79–92. Springer. 27

Baratchart, L. (1986). Existence and generic properties of l^2 approximants for linear systems. *IMA Journal of Mathematical Control and Information*, 3(2-3):89–101. 58

Baratchart, L., Cardelli, M., and Olivi, M. (1991). Identification and rational l^2 approximation a gradient algorithm. *Automatica*, 27(2):413–417. 58

Baur, U., Benner, P., and Feng, L. (2014). Model order reduction for linear and nonlinear systems: a system-theoretic perspective. *Archives of Computational Methods in Engineering*, 21(4):331–358. 35

Beattie, C. and Gugercin, S. (2009a). Interpolatory projection methods for structure-preserving model reduction. *Systems & Control Letters*, 58(3):225–232. 44, 126

Beattie, C. and Gugercin, S. (2009b). A trust region method for optimal $\mathcal{H}_2$ model reduction. In *Proceedings of the 48th IEEE Conference on Decision and Control*, pages 5370–5375. 58

Beattie, C. and Gugercin, S. (2011). Structure-preserving model reduction for nonlinear port-Hamiltonian systems. In *Proceedings of the IEE 50th IEEE Conference on Decision and Control and European Control Conference*, pages 6564–6569. 4, 6, 7, 45, 108

Beattie, C. and Gugercin, S. (2012). Realization-independent $\mathcal{H}_2$-approximation. In *Proceedings of the 51st IEEE Conference on Decision and Control*, pages 4953–4958. 12, 14, 58, 76, 78, 79, 94, 113, 114, 120, 137

Beattie, C. and Gugercin, S. (2015). Model reduction by rational interpolation. *To appear in Model Reduction and Approximation: Theory and Algorithms*. 35

Beattie, C. A. and Gugercin, S. (2007). Krylov-based minimization for optimal $\mathcal{H}_2$ model reduction. In *Proceedings of the 46th IEEE conference on Decision and Control*, pages 4385–4390. 58

Bellman, R. E. and Cooke, K. L. (1963). Differential-difference equations. 27, 29, 139, 140

Bender, D. (1987). Lyapunov-like equations and reachability/observabiliy gramians for descriptor systems. *IEEE Transactions on Automatic Control*, 32(4):343–348. 40

Benner, P. and Breiten, T. (2012). Krylov-subspace based model reduction of nonlinear circuit models using bilinear and quadratic-linear approximations. In *Progress in Industrial Mathematics at ECMI 2010*, pages 153–159. Springer Berlin Heidelberg. 45

Benner, P., Gugercin, S., and Willcox, K. (2015). A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM review*, 57(4):483–531. 35, 45

Benner, P., Li, R.-C., and Truhar, N. (2009). On the ADI method for Sylvester equations. *Journal of Computational and Applied Mathematics*, 233(4):1035–1045. 26

Benner, P., Quintana-Ortí, E. S., and Quintana-Ortí, G. (2006). Solving stable Sylvester equations via rational iterative schemes. *Journal of Scientific Computing*, 28(1):51–83. 26

Benner, P. and Saak, J. (2013). Numerical solution of large and sparse continuous time algebraic matrix Riccati and Lyapunov equations: a state of the art survey. *GAMM-Mitteilungen*, 36(1):32–52. 40

Benner, P. and Sokolov, V. I. (2006). Partial realization of descriptor systems. *Systems & Control Letters*, 55(11):929–938. 22, 25

Benner, P. and Stykel, T. (2015). Model order reduction for differential-algebraic equations: a survey. *preprint MPIMD/15-19, Max Planck Institute Magdeburg*. 35

Breda, D. (2006). Solution operator approximations for characteristic roots of delay differential equations. *Applied Numerical Mathematics*, 56(3):305–317. 139

Breda, D., Maset, S., and Vermiglio, R. (2005). Pseudospectral differencing methods for characteristic roots of delay differential equations. *SIAM Journal on Scientific Computing*, 27(2):482–495. 139

Breiten, T., Beattie, C., and Gugercin, S. (2015). Near-optimal frequency-weighted interpolatory model reduction. *Systems & Control Letters*, 78:8–18. 33, 35

Breiten, T. and Damm, T. (2010). Krylov subspace methods for model order reduction of bilinear control systems. *Systems & Control Letters*, 59(8):443–450. 45

Bresch-Pietri, D., Leroy, T., and Petit, N. (2014). Control-oriented input-delay model of the distributed temperature of a si engine exhaust catalyst. In *Low-Complexity Controllers for Time-Delay Systems*, pages 173–188. Springer. 82

Briat, C. (2015). *Linear Parameter-Varying and Time-Delay Systems – Analysis, Observation, Filtering & Control*. Advances on Delays and Dynamics, Vol. 3. Springer-Verlag, Heidelberg, Germany. 11, 27, 153

Bruinsma, N. and Steinbuch, M. (1990). A fast algorithm to compute the $\mathcal{H}_\infty$-norm of a transfer function matrix. *Systems & Control Letters*, 14(4):287–293. 20

Bryson, A. E. and Carrier, A. (1990). Second-order algorithm for optimal model order reduction. *Journal of Guidance, Control, and Dynamics*, 13(5):887–892. 58

Byrnes, C. I. and Falb, P. L. (1979). Applications of algebraic geometry in system theory. *American Journal of Mathematics*, pages 337–363. 119

Cantoni, M. (2001). On model reduction in the ν-gap metric. In *Proceedings of the 40th IEEE Conference on Decision and Control, 2001*, volume 4, pages 3665–3670. 57

Cepeda-Gomez, R. and Michiels, W. (2015). Some special cases in the stability analysis of multi-dimensional time-delay systems using the matrix Lambert W function. *Automatica*, 53:339–345. 128

Chahlaoui, Y., Lemonnier, D., Vandendorpe, A., and Van Dooren, P. (2006). Second-order balanced truncation. *Linear Algebra and its Applications*, 415(2):373–384. 40

Chaturantabut, S., Beattie, C., and Gugercin, S. (2016). Structure-preserving model reduction for nonlinear port-Hamiltonian systems. *arXiv preprint arXiv:1601.00527*. 45

Chellaboina, V. and Haddad, W. M. (1995). Is the Frobenius matrix norm induced? *IEEE Transactions on Automatic Control*, 40(12):2137–2139. 56

Corless, R. M., Gonnet, G. H., Hare, D. E., Jeffrey, D. J., and Knuth, D. E. (1996). On the Lambert W function. *Advances in Computational mathematics*, 5(1):329–359. 115, 118

Curtain, R. and Morris, K. (2009). Transfer functions of distributed parameter systems: A tutorial. *Automatica*, 45(5):1101–1116. 17

Curtain, R. F. and Zwart, H. (2012). *An introduction to infinite-dimensional linear systems theory*, volume 21. Springer Science & Business Media. 138

Dai, L. (1989). *Singular control systems*. Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin. 22

Dalmas, V., Robert, G., Poussot-Vassal, C., Pontes Duff, I., and Seren, C. (2016). From infinite dimensional modelling to parametric reduced-order approximation: Application to open-channel flow for hydroelectricity. In *Proceedings of the IEEE 15th European Control Conference*. 4, 7, 8, 109

Datko, R. (1978). A procedure for determination of the exponential stability of certain differential-difference equations. *Quarterly of Applied Mathematics*, pages 279–292. 139

Dym, H., Georgiou, T. T., and Smith, M. C. (1995). Explicit formulas for optimally robust controllers for delay systems. *IEEE Transactions on Automatic Control*, 40(4):656–669. 27

Engelborghs, K., Luzyanina, T., and Samaey, G. (2001). DDE-BIFTOOL v. 2.00: a Matlab package for bifurcation analysis of delay differential equations. 153

Engelborghs, K. and Roose, D. (2002). On stability of lms methods and characteristic roots of delay differential equations. *SIAM Journal on Numerical Analysis*, 40(2):629–650. 139

Feldmann, P. and Freund, R. W. (1995). Efficient linear circuit analysis by Padé approximation via the lanczos process. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 14(5):639–649. 43

Flagg, G., Beattie, C., and Gugercin, S. (2012). Convergence of the iterative rational Krylov algorithm. *Systems & Control Letters*, 61(6):688–691. 78

Flagg, G., Beattie, C. A., and Gugercin, S. (2013). Interpolatory $\mathcal{H}_\infty$ model reduction. *Systems & Control Letters*, 62(7):567–574. 33, 58

Foias, C., Özbay, H., and Tannenbaum, A. (1996). *Robust control of infinite dimensional systems: A frequency domain method*, volume 209 of *Lecture notes in control and information sciences*. Springer, Berlin. 27

Freund, R. W. (2003). Model reduction methods based on Krylov subspaces. *Acta Numerica*, 12(1):267–319. 43

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. 58

Gallivan, K., Grimme, E., and Van Dooren, P. (1994). Asymptotic waveform evaluation via a lanczos method. *Applied Mathematics Letters*, 7(5):75–80. 43

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. 41, 42, 43, 77

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. 34, 43

Georgiou, T. T. and Smith, M. C. (1990). Optimal robustness in the gap metric. *IEEE Transactions on Automatic Control*, 35(6):673–686. 57

Glover, K. (1984). All optimal Hankel-norm approximations of linear multivariable systems and their L_∞-error bounds. *International Journal of Control*, 39(6):1115–1193. 4, 34

Glover, K., Curtain, R. F., and Partington, J. R. (1988). Realisation and approximation of linear infinite-dimensional systems with error bounds. *SIAM Journal on Control and Optimization*, 26(4):863–898. 34, 40

Gosea, I. V. and Antoulas, A. C. (2015). Model reduction of linear and nonlinear systems in the Loewner framework: A summary. In *Proceedings of the IEE European Control Conference*, pages 345–349. 50

Green, M. and Limebeer, D. J. N. (2012). *Linear robust control*. Prentice-Hall. 56

Grimme, E. J. (1997). *Krylov projection methods for model reduction*. PhD thesis, University of Illinois, Urbana-Champaign, Urbana, IL. 42, 43, 77

Gu, K., Chen, J., and Kharitonov, V. L. (2003). *Stability of time-delay systems*. Springer Science & Business Media. 27, 138, 139

Gugercin, S. and Antoulas, A. C. (2004). A survey of model reduction by balanced truncation and some new results. *International Journal of Control*, 77(8):748–766. 35, 38, 40

Gugercin, S., Antoulas, A. C., 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. 4, 26, 43, 55, 58, 60, 74, 76, 77, 78, 83, 88, 93, 100, 101, 102, 111, 114

Gugercin, S., Beattie, C., and Antoulas, A. (2006). Rational Krylov methods for optimal $\mathcal{H}_2$ model reduction. In *Proceedings of the International Symposium on Mathematical Theory of Networks and Systems*, page 1665 – 1667. 58, 77

Gugercin, S., Polyuga, R. V., Beattie, C., and Van Der Schaft, A. (2012). Structure-preserving tangential interpolation for model reduction of port-Hamiltonian systems. *Automatica*, 48(9):1963–1974. 6, 45, 108

Guiver, C. and Opmeer, M. R. (2014). Model reduction by balanced truncation for systems with nuclear Hankel operators. *SIAM Journal on Control and Optimization*, 52(2):1366–1401. 40

Hale, J. K. and Lunel, S. M. V. (2013). *Introduction to functional differential equations*, volume 99. Springer Science & Business Media. 139

Halevi, Y. (1992). Frequency weighted model reduction via optimal projection. *IEEE Transactions on Automatic Control*, 37(10):1537–1542. 35, 58

Halevi, Y. (1996). Reduced-order models with delay. *International Journal of Control*, 64(4):733–744. 83

Heinkenschloss, M., Reis, T., and Antoulas, A. C. (2011). Balanced truncation model reduction for systems with inhomogeneous initial conditions. *Automatica*, 47(3):559–564. 40

Hoffman, K. (1962). *Banach spaces of analytic functions*. Prentice-Hall. 17

Hsu, C. S., Desai, U. B., and Darden, R. J. (1983). Reduction of large-scale systems via generalized gramians. In *Proceedings of the IEE 22nd IEEE Conference on Decision and Control*, pages 1409–1410. 38

Hyland, D. C. and Bernstein, D. S. (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. 58, 83

Ionescu, T. C. and Astolfi, A. (2013). Families of moment matching based, structure preserving approximations for linear port-Hamiltonian systems. *Automatica*, 49(8):2424–2434. 45

Ionita, A. and Antoulas, A. (2014a). Data-driven parametrized model reduction in the Loewner framework. *SIAM Journal on Scientific Computing*, 36(3):A984–A1007. 50

Ionita, A. C. and Antoulas, A. C. (2014b). Case study: Parametrized reduction using reduced-basis and the Loewner framework. In *Reduced Order Methods for Modeling and Computational Reduction*, pages 51–66. Springer. 50

Ionita, C. (2013). *Lagrange rational interpolation and its applications to model reduction and system identification*. PhD thesis, Rice University. 46

Izmailov, R. (1996). Analysis and optimization of feedback control algorithms for data transfers in high-speed networks. *SIAM journal on control and optimization*, 34(5):1767–1780. 168

Jarlebring, E., Vanbiervliet, J., and Michiels, W. (2011). Characterizing and computing the $\mathcal{H}_2$ norm of time-delay systems by solving the delay Lyapunov equation. *IEEE Transactions on Automatic Control*, 56(4):814–825. 132

Krajewski, W., Lepschy, A., Redivo-Zaglia, M., and Viaro, U. (1995). A program for solving the L_2 reduced-order model problem with fixed denominator degree. *Numerical Algorithms*, 9(2):355–377. 77, 155

Krasovskii, N. N. and Brenner, J. L. (1963). *Stability of motion: applications of Lyapunov's second method to differential systems and equations with delay*. Stanford university press. 139

Kreyszig, E. (1989). *Introductory functional analysis with applications*, volume 81. Wiley New York. 59, 67

Kunkel, P. and Mehrmann, V. L. (2006). *Differential-algebraic equations: analysis and numerical solution*. European Mathematical Society. 22

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. 38

Leibfritz, F. and Lipinski, W. (2003). Description of the benchmark examples in COMPleib 1.0. *Dept. Math., Univ. Trier, Trier, Germany, Tech. Rep.* 4, 123, 163

Lepschy, A., Mian, G., Pinato, G., and Viaro, U. (1991). Rational L_2 approximation: A non-gradient algorithm. In *Proceedings of the 30th IEEE Conference on Decision and Control*, pages 2321–2323. IEEE. 58

Litrico, X. and Fromion, V. (2009). *Modeling and control of hydrosystems*. Springer Science & Business Media. 82

Litrico, X., Pomet, J.-B., and Guinot, V. (2010). Simplified nonlinear modeling of river flow routing. *Advances in Water Resources*, 33(9):1015–1023. 82

Magruder, C., Beattie, C., and Gugercin, S. (2010). Rational Krylov methods for optimal $\mathcal{L}_2$ model reduction. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 6797–6802. 147, 148

Marmorat, J.-P., Olivi, M., Hanzon, B., and Peeters, R. L. M. (2002). Matrix rational $\mathcal{H}_2$ approximation: a state-space approach using schur parameters. In *Proceedings of the 41st IEEE Conference on Decision and Control*, volume 4, pages 4244–4249. 58

Martins, N., Lima, L. T., and Pinto, H. (1996). Computing dominant poles of power system transfer functions. *IEEE Transactions on Power Systems*, 11(1):162–170. 35

Mayo, A. J. and Antoulas, A. C. (2007). A framework for the solution of the generalized realization problem. *Linear Algebra and its Applications*, 425(2):634–662. 4, 46, 47, 48, 50, 83, 100, 106, 113, 114, 124

Mehrmann, V. and Stykel, T. (2005). Balanced truncation model reduction for large-scale systems in descriptor form. In *Dimension Reduction of Large-Scale Systems*, pages 83–115. Springer. 26, 40

Meier III, L. and Luenberger, D. G. (1967). Approximation of linear constant systems. *IEEE Transactions on Automatic Control*, 12(5):585–588. 57

Meinsma, G. and Zwart, H. (2000). On $\mathcal{H}_\infty$ control for dead-time systems. *IEEE Transactions on Automatic Control*, 45(2):272–285. 27

Meyer, D. G. and Srinivasan, S. (1996). Balancing and model reduction for second-order form linear systems. *IEEE Transactions on Automatic Control*, 41(11):1632–1644. 40

Mi, W., Qian, T., and Wan, F. (2012). A fast adaptive model reduction method based on takenaka–malmquist systems. *Systems & Control Letters*, 61(1):223–230. 68

Michiels, W., Jarlebring, E., and Meerbergen, K. (2011). Krylov-based model order reduction of time-delay systems. *SIAM Journal on Matrix Analysis and Applications*, 32(4):1399–1421. 45

Michiels, W. and Niculescu, S.-I. (2014). *Stability, Control, and Computation for Time-delay Systems: An Eigenvalue-based Approach*, volume 27. SIAM, Singapore. 11, 29, 139, 140

Moore, B. (1981). Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, 26(1):17–32. 4, 37

Mullis, C. and Roberts, R. (1976). Synthesis of minimum roundoff noise fixed point digital filters. *IEEE Transactions on Circuits and Systems*, 23(9):551–562. 37

Niculescu, S.-I. (2001a). *Delay effects on stability: a robust control approach*, volume 269. Springer Science & Business Media. 27, 139

Niculescu, S.-I. (2001b). *Delay effects on stability. A robust control approach*, volume 269. Springer-Verlag: Heidelbeg. 27

Niculescu, S.-I. (2002). On delay robustness analysis of a simple control algorithm in high-speed networks. *Automatica*, 38(5):885–889. 168, 170

Olivi, M., Seyfert, F., and Marmorat, J.-P. (2013). Identification of microwave filters by analytic and rational $\mathcal{H}_2$ approximation. *Automatica*, 49(2):317–325. 58

Opmeer, M. (2015). Optimal model reduction for non-rational functions. In *Proceedings of the 14th IEE European Control Conference*, pages 362–367. 58

Partington, J. R. (1988). *An introduction to Hankel operators*, volume 13. Cambridge University Press. 19

Partington, J. R. (1997). *Interpolation, identification, and sampling*. Number 17. Oxford University Press. 15, 17, 19

Partington, J. R. (2004). *Linear operators and linear systems: an analytical approach to control theory*, volume 60. Cambridge University Press. 15, 127, 138

Partington, J. R. and Bonnet, C. (2004). $\mathcal{H}_\infty$ and BIBO stabilization of delay systems of neutral type. *Systems & Control Letters*, 52(3):283–288. 141

Petersson, D. (2013). A nonlinear optimization approach to $\mathcal{H}_2$-optimal modeling and control. 35, 58

Petersson, D. and Löfberg, J. (2014). Model reduction using a frequency-limited $\mathcal{H}_2$-cost. *Systems & Control Letters*, 67:32–39. 33, 35

Pontes Duff, I., Gugercin, S., Beattie, C., Poussot-Vassal, C., and Seren, C. (2016a). $\mathcal{H}_2$-optimality conditions for reduced time-delay systems of dimension one. In *Proceedings of the 13th IFAC workshop on Time-Delay Systems*, volume 49, pages 7 – 12. 12, 125, 130, 131

Pontes Duff, I., Poussot-Vassal, C., and Seren, C. (2015a). Realization independent single time-delay dynamical model interpolation and $\mathcal{H}_2$-optimal approximation. In *Proceedings of the 54st IEEE Conference on Decision and Control*. 12, 50, 83

Pontes Duff, I., Poussot-Vassal, C., and Seren, C. (2016b). Optimal $\mathcal{H}_2$ model approximation based on multiple input/output delays systems. *Submitted, Arxiv version available at http://arxiv.org/abs/1511.05252*. 11, 81

Pontes Duff, I., Vuillemin, P., Poussot-Vassal, C., Briat, C., and Seren, C. (2015b). Approximation of stability regions for large-scale time-delay systems using model reduction techniques. In *Proceedings of the 14th European Control Conference*, pages 356–361. 12, 14, 79

Pontes Duff, I., Vuillemin, P., Poussot-Vassal, C., Briat, C., and Seren, C. (2015c). *Model reduction for norm approximation: An application to large-scale time-delay systems*. Advances in Dynamics and Delays. Springer-Verlag. 128

Pontryagin, L. S. (1955). On the zeros of some elementary transcendental functions. *American Mathematical Society Translations*, 2:95–110. 139

Poussot-Vassal, C., Loquen, T., Vuillemin, P., Cantinaud, O., and Lacoste, J.-P. (2013). Business jet large-scale model approximation and vibration control. In *Proceedings of the IFAC ALCOSP*, volume 11, pages 199–204. 57

Poussot-Vassal, C., Seren, C., Pontes Duff, I., and Vuillemin, P. (2016). Numerical approach to time-delay dynamical systems stability chart approximation through interpolatory techniques. *Submitted*. 177

Razumikhin, B. S. (1956). On the stability of systems with a delay. *Prikladnava Matematika i Mekhanika*, 20(4):500–512. 139

Reis, T. and Selig, T. (2014). Balancing transformations for infinite-dimensional systems with nuclear Hankel operator. *Integral Equations and Operator Theory*, 79(1):67–105. 40

Richard, J.-P. (2003). Time-delay systems: an overview of some recent advances and open problems. *Automatica*, 39(10):1667–1694. 11, 27, 29

Riggs, J. B. and Edgar, T. F. (1974). Least squares reduction of linear systems using impulse response. *International Journal of Control*, 20(2):213–223. 57

Rommes, J. and Martins, N. (2006). Efficient computation of multivariable transfer function dominant poles using subspace acceleration. *IEEE Transactions on Power Systems*, 21(4):1471–1483. 35

Rommes, J. and Martins, N. (2008). Computing transfer function dominant poles of large-scale second-order dynamical systems. *SIAM Journal on Scientific Computing*, 30(4):2137–2157. 35

Rommes, J. and Sleijpen, G. L. (2008). Convergence of the dominant pole algorithm and rayleigh quotient iteration. *SIAM Journal on Matrix Analysis and Applications*, 30(1):346–363. 35

Saad, Y. (2003). *Iterative methods for sparse linear systems*. SIAM, Philadelphia. 34, 43

Sandberg, H. (2006). A case study in model reduction of linear time-varying systems. *Automatica*, 42(3):467–472. 40

Sandberg, H. and Rantzer, A. (2004). Balanced truncation of linear time-varying systems. *IEEE Transactions on Automatic Control*, 49(2):217–229. 40

Sasane, A. (2002). *Hankel norm approximation for infinite-dimensional systems*, volume 277. Springer Science & Business Media. 34

Scarciotti, G. and Astolfi, A. (2014). Model reduction by moment matching for linear time-delay systems. In *Proceedings of the IEE 19th IFAC World Congress*, volume 29. 83

Scarciotti, G. and Astolfi, A. (2016). Model reduction of neutral linear and nonlinear time-invariant time-delay systems with discrete and distributed delays. *IEEE Transactions on Automatic Control*, 61(6):1438–1451. 45

Scherpen, J. M. (1993). Balancing for nonlinear systems. *Systems & Control Letters*, 21(2):143–153. 40

Schilders, W. H., Van der Vorst, H. A., and Rommes, J. (2008). *Model order reduction: theory, research aspects and applications*, volume 13. Springer. 35

Schulze, P. and Unger, B. (2015). Data-driven interpolation of dynamical systems with delay. *Preprint-Reihe des Instituts fur Mathematik, Technische Universitat Berlin*. 45, 50, 83, 124, 175

Seuret, A. and Gouaisbaut, F. (2015). Hierarchy of LMI conditions for the stability analysis of time-delay systems. *Systems & Control Letters*, 81:1–7. 153, 160

Sinani, K. (2015). Iterative rational Krylov algorithm for unstable dynamical systems and generalized coprime factorization. Master's thesis, Virginia Tech. 148

Sipahi, S., Niculescu, S., Abdallah, C., Michiels, W., and Gu, K. (2011). Stability and stabilization of systems with time delay. *IEEE Control Systems Magazine*, 31(1):38–65. 152, 153, 160, 168

Smith, O. (1957). Closer control of loops with dead time. *Chem. Eng. Progr.*, 53(5):217–219. 27

Sontag, E. D. (2013). *Mathematical control theory: deterministic finite dimensional systems*, volume 6. Springer Science & Business Media. 16

Sootla, A. (2014). ν-gap model reduction in the frequency domain. *IEEE Transactions on Automatic Control*, 59(1):228–233. 57

Sorensen, D. C. and Antoulas, A. (2002). The Sylvester equation and approximate balanced reduction. *Linear Algebra and its Applications*, 351:671–700. 26

Spanos, J. T., Milman, M. H., and Mingori, D. L. (1992). A new algorithm for l_2 optimal model reduction. *Automatica*, 28(5):897–909. 58

Stykel, T. (2004). Gramian-based model reduction for descriptor systems. *Mathematics of Control, Signals and Systems*, 16(4):297–319. 40

Stykel, T. (2006). Balanced truncation model reduction for semidiscretized Stokes equation. *Linear Algebra and its Applications*, 415(2):262–289. 40

Tadmor, G. (1997). Robust control in the gap: a state-space solution in the presence of a single input delay. *IEEE Transactions on Automatic Control*, 42(9):1330–1335. 27

Therapos, C. (1989). Balancing transformations for unstable nonminimal linear systems. *IEEE Transactions on Automatic Control*, 34(4):455–457. 40

Van Dooren, P., Gallivan, K. A., and Absil, P.-A. (2008). $\mathcal{H}_2$-optimal model reduction of MIMO systems. *Applied Mathematics Letters*, 21(12):1267–1273. 58, 101, 111, 114, 119

Van Dooren, P., Gallivan, K. A., and Absil, P.-A. (2010). $\mathcal{H}_2$-optimal model reduction with higher-order poles. *SIAM Journal on Matrix Analysis and Applications*, 31(5):2738–2753. 58

Vandendorpe, A. (2004). *Model reduction of linear systems, an interpolation point of view*. PhD thesis, Universite Catholique De Louvain Faculte des sciences appliquees Departement d'ingenierie mathematique. 43

Villemagne, C. d. and Skelton, R. E. (1987). Model reductions using a projection formulation. *International Journal of Control*, 46(6):2141–2169. 34, 42

Vinnicombe, G. (1993). Frequency domain uncertainty and the graph topology. *IEEE Transactions on Automatic Control*, 38(9):1371–1383. 57

Vuillemin, P. (2014). *Frequency-limited model approximation of large-scale dynamical models*. PhD thesis, Université de Toulouse. 23, 33, 35, 37, 39, 69

Vuillemin, P., Demourant, F., Biannic, J. M., and Poussot-Vassal, C. (2016). Stability analysis of a set of uncertain large-scale dynamical models with saturations: Application to an aircraft system. *IEEE Transactions on Control Systems Technology*, PP(99):1–8. 57

Vuillemin, P., Poussot-Vassal, C., and Alazard, D. (2014a). Poles Residues Descent Algorithm for Optimal Frequency-Limited $\mathcal{H}_2$ Model Approximation. In *Proceedings of the European Control Conference*, pages 1080–1085. 35, 58

Vuillemin, P., Poussot-Vassal, C., and Alazard, D. (2014b). Spectral expression for the frequency-limited $\mathcal{H}_2$-norm of LTI dynamical systems with high order poles. In *Proceedings of the IEEE European Control Conference*, pages 55–60. 23

Walsh, J. L. (1932). On interpolation and approximation by rational functions with preassigned poles. *Transactions of the American Mathematical Society*, 34(1):22–74. 69

Watanabe, K. and Ito, M. (1981). A process-model control for linear systems with delay. *IEEE Transactions on Automatic Control*, 26(6):1261–1269. 27

Wilson, D. A. (1970). Optimum solution of model-reduction problem. In *Proceedings of the Institution of Electrical Engineers*, volume 117, pages 1161–1165. 58

Wolf, T. (2014). *$\mathcal{H}_2$ pseudo-optimal model order reduction*. PhD thesis, Universität München. 69

Wolf, T., Panzer, H. K., and Lohmann, B. (2013). $\mathcal{H}_2$ pseudo-optimality in model order reduction by Krylov subspace methods. In *Proceedings of the IEE European Control Conference*, pages 62–77. 69

Wu, Z. and Michiels, W. (2012). Reliably computing all characteristic roots of delay differential equations in a given right half plane using a spectral method. *Journal of Computational and Applied Mathematics*, 236(9):2499–2514. 139

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. 58

Yi, S. (2009). *Time-delay systems: Analysis and control using the Lambert W function*. PhD thesis, The University of Michigan. 127

Zhou, K. and Doyle, J. C. (1998). *Essentials of robust control*, volume 180. Prentice-Hall Upper Saddle River, NJ. 56

Zhou, K., Salomon, G., and Wu, E. (1999). Balanced realization and model reduction for unstable systems. *International Journal of Robust and Nonlinear Control*, 9(3):183–198. 40