Approximation, Analysis and Control of Large-Scale Systems

Theory and Applications

by
Giordano Scarciotti

A Thesis submitted in fulfillment of requirements for the degree of
Doctor of Philosophy

Control and Power Research Group
Department of Electrical and Electronic Engineering
Imperial College London
2016

Abstract

This work presents some contributions to the fields of approximation, analysis and control of large-scale systems. Consequently the Thesis consists of three parts. The first part covers approximation topics and includes several contributions to the area of model reduction. Firstly, model reduction by moment matching for linear and nonlinear time-delay systems, including neutral differential time-delay systems with discrete-delays and distributed delays, is considered. Secondly, a theoretical framework and a collection of techniques to obtain reduced order models by moment matching from input/output data for linear (time-delay) systems and nonlinear (time-delay) systems is presented. The theory developed is then validated with the introduction and use of a low complexity algorithm for the fast estimation of the moments of the NETS-NYPS benchmark interconnected power system. Then, the model reduction problem is solved when the class of input signals generated by a linear exogenous system which does not have an implicit (differential) form is considered. The work regarding the topic of approximation is concluded with a chapter covering the problem of model reduction for linear singular systems. The second part of the Thesis, which concerns the area of analysis, consists of two very different contributions. The first proposes a new "discontinuous phasor transform" which allows to analyze in closed-form the steady-state behavior of discontinuous power electronic devices. The second presents in a unified framework a class of theorems inspired by the Krasovskii-LaSalle invariance principle for the study of "lim inf" convergence properties of solutions of dynamical systems. Finally, in the last part of the Thesis the problem of finite-horizon optimal control with input constraints is studied and a methodology to compute approximate solutions of the resulting partial differential equation is proposed.

Declaration of Originality

I hereby declare that the work presented in this thesis is my own unless otherwise stated. To the best of my knowledge the work is original and ideas developed in collaboration with others have been appropriately referenced.

Copyright Declaration

Acknowledgments

The difficulty of a PhD is not in the study or work that it requires, but in the psychological burden created by the continuous uncertainty of the research investigation. Fortunately, during my PhD I had the luck of having around exceptional people, who supported me and alleviated the burden of this experience. My deepest gratitude goes to Alessandro Astolfi for an uncountable number of reasons. For sure being supervised by someone who is an exceptionally clever and talented researcher helped my PhD. But Alessandro, in addition to this, is also an extraordinary human being. It is this latter quality that made my PhD a very enjoyable experience (which is a quite rare occurrence). I owe my gratitude to Laurent Praly, first for accepting to work with me and then for the week I spent in Fontainebleau where he crashed my ego in less than 10 minutes, and to Zhong-Ping Jiang, for the time spent at New York University. Although I am still waiting for the final decision, I wish to thank all the persons who believed in me and directly contributed to my preparation for the JRF: Alessandro Astolfi, Paul Mitcheson, Thomas Parisini, Richard Syms, Oleksiy Sydoruk, Eric Yeatman, Laurent Praly, Ron Hui, Hubregt Visser, Thulasi Mylvaganam and Alberto Padoan. I wish to thank my two examiners, David Angeli and Malcolm Smith, for the time spent reading my thesis and for the interesting discussions during my *viva voce*. I wish to express my gratitude to all the persons who loved me and supported me during these years. Although things did not work out, I owe my gratitude to Valeria who followed me to UK where she has been an invaluable support for me. I wish to thank the group of PhD students and researchers that I met when I started my PhD and that readily accepted me: Adriano, Andrea, Antonio, Inma, Jaehwa, Jose, Michele, Thulasi; and the people who joined along the way: Alberto and Paolo. Among these, I want to make a special mention to a few people. I wish to thank Michele and his “cover” Emma,

my perfect flatmate, for their friendship and help in difficult times. Thanks to Tessa who welcomed me in her house and made me feel home. I wish to thank my old friends Daniela, Emanuele, Francesco, Giacomo, Luca and Martina for being always there. Thanks to Max, first as my GTA coordinator, then as an extremely witty friend. Thanks to Alberto for our climbing adventures. A special thanks goes to Thulasi for the mind-blowing year we had together and for her irreplaceable help and contagious energy. I owe sincere gratitude to my family for their support and love even though they are still quite puzzled regarding what their grumpy son is doing.

Contents

Chapter 1

Introduction

1.1 Motivation and objectives

In this section we motivate the work of the Thesis and set its objectives. We begin giving an intuitive explanation of some basic terminology and concepts that are used through the Thesis.

1.1.1 Representing the evolution of reality

Evolving phenomena are characterized by quantities that change with time. Since the variation of certain quantities depends on the variation of other quantities, we can distinguish between *outputs* (the first quantities) and *inputs* (the second quantities). With the expression "*dynamical system*" we refer to the mathematical relation that describes the effect of the inputs on the outputs when all the other unrelated quantities are neglected. A very simple example of dynamical system is represented by the equation

$$M\ddot{x}(t) = u(t) + Mg + Kx(t) - F\dot{x}(t). \tag{1.1}$$

This equation represents the dynamics of a mass-spring-damper system in which M is the mass, K is the elastic constant, F in the friction constant and g is the gravity. The term u represents an input to the system that can be freely chosen and for this reason it is called control variable or, more simply, *control*. The quantity Mg represents an external input that cannot be chosen and therefore is referred to as a *disturbance*. In system (1.1) it can be noticed that the variable x, along with its time derivatives, can be used to represent

the time history of the system from the initial time t_0 until time t. This variable x is called state at time t, or simply *state* of the system.

Dynamical systems can be represented by several different classes of mathematical relations. If the time variable t takes real values, then the resulting system is a *continuous-time system*, *e.g.* system (1.1). On the other hand if the time variable takes integer values, then the system is a *discrete-time system*. Both continuous and discrete systems are powerful mathematical representations of evolving phenomena and both of them have interesting and characteristic properties. The reason one representation may be preferred over the other depends on the particular phenomenon studied. For instance, it is far more intuitive to represent the dynamics of digital devices, *e.g.* microcontrollers, with discrete systems since the time variable is related to the period of the clock inside each of these devices and thus these devices are themselves discrete-time by nature.

Another important property which is used to distinguish between classes of dynamical systems is linearity and nonlinearity. Again, the same phenomenon can be represented by a linear or a nonlinear system; however, these two representations are not equivalent. A *linear system* is usually represented by the equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t), \end{aligned} \tag{1.2}$$

where A, B and C are constant matrices and y is the output of the system. As a contrast, *nonlinear systems* are represented by means of nonlinear differential equations of the form

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)), \\ y(t) &= h(x(t)), \end{aligned} \tag{1.3}$$

where f and h are nonlinear mappings. Linear systems have the important advantage that they are mathematically "tractable". A large amount of results are known for these systems and strong claims can be proved easily. However, they have the disadvantage that they are very rarely, if ever, faithful representations of the original phenomenon since most real-life dynamics are nonlinear. Thus, the results that can be achieved with linear analysis often do not describe accurately the behavior of reality. On the other hand,

results developed for nonlinear systems, which provide a more faithful representation of the actual phenomenon and show a richer dynamics, are in comparison weak.

It is timely now to explain the meaning of the words "richer dynamics" of nonlinear systems with respect to linear systems. A first notable difference concerns the properties of the equilibrium points and the presence of the finite escape time phenomenon. *Equilibrium points* are the solutions of the differential equations (1.2) or (1.3), when $\dot{x}$ is equal to zero and u is constant and represent those states of the system from which the system does not depart unless a variation of the state, however little, happens. It is straightforward to see that equation (1.2) can have only one isolated equilibrium point because this is the solution of a linear equation. On the other hand, since the mapping f in (1.3) can be a high order algebraic equation or a transcendental equation, nonlinear systems can have *multiple isolated equilibrium points* and, as a consequence, the asymptotic behavior of the trajectories may vary depending on the initial state of the system. It is easy to show that if there is an unstable equilibrium point, namely an equilibrium point for which any variation of the state causes a departure of the state of the system from that point, the trajectories of (1.2) can only (and will) approach infinity asymptotically as time goes to infinity, because the solution of the differential equation (1.2) is governed by exponential functions. On the other hand, nonlinear systems may have trajectories that approach infinity in a finite time interval, amid several other different possible behaviors, a property which is referred to as *finite escape time* [1].

A second notable difference regards oscillation and periodic behaviors. A linear system can oscillate if and only if there are pairs of purely imaginary poles (a subset of the eigenvalues of A). This is a condition that is extremely sensitive to perturbations of the model of the system. On the contrary nonlinear systems can generate oscillations that are robust to variations of the model. In addition, the amplitude of the oscillations of linear systems depends upon the initial conditions. As a contrast, the amplitude and the period of the oscillations generated by nonlinear systems may be independent from the initial conditions. This kind of oscillations is called *limit cycle*.

The last remarkable difference shown by nonlinear systems is *chaos* and *multiple modes of behavior*. The trajectories of (1.2) can only approach an equilibrium point, a periodic orbit

or an almost-periodic orbit. The dynamics of (1.3) can be so rich that they may exhibit chaotic behavior. As a consequence, the value of the state at time T may be substantially different if an infinitesimal variation of the initial condition is considered. An example of chaotic dynamical system is the Lorenz system that can generate a *strange attractor* [2]. In addition to chaos, nonlinear systems can change mode of behavior for little variations of the input u, whereas linear systems change behavior smoothly with respect to u. Thus equilibrium points, limit cycles and strange attractors may disappear, or appear, varying the input of the nonlinear system.

Linear and nonlinear systems are not the only two possible representations of the evolution of phenomena. Another distinction that can be made is between time-varying and time-invariant systems. As it can be noted by equation (1.2), the mathematical law that expresses the dynamics of the state x does not depend on the time. Of course, the state x and the control u are time dependent, but the mathematical law is fixed. This is an example of a *time-invariant system*, *i.e.* the dynamics of the system does not depend on time. Replacing equation (1.2) with

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t), \\ y(t) &= C(t)x(t), \end{aligned} \tag{1.4}$$

we obtain a rather general representation, known as *time-varying system*, in which the matrices A, B and C change with time. The relation in this case is still linear, but the form (1.4) is more general than the ones expressed by (1.2) since it represents a dynamical system in which the parameters that are present in the mathematical equations are time-varying.

A last class of systems that is relevant for the Thesis is the class of *time-delay systems*. A very simple linear time-delay system can be represented as

$$\dot{x}(t) = Ax(t) + A_1x(t-\tau) + Bu(t), \tag{1.5}$$

in which τ is called *delay*. This class of systems represents the idea that the evolution of the phenomenon at time t does not depend only on the present state and input, but

depends also on the past value, delayed by τ, of the state.
Combinations of these classes of systems can be obtained resulting in different level of generality of the phenomena which they represent and difficulty of analysis and control.

1.1.2 Objectives of the Thesis

The objective of the Thesis is to shed light on fundamental problems regarding the study of dynamical systems, namely modeling, analysis and control. In fact, almost any problem concerning dynamical systems can be grouped in one of these three categories. *Modeling* is the problem of the selection of the mathematical representation of a phenomenon and of the determination of the parameters of this representation. *Analysis* is the problem of studying the behavior of dynamical systems, *e.g.* determining the equilibrium points, stability or instability, transient behavior and asymptotic behavior. *Control* is the problem of changing the current behavior of the system to obtain an alternative behavior.
Whichever representation we are using, and whatever the objective, the problem becomes more complicated when *large-scale systems* are considered. Large-scale systems are dynamical systems composed by a large number of differential equations and may results, for instance, from the interconnection of several simple subsystems. When considering large-scale systems another problem becomes fundamental: the *approximation* of the large-scale system with a lower dimensional representation.
The Thesis explores the topics of approximation, analysis and control of large-scale dynamical systems proposing novel solutions to some of the many open problems in this wide area of research. In particular the topics covered consist of the model reduction for the classes of systems we have introduced (and a few more), the steady-state analysis of switching circuits, the description of the asymptotic behavior of nonlinear systems and the problem of optimal control.

1.2 Contribution of the Thesis

The Thesis is organized in three parts, each divided in several chapters. Every chapter has its own introduction, presentation of the problem, theoretical development and examples. The examples are always organized in a subsection titled “Example: ...”.

The only exception is Chapter 5 which is an "application" chapter. This chapter builds upon the theoretical results of the previous chapters and presents the validation of these results from an applicative point of view, *i.e.* the practical problem takes the majority of the attention over the mathematical technicalities. A brief description of each part and chapter follows.

Part I is the most substantial part of the Thesis, consists of six chapters and covers the topic of approximation.

Chapter 2 provides an introduction to the problem of model reduction by moment matching. We introduce the interpolation approach to moment matching, which is how moment matching has been classically interpreted in the linear framework. Then, we move to the steady-state approach introduced in [3]. Exploiting this description of moment, we give the solution of the problem of model reduction by moment matching achieving additional properties for linear and nonlinear systems. These basic results are instrumental for the development of the remaining of Part I.

Chapter 3 deals with the problem of model reduction for linear and nonlinear time-delay systems. The notion of moment is extended to linear time-delay systems, the solvability of the resulting Sylvester-like equation is discussed and the equivalent steady-state description is given. These results are subsequently extended to neutral systems and systems with distributed-delays. A family of systems achieving moment matching is presented and a simple example based on the model of a LC transmission line is given. The possibility of interpolating a larger number of points maintaining the same "number of equations" is investigated. Two examples, one inspired by [4] and one inspired by the problem of the automatic control of a platoon of vehicles, illustrate the results. Then, the center manifold theory for time-delay systems is used to extend the definition of moment to nonlinear time-delay systems and a family of systems achieving moment matching for nonlinear time-delay systems is given. The possibility of interpolating multiple moments is investigated and the problem of obtaining a reduced order model of an unstable system

is discussed. The theory is illustrated by the reduction of a nonlinear neutral time-delay equation describing the torsional dynamics of an oilwell drillstring. Finally, the characterization of the moments at infinity for linear and nonlinear time-delay systems is presented.

Chapter 4 presents a theoretical framework and a collection of techniques to obtain reduced order models by moment matching from input/output data for linear (time-delay) systems and nonlinear (time-delay) systems. We begin by giving a preliminary analysis to compute on-line estimates of the moments of a linear system. Then approximations which converge asymptotically to the moments of the linear (nonlinear) system are given. A discussion on the computational complexity associated with the evaluation of these approximations is presented, a recursive least-square formula is given and a moment estimation algorithm is provided. A simple example in which the average heartbeat under "stress" is measured is used to show how to estimate the moment of a nonlinear system. Then, a family of reduced order models for linear (nonlinear) systems and linear (nonlinear) time-delay systems, respectively, is given. We discuss how several properties, *e.g.* matching with prescribed eigenvalues or zeros, can be enforced in the present scenario. A linear reduced order model computed with the method proposed in the chapter is estimated for a system of order $n = 1010$ [4, 5]. A nonlinear reduced order model constructed using an approximation of the moment of the DC-to-DC Ćuk converter provides a further example.

Chapter 5 provides a validation of the theory developed in Chapters 2 and 4. We present a low complexity algorithm for the fast estimation of the moments of MIMO systems. The estimated moments are exploited for the model reduction of large-scale interconnected power systems. The technique offers, simultaneously, a low computational complexity approximation of the moments and the possibility to easily enforce constraints on the reduced order model. This possibility is used to preserve selected slow and poorly damped modes which are important both from a mathematical and physical point of view. The problem of the choice of the so-called tangential directions is also studied and an heuristic for their approximation is given. The techniques have been validated with

the study of the dynamic response of the NETS-NYPS benchmark system.

Chapter 6 investigates the limitations of the description of moment based on a signal generator described by differential equations. With the final aim of solving the model reduction problem for a class of input signals generated by a linear exogenous system which does not have an implicit (differential) form, a time-varying parametrization of the steady-state of the system is used to extend, exploiting an integral matrix equation, the definition of moment to this class of input signals. The equivalence of the new definition and the one based on the Sylvester equation is proved under specific conditions. Special attention is given to periodic signals due to the wide range of practical applications where these are used. Reduced order models matching the steady-state response of explicit signal generators are given for linear systems and several connections with the classical reduced order models are drawn.

Chapter 7 extends the model reduction techniques to linear singular systems. Combining the interpolation-based and the steady-state-based description of moment a partitioned projector is constructed. The contribution of the slow subsystem and the contribution of the fast subsystem to the moment are separated. The information on the fast subsystem is encoded in the projector and it is not lost by the moment matching technique. Moreover, the output of reduced order models based on this projector approximates the output of the system only when consistent initial conditions are taken into account. Exploiting this partitioned projector, several families of reduced order models are proposed. In particular, a purely fast, a purely slow and a "simple" family of reduced order models are given. The possibility of maintaining the impulsive controllability property is investigated and a few examples are used to illustrate the results.

Part II is another significant part of the Thesis, consists of two chapters and covers analysis topics.

Chapter 8 shows the somewhat obvious fact that the phasors of an electric circuit are the moments on the imaginary axis of the linear system describing the

circuit. Exploiting this relation, we can analyze circuits powered by discontinuous sources. A new "discontinuous phasor transform", which allows to analyze in closed-form the steady-state behavior of discontinuous power electronic devices, is defined and the v-i characteristics for inductors, capacitors and resistors is described in terms of this new phasor transform. The new quantities maintain their physical meaning: the instantaneous power, the average power and the reactive power in the phasor domain are defined. The analytic potential of the new tool is illustrated studying the steady-state response of power inverters and of wireless power transfer systems with non-ideal switches.

Chapter 9 presents in a unified framework a class of theorems inspired by the Krasovskii-LaSalle invariance principle. The contribution of the chapter is a tool to study "$\liminf$" convergence properties of solutions of dynamical systems. In particular the theorems give sufficient conditions to determine the convergence in the mean and the "$\liminf$" convergence. These theorems are derived by a relaxation of Matrosov and Small-gain Theorems, and they are based on a "$\liminf$" Barbalat's Lemma. Additional technical assumptions to have "lim" convergence are given. The "$\liminf$" / "lim" relation and the role of some of the assumptions are illustrated by means of examples.

Part III is the shortest part of the Thesis, consists of one chapter and covers a control topic.

Chapter 10 studies the finite-horizon optimal control problem with input constraints, which consists of controlling the state of a dynamical system over a finite time interval (possibly unknown) minimizing a cost functional, while satisfying hard constraints on the input. In this framework, the minimum time optimal control problem and some related problems are of interest for both theory and applications. For linear systems the solution of the problem often relies upon the use of bang-bang control signals. For nonlinear systems the "shape" of the optimal input is in general not known. The control input can be found solving a Hamilton-Jacobi-Bellman (HJB) partial differential equation (pde): it typically consists of a combination of bang-bang controls and singular arcs. In

this chapter a methodology to approximate the solution of the HJB pde is proposed. This approximation yields a dynamic state feedback law. The theory is illustrated by means of two examples: the minimum time optimal control problem for an industrial wastewater treatment plant and the Goddard problem, *i.e.* a maximum range optimal control problem.

Finally, the Thesis is concluded with a summary of the work and directions for future work.

1.3 Notation

Standard notation has been adopted in the Thesis, most of which is defined in this section and used throughout the remainder of the Thesis. When new notation, not included in this section is introduced, this is defined in the relevant parts of the Thesis.

The symbol $\mathbb{R}_{\geq 0}$ ($\mathbb{R}_{>0}$) denotes the set of non-negative (positive) real numbers; $\mathbb{C}_{<0}$ denotes the set of complex numbers with strictly negative real part; $\mathbb{C}_0$ denotes the set of complex numbers with zero real part and $\mathbb{D}_{<1}$ the set of complex numbers with modulo less than one.

The symbol I denotes the identity matrix and $\sigma(A)$ denotes the spectrum of the matrix $A \in \mathbb{R}^{n\times n}$. The symbol $\otimes$ indicates the Kronecker product and $||A||$ indicates the induced Euclidean matrix norm. Given a list of n elements a_i, $\text{diag}(a_i)$ indicates a diagonal matrix with diagonal elements equal to the a_i's. The vectorization of a matrix $A \in \mathbb{R}^{n\times m}$, denoted by $\text{vec}(A)$, is the $nm \times 1$ vector obtained by stacking the columns of the matrix A one on top of the other, namely $\text{vec}(A) = [a_1^\top, a_2^\top, \ldots, a_m^\top]^\top$, where $a_i \in \mathbb{R}^n$ are the columns of A and the superscript $\top$ denotes the transposition operator. The superscript $*$ indicates the conjugate transpose operator.

The symbol $\Re[z]$ indicates the real part of the complex number z, $\Im[z]$ denotes its imaginary part and ι denotes the imaginary unit. The symbol ϵ_k indicates a vector with the k-th element equal to 1 and with all the other elements equal to 0. Given a function f, $\overline{F}$ represents its phasor at ω, whereas $< f(t) >$ indicates its time average.

Given a set of delays $\{\tau_j\}$, the symbol $\mathfrak{R}^n_T = \mathfrak{R}^n_T([-T, 0], \mathbb{R}^n)$, with $T = \max_j\{\tau_j\}$, indicates the set of continuous functions mapping the interval $[-T, 0]$ into $\mathbb{R}^n$ with the

topology of uniform convergence [6]. The subscripts "τ_j" and "χ_j" denote the translation operator, *e.g.* $x_{\tau_j}(t) = x(t - \tau_j)$.

Let $\bar{s} \in \mathbb{C}$ and $A(s) \in \mathbb{C}^{n\times n}$. Then $\bar{s} \notin \sigma(A(s))$ means that $\det(\bar{s}I - A(\bar{s})) \neq 0$. $\sigma(A(s)) \subset \mathbb{C}_{<0}$ means that for all $\bar{s}$ such that $\det(\bar{s}I - A(\bar{s})) = 0$, $\bar{s} \in \mathbb{C}_{<0}$.

The symbol $\mathcal{L}(f(t))$ denotes the Laplace transform of the function $f(t)$ (provided that $f(t)$ is Laplace transformable) and $\mathcal{L}^{-1}\{F(s)\}$ denotes the inverse Laplace transform of $F(s)$ (provided it exists). With some abuse of notation, $\sigma(\mathcal{L}(f(t)))$ denotes the set of poles of $\mathcal{L}(f(t))$. Given two functions, $f : Y \to Z$ and $g : X \to Y$, with $f \circ g : X \to Z$ we denote the composite function $(f \circ g)(x) = f(g(x))$ which maps all $x \in X$ to $f(g(x)) \in Z$.

1.4 Published material

The introduction to the problem of model reduction by moment matching given in Chapter 2 is partly contained in the book chapter [7]. The problem of the model reduction of time-delay systems covered in Chapter 3 has been published in the conference papers [8,9], in the journal paper [10] and is partly included in [7]. The topic of data-driven model reduction discussed in Chapter 4 is contained in the conference papers [11, 12] and in the journal paper [13]. The conference paper [14] and the journal paper [15] encompass the application of model reduction to power systems given in Chapter 5. Chapter 6, which deals with the problem of model reduction for explicit signal generators, has been published in the conference paper [16] and in the journal paper [17]. The topic in Chapter 7 is included in the conference paper [18]. The theory regarding the discontinuous phasor transform and its applications in power electronics presented in Chapter 8 are contained in the conference paper [19] and in the journal paper [20]. The theorems for the study of weak converge properties given in Chapter 9 are published in the conference paper [21] and in the journal paper [22]. The work reported in Chapter 10 on the approximate finite-horizon optimal control with input constraints has been published in the conference paper [23] and in the journal paper [24]. Finally, note that a few contributions of the author, submitted soon after the writing of the draft of the Thesis, have not been included [25–27].

Part I

Approximation

Chapter 2

Model reduction by moment matching

2.1 Introduction

The availability of mathematical models is essential for the analysis, control and design of modern technological devices. As the computational power has advanced, the complexity of these mathematical descriptions has increased. This has maintained the computational needs at the top or over the available possibilities [28]. A solution to this problem is represented by the use of reduced order models, which are exploited in the prediction, analysis and control of a wide class of physical behaviors. For instance, reduced order models are used to simulate or design weather forecast models, very large scale integrated circuits or networked dynamical systems [5]. The model reduction problem can be informally formulated as the problem of finding a simplified description of a dynamical system in specific operating conditions, preserving at the same time specific properties, *e.g.* stability. For linear systems, the problem has been addressed from several perspectives which can be divided into two main groups: singular value decomposition (SVD)-based approximation methods and moment matching, or Krylov-based, approximation methods. The use of the notion of Hankel operators [29–31] and the theory of balanced realizations [32–35], belong to the first group, whereas the use of interpolation theory and the notion of projections [36–44] belong to the latter. For further detail and an extensive list of references see the monograph [5]. The additional difficulties of the reduction of nonlinear systems carry

the need to develop different or "enhanced" techniques. The problem of model reduction for special classes of systems, such as differential-algebraic systems, bilinear systems and mechanical/Hamiltonian systems has been studied in [45–48]. Energy-based methods have been proposed in [34,49,50]. Other techniques, based on the reduction around a limit cycle or a manifold, have been presented in [51, 52]. Model reduction methods based on proper orthogonal decomposition have been developed for linear and nonlinear systems, see *e.g.* [53–57]. Finally, note that some computational aspects have been investigated in [52, 55, 58, 59].

The goal of this chapter is to review the model reduction techniques for linear and nonlinear systems based on the "steady-state" notion of moment introduced in [60] and [3] exploiting the center manifold theory. The advantage of moment matching over the SVD-based methods is that the numerical implementation is much more efficient [5, Section 14.1]. The major drawback, however, is the difficulty in the moment matching methods to preserve important properties of the original system. The first outcome of the "steady-state" notion of moment is the ability to preserve some properties of the original system overcoming some of the drawbacks of the moment matching methods. The second outcome is the extension of the model reduction theory by moment matching to nonlinear systems. This has led to new results in the area of model reduction, see *e.g.* [61–63].

The rest of the chapter is organized as follows. In Section 2.2 we introduce the interpolation approach to moment matching, which is how moment matching has been classically interpreted in the linear framework. In Section 2.3 we move to the steady-state approach introduced in [3]. In Section 2.4 we provide families of linear reduced order models and we study the possibility to achieve moment matching with additional properties. In Section 2.5 we present some results on the model reduction problem by moment matching for nonlinear systems.

The material presented in this chapter is not original work of the author but it is due to [3, 60]. It is herein reported because it is instrumental for the development of the rest of the Approximation Part of the Thesis. Note finally that part of this chapter has been published in the book chapter [7] as a review report on model reduction by moment matching.

2.2 The interpolation approach

In this section we briefly recall the notion of moment and the related model reduction techniques as presented in [5]. We refer to this family of methods as "interpolation-based" methods. The key element to understand this framework is that the moment matching problem is interpreted as a problem of interpolation of points on the complex plane.

Consider a linear, single-input, single-output, continuous-time, system described by the equations

$$\dot{x} = Ax + Bu, \qquad y = Cx, \tag{2.1}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$ and $C \in \mathbb{R}^{1\times n}$. Let

$$W(s) = C(sI - A)^{-1}B$$

be the associated transfer function and assume that (2.1) is minimal, *i.e.* controllable and observable. The k-moment of system (2.1) at s_i is defined as the k-th coefficient of the Laurent series expansion of the transfer function $W(s)$ in a neighborhood of $s_i \in \mathbb{C}$ (see [5, Chapter 11]), provided it exists.

Definition 1. [5] Let $s_i \in \mathbb{C} \setminus \sigma(A)$. The 0-*moment of system (2.1) at* s_i is the complex number

$$\eta_0(s_i) = C(s_iI - A)^{-1}B.$$

The k-*moment of system (2.1) at* s_i is the complex number

$$\eta_k(s_i) = \frac{(-1)^k}{k!}\left[\frac{d^k}{ds^k}(C(sI - A)^{-1}B)\right]_{s=s_i},$$

with $k \geq 1$ integer.

In the interpolation-based notion of moment a reduced order model is such that its transfer function (and derivatives of this) takes the same values of the transfer function (and derivatives of this) of system (2.1) at s_i. This is graphically represented in Fig. 2.1 in which the magnitude (top) and phase (bottom) of the transfer function of a reduced order model (dashed/red line) matches the respective quantities of a given system (solid/blue line) at the point $s_i = 30\imath$. Since a minimal system can be entirely described by its transfer

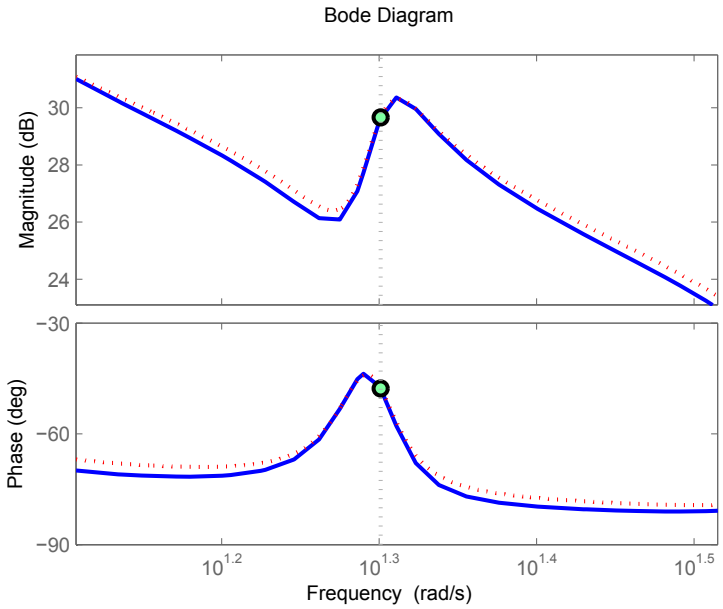

Figure 2.1: Diagrammatic illustration of the interpolation approach. Magnitude (top graph) and phase (bottom graph) plots of a given system (solid/blue line) and of a reduced order model (dashed/red line). The green circle represents the interpolation point.

function, such a system can be effectively reduced using this technique.

In this framework the problem of model reduction by moment matching can be formulated as the problem of finding the Petrov-Galerkin projectors $V \in \mathbb{R}^{n\times\nu}$ and $W \in \mathbb{R}^{n\times\nu}$, with $W^*V = I$, such that the model described by the equations

$$\dot{\xi} = F\xi + Gu, \qquad \psi = H\xi, \tag{2.2}$$

with $\xi(t) \in \mathbb{R}^\nu$, $u(t) \in \mathbb{R}$, $\psi(t) \in \mathbb{R}$, $F \in \mathbb{R}^{\nu\times\nu}$, $G \in \mathbb{R}^{\nu\times 1}$, $H \in \mathbb{R}^{1\times\nu}$ and

$$F = W^*AV, \quad G = W^*B, \quad H = CV, \tag{2.3}$$

matches the moments of the given systems at a set of points s_i. The problem of model reduction by moment matching using the Petrov-Galerking projectors is thoroughly described in [5] and it is the subject of intensive research, see *e.g.* [36–40, 42, 43]. Herein we report a few results which are instrumental for the aims of the Thesis. We invite the reader to refer to [5] for additional details.

Proposition 1. [5] Consider $s_j \in \mathbb{C} \setminus \sigma(A)$, with $j = 1, \ldots, \nu$. The transfer function of the reduced order model (2.2), with

$$V = \left[\begin{array}{ccc} (s_1 I - A)^{-1} B & \cdots & (s_\nu I - A)^{-1} B \end{array} \right] \tag{2.4}$$

a generalized reachability matrix and W any left inverse of V, interpolates the transfer function of system (2.1) at the points s_j, with $j = 1, \ldots, \nu$.

Proposition 2. [5] Consider $s_0 \in \mathbb{C} \setminus \sigma(A)$. The transfer function of the reduced order model (2.2), with

$$V = \left[\begin{array}{cccc} (s_0 I - A)^{-1} B & (s_0 I - A)^{-2} B & \cdots & (s_0 I - A)^{-\nu} B \end{array} \right] \tag{2.5}$$

a generalized reachability matrix and W any left inverse of V, interpolates the transfer function of system (2.1) and its $\nu - 1$ derivatives at the point s_0.

The techniques which result from these propositions are called *rational interpolation methods* by projection, or *Krylov methods*. We note that the matrix W is a free parameter since it has to satisfy only a "mild" constraint, namely that it is a left inverse of V. However, the selection of W such that the reduced order model exhibits specific properties is in general a difficult problem. All the results presented to exploit the free parameters of the matrix W play, with different aims, on the possibility of interpolating more, somewhat special, points. The first of these results, which we recall here, provides a method for the so-called two-sided interpolation.

Proposition 3. [5] Consider $s_j \in \mathbb{C} \setminus \sigma(A)$, with $j = 1, \ldots, 2\nu$, the generalized reachability matrix

$$\bar{V} = \left[\begin{array}{ccc} (s_1 I - A)^{-1} B & \cdots & (s_\nu I - A)^{-1} B \end{array} \right], \tag{2.6}$$

and the generalized observability matrix

$$\bar{W} = \left[\begin{array}{ccc} (s_{\nu+1} I - A^*)^{-1} C^* & \cdots & (s_{2\nu} I - A^*)^{-1} C^* \end{array} \right]. \tag{2.7}$$

Assume that $\det(\bar{W}^* \bar{V}) \neq 0$, then the transfer function of the reduced order model (2.2) with $V = \bar{V}$ and $W = \bar{W}(\bar{V}^* \bar{W})^{-1}$ interpolates the transfer function of system (2.1) at the points s_j, with $j = 1, \ldots, 2\nu$.

From a numerical point of view, the computation of the matrices V and W defined in this chapter can be performed using the two-sided Lanczos or the Arnoldi algorithms [5], which are algorithms originally devised for the iterative approximation of the eigenvalues of a matrix. We can now summarize the main advantages of the Krylov methods [5]:

- the number of operations needed to compute a reduced order model of order k given a system of order n using Lanczos or Arnolfi factorization is $\mathcal{O}(\nu n^2)$ for dense systems and $\mathcal{O}(\nu^2 n)$ for sparse systems;
- only matrix-vector multiplications are required, *i.e.* no matrix factorizations or inversions are needed;
- Krylov methods can also be applied to multi-input, multi-output systems;
- Krylov methods have simple algorithms and high convergence rate.

However, the drawbacks of the Krylov methods are [5]:

- the reduced order model may be unstable even though the original system is stable;
- the algorithms break down if during the iteration some rank conditions (see [5, Pag. 350]) are not satisfied;
- there is no systematic technique to preserve important properties of the system, for instance maintaining prescribed eigenvalues, relative degree, zeros, L_2-gain, or preserving compartmental constraints;
- the interpolation-based methods cannot be applied to nonlinear systems (or more general classes of systems), since for these we cannot define a transfer function.

Note that a solution to the problem of preservation of passivity and stability has been proposed in [64, 65], as reported here.

Lemma 1. [5] If the interpolation points in Proposition 3 are chosen so that s_j, with $j = 1, \ldots, \nu$, are stable spectral zeros, *i.e.* they are such that $W^*(-s_j) + W(s_j) = 0$, and $s_{j+\nu} = -s_j$, with $j = 1, \ldots, \nu$, *i.e.* the interpolation points are chosen as zeros of the spectral factors and their mirror images, the projected system is both stable and passive.

However, this lemma requires that specific moments are matched. Hence, the designer cannot choose arbitrary moments which, as we see later, should be in general considered a drawback in the moment matching approach. Moreover, there is a lack of system theoretic understanding behind why a particular interpolation point is related to a property like passivity. Finally, in Lemma 1 all the free parameters (the matrix W) are used and no additional property can be preserved.

A possible solution to these issues is offered by the "steady-state-based" approach to moment matching, which is introduced in the next section.

2.3 The steady-state approach

As just observed the interpolation approach cannot be extended to nonlinear systems for which the idea of interpolating points in the complex plane partially loses its meaning (see, however, [66] and [67] for some results on the interpolation problem for nonlinear systems). In [3] (see also [41], [42]) a characterization of moment for system (2.1) has been given in terms of the solution of a Sylvester equation as follows.

Lemma 2. [3] Consider system (2.1), $s_i \in \mathbb{C} \setminus \sigma(A)$, for all $i = 1, \ldots, \eta$. There exists a one-to-one[1] relation between the moments $\eta_0(s_1), \ldots, \eta_{k_1-1}(s_1), \ldots, \eta_0(s_\eta), \ldots, \eta_{k_\eta-1}(s_\eta)$ and the matrix $C\Pi$, where Π is the unique solution of the Sylvester equation

$$A\Pi + BL = \Pi S, \tag{2.8}$$

with $S \in \mathbb{R}^{\nu\times\nu}$ any non-derogatory[2] matrix with characteristic polynomial

$$p(s) = \prod_{i=1}^{\eta}(s - s_i)^{k_i}, \tag{2.9}$$

where $\nu = \sum_{i=1}^{\eta} k_i$, and L is such that the pair (L, S) is observable.

The importance of this formulation, which resulted in several developments in the area of model reduction by moment matching, see *e.g.* [61, 63] and [10–12, 16, 17], is that

[1] The matrices A, B, C and the zeros of (2.9) fix the moments. Then, given any observable pair (L, S) with S a non-derogatory matrix with characteristic polynomial (2.9), there exists an invertible matrix $T \in \mathbb{R}^{\nu\times\nu}$ such that the elements of the vector $C\Pi T^{-1}$ are equal to the moments.

[2] A matrix is non-derogatory if its characteristic and minimal polynomials coincide.

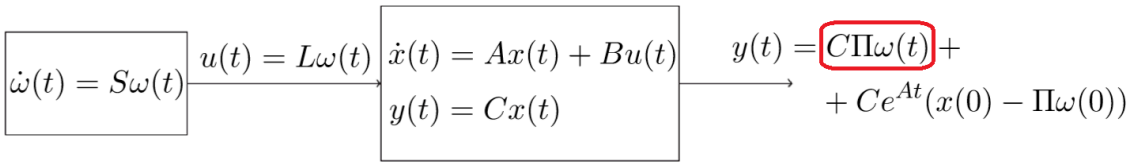

Figure 2.2: Diagrammatic illustration of Theorem 1. The term denoting the steady-state response is circled.

it establishes, through the Sylvester equation (2.8), a relation between the moments and the steady-state response of the output of the system. This is summarized in the following result and illustrated in Fig. 2.2.

Theorem 1. [3] Consider system (2.1), $s_i \in \mathbb{C}\backslash\sigma(A)$, for all $i = 1, \ldots, \eta$, and $\sigma(A) \subset \mathbb{C}_{<0}$. Let $S \in \mathbb{R}^{\nu\times\nu}$ be any non-derogatory matrix with characteristic polynomial (2.9). Consider the interconnection of system (2.1) with the system

$$\dot{\omega} = S\omega, \qquad u = L\omega, \tag{2.10}$$

with L and $\omega(0)$ such that the triple $(L, S, \omega(0))$ is minimal. Then there exists a one-to-one relation between the moments $\eta_0(s_1)$, $\ldots$, $\eta_{k_1-1}(s_1)$, $\ldots$, $\eta_0(s_\eta)$, $\ldots$, $\eta_{k_\eta-1}(s_\eta)$ and the steady-state response of the output y of such interconnected system.

Remark 1. By one-to-one relation we mean that the moments are uniquely determined by the steady-state response of $y(t)$ and *vice versa*. Exploiting this fact, in Chapter 4 the problem of model reduction for unknown linear systems from input/output data is studied. Therein, an algorithm that, given the signal ω and the output y, retrieves the moments of a system for which the matrices A, B and C are not known is devised. ■

The reduction technique based on this notion of moment consists in the interpolation of the steady-state response of the output of the system: a reduced order model is such that its steady-state response is equal to the steady-state response of the output of system (2.1) (provided it exists). Thus, the problem of model reduction by moment matching has been changed from a problem of interpolation of points to a problem of interpolation of signals. The output of the reduced order model has to behave as the output of the original system for a class of input signals, a concept which can be translated to nonlinear systems, time-delay systems and infinite dimensional systems, [3, 10]. This fact

also highlights the importance for the moment matching techniques of letting the designer free to choose the interpolation points, which are related to the class of inputs to the system.

2.4 Model reduction by moment matching for linear systems

2.4.1 Moment matching

Exploiting this new notion of moment, in [3] a family of reduced order models that achieve moment matching is given. The main advantage of this family lies on the ease of choosing the free parameters to preserve additional properties.

Definition 2. [3] Consider the signal generator (2.10). The system described by the equations

$$\dot{\xi} = F\xi + Gu, \qquad \psi = H\xi, \tag{2.11}$$

with $\xi(t) \in \mathbb{R}^\nu$, $\psi(t) \in \mathbb{R}$, $F \in \mathbb{R}^{\nu\times\nu}$, $G \in \mathbb{R}^{\nu\times 1}$, $H \in \mathbb{R}^{1\times\nu}$ is a *model of system (2.1) at S* if system (2.11) has the same moments at S as (2.1).

Lemma 3. [3] Consider system (2.1) and the signal generator (2.10). Suppose $\sigma(A) \cap \sigma(S) = \emptyset$. Then the system (2.11) is a model of system (2.1) at S if there exists a unique solution P of the equation

$$FP + GL = PS, \tag{2.12}$$

such that

$$HP = C\Pi, \tag{2.13}$$

where Π is the unique solution of (2.8).

Several methods and algorithms to solve equation (2.8) are presented in [5]. However, note that the determination of the matrix Π may be computationally expensive because one of its dimensions depends on the order of the system to be reduced. An alternative approach proposed in [3] to avoid solving equation (2.8) when the solution Π has dimension $n \times \nu$ consists in using Arnoldi or Lanczos algorithms as follows. Consider system (2.1) and construct a reduced order model achieving moment matching at S with any efficient algorithm [5, 68, 69]. At this point it is sufficient to apply the steady-state-

based model reduction techniques discussed in this chapter considering the reduced order model just obtained as the system to be reduced. This time the matrix Π will have the smaller dimensions $\nu \times \nu$. Note also that the results in Chapter 4 can be used to determine a reduced order model in a even more computationally efficient way. This last strategy is exploited in Chapter 5 for the reduction of large-scale interconnected power system.
The computation of the solution P of the Sylvester equation (2.12) of the reduced order model can be avoided altogether. In fact, as shown in [3], the family of systems

$$\dot{\xi} = (S - GL)\xi + Gu, \qquad \psi = C\Pi\xi, \tag{2.14}$$

with G any matrix such that $\sigma(S) \cap \sigma(S - GL) = \emptyset$, belongs to the family (2.11) and contains all the models of dimension ν interpolating the moments of system (2.1) at S.

Remark 2. With the expression "the family (2.11) contains all the models of dimension ν" we mean that all the models that can be obtained using Krylov projectors are encoded in the family of systems (2.11) [3]. Thus the models obtained with the two approaches are equivalent. The advantage of this formulation is that the family of systems (2.11) is parametrized in G, which allows to set with ease several properties of the reduced order model, as shown in the next section. For instance, setting the eigenvalues of the reduced order model is an easy task, whereas with the classic Krylov method this is rather difficult.
■

It can be observed that the family of models (2.14) is built on three ideas: avoiding to solve equation (2.12), selecting the solution as I; copying the dynamics of the signal generator (2.10), *i.e.* the relation $\xi = \omega$ holds for the steady-state of ξ; and having a convenient parametrization for the family of reduced order models for which additional constraints can be easily imposed.

Remark 3. The state ξ of system (2.14) taken in isolation does not hold any information regarding system (2.1), but it contains only information about the signal generator (S and L) and ν free parameters (G). This suggests that any use of the reduced order model has to take into account both the state ξ and the output ψ. In fact, model reduction by moment matching consists in the problem of interpolating the steady-state *output* response of the system [16]. Even though not as evident, this holds true also for the

family of models (2.11). ■

Remark 4. [3] This approach differs from the interpolation-based approach presented in Section 2.2. In fact, the relation between the projector W and the properties of the reduced order models is nontrivial because it does not possess a system theoretic interpretation, see [43, 44, 64, 70]. On the contrary, in the approach presented in [3] and reported in the remainder of this section, the family of reduced order models achieving moment matching is parameterized explicitly by the matrix G. Consequently, procedures to construct the matrix G achieving additional properties can be easily obtained. ■

Remark 5. In addition to the Krylov interpolation theory, other connections can be established with the so-called Georgiou-Kimura parametrization [36,71,72], and the Nevanlinna-Pick interpolation problem [73]. A discussion of these connections can be found in [3]. ■

2.4.2 Model reduction by moment matching with additional properties

We recall here the solution to the problem of achieving moment matching with additional constraints on the reduced order model which is relevant for Chapters 3 and 4. The proofs are omitted and can be found in [3]. Note also that therein other problems, such as matching with a passivity constraint and matching with L_2-gain, are discussed and solved.

Matching with prescribed eigenvalues [3]

Given a set $\Lambda = \{\lambda_1, \cdots, \lambda_\nu\}$ of ν values $\lambda_i \in \mathbb{C}$, with $\Lambda \cap \sigma(S) = \emptyset$, we consider the problem of determining G such that system (2.14) has eigenvalues equal to the elements of Λ. This objective has been achieved with the interpolation-based notion of moment in [70,74]. In the steady-state-based framework this problem can be solved simply selecting G such that

$$\sigma(S - GL) = \Lambda. \tag{2.15}$$

Matching with prescribed relative degree, zeros, or a compartmental constraint

Consider system (2.14) and the problem of selecting G such that the system has a given relative degree $r \in [1, \nu]$, prescribed zeros or satisfies a compartmental constraint [75–77].

Definition 3. [3] System (2.14) has relative degree r if

$$\begin{bmatrix} C\Pi \\ \vdots \\ C\Pi S^{r-2} \\ C\Pi S^{r-1} \end{bmatrix} G = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \gamma \end{bmatrix}, \tag{2.16}$$

for some non-zero γ.

The following result holds.

Theorem 2. [3] The following statements are equivalent.

(RD) Equation (2.16) has a solution G for all $r \in [1, \nu]$.

(Z1) The zeros of system (2.14) can be arbitrarily assigned by a proper selection of G.

(Z2) The zeros of the system

$$\dot{\xi} = S\xi + Gu, \qquad \psi = C\Pi\xi, \tag{2.17}$$

can be arbitrarily assigned by a selection of G.

(C) There is a matrix G such that system (2.17) has a diagonal positive realization.

(O1) The system

$$\dot{\xi} = S\xi, \qquad \psi = C\Pi\xi, \tag{2.18}$$

is observable.

(O2) The system

$$\dot{\omega} = S\omega, \quad \dot{x} = Ax + BL\omega, \quad y = Cx, \tag{2.19}$$

is observable.

2.5 Model reduction by moment matching for nonlinear systems

2.5.1 The notion of moment

An important contribution of [3] is the extension of the model reduction techniques by moment matching to nonlinear systems. Consider a nonlinear, single-input, single-output,

continuous-time system described by the equations

$$\dot{x} = f(x, u), \qquad y = h(x), \tag{2.20}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, f and h smooth mappings, a signal generator described by the equations

$$\dot{\omega} = s(\omega), \qquad u = l(\omega), \tag{2.21}$$

with $\omega(t) \in \mathbb{R}^v$, s and l smooth mappings, and the interconnected system

$$\dot{\omega} = s(\omega), \qquad \dot{x} = f(x, l(\omega)), \qquad y = h(x). \tag{2.22}$$

In addition, suppose that $f(0, 0) = 0$, $s(0) = 0$, $l(0) = 0$ and $h(0) = 0$. Similarly to the linear case the interconnection of system (2.20) with the signal generator captures the property that we are interested in preserving the behavior of the system only for specific input signals. The following assumptions and definitions provide a generalization of the notion of moment.

Assumption 1. The signal generator (2.21) is observable, *i.e.* for any pair of initial conditions $\omega_a(0)$ and $\omega_b(0)$, such that $\omega_a(0) \neq \omega_b(0)$, the corresponding output trajectories $l(\omega_a(t))$ and $l(\omega_b(t))$ are such that $l(\omega_a(t)) - l(\omega_b(t)) \not\equiv 0$, and Poisson stable[3] with $\omega(0) \neq 0$.

Assumption 2. The zero equilibrium of the system $\dot{x} = f(x, 0)$ is locally exponentially stable.

Lemma 4. [3] Consider system (2.20) and the signal generator (2.21). Suppose Assumptions 1 and 2 hold. Then there is a unique mapping $\pi(\omega)$, locally defined in a neighborhood of $\omega = 0$, which solves the partial differential equation

$$\frac{\partial \pi}{\partial \omega} s(\omega) = f(\pi(\omega), l(\omega)). \tag{2.23}$$

Remark 6. Lemma 4 implies that the interconnected system (2.22) possesses an invariant manifold described by the equation $x = \pi(\omega)$. ■

[3] See [78, Chapter 8] for the definition of Poisson stability.

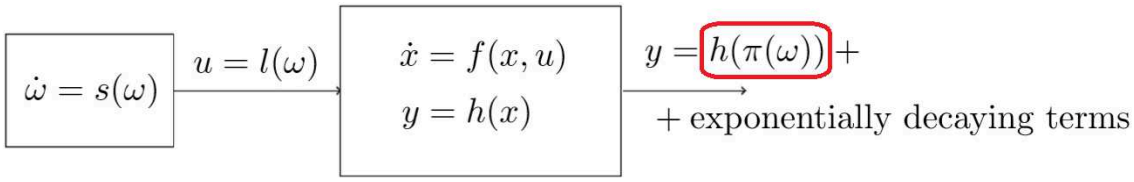

Figure 2.3: Diagrammatic illustration of Theorem 3. The term denoting the steady-state response is circled.

Definition 4. [3] Consider system (2.20) and the signal generator (2.21). Suppose Assumption 1 holds. The function $h \circ \pi$, with π solution of equation (2.23), is the *moment of system (2.20) at* (s, l).

Theorem 3. [3] Consider system (2.20) and the signal generator (2.21). Suppose Assumptions 1 and 2 hold. Then the moment of system (2.20) at (s, l) coincides with the steady-state response of the output of the interconnected system (2.22).

The result is illustrated in Fig. 2.3 which represents the nonlinear counterpart of Fig. 2.2.

Remark 7. [3] While for linear systems it is possible to define k-moments for every $s_i \in \mathbb{C}$ and for any $k \geq 0$, for nonlinear systems it may impossible to provide general statements if the input signal generated by system (2.21) is unbounded. Hence, we assume that the trajectories generated by the signal generator are bounded. ■

2.5.2 Markov parameters of a nonlinear system

In [3] the steady-state-based notion of moment has been extended also to moment at infinity. This notion is relevant for the results in Chapter 3 and thus it is herein recalled. For the linear system (2.1) the k-moments at infinity are defined as $\eta_k(\infty) = CA^{k-1}B$ (with $\eta_0(\infty) = 0$), *i.e.* the first $k+1$ moments at infinity coincide with the first $k+1$ Markov parameters [5]. Recall also that[4]

$$CA^kB = \frac{d^k}{dt^k}(Ce^{At}B)\bigg|_{t=0} = y_I^{(k)}(0^+),$$

[4]In this section we use $y^{(i)}(t)$ to denote the i-th order time derivative of y at time t.

where y_I denotes the impulse response of the system. Consider now a nonlinear affine system[5] described by equations of the form

$$\dot{x} = f(x) + g(x)u, \qquad y = h(x), \tag{2.24}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$ and f, g and h smooth mappings. The k-moment at infinity, for $k \geq 0$, of the nonlinear system (2.24) are defined as $\eta_k(\infty) = y_I^{(k)}(0^+)$ (with $\eta_0(\infty) = 0$). Reduced order models which match the 0, ..., k moments at infinity of system (2.24) can be constructed exploiting these definitions. See [3] for more details and an example.

2.5.3 Moment matching

We are now ready to introduce the notion of reduced order model by moment matching for nonlinear systems.

Definition 5. [3] Consider the signal generator (2.21). The system described by the equations

$$\dot{\xi} = \phi(\xi, u), \qquad \psi = \kappa(\xi), \tag{2.25}$$

with $\xi(t) \in \mathbb{R}^\nu$, is a *model at* (s, l) *of system (2.20)* if system (2.25) has the same moment at (s, l) as (2.20). In this case, system (2.25) is said to *match* the moment of system (2.20) at (s, l). Furthermore, system (2.25) is a reduced order model of system (2.20) if $\nu < n$.

Lemma 5. [3] Consider system (2.20), system (2.25) and the signal generator (2.21). Suppose Assumptions 1 and 2 hold. System (2.25) matches the moments of (2.20) at (s, l) if the equation

$$\phi(p(\omega), l(\omega)) = \frac{\partial p}{\partial \omega} s(\omega) \tag{2.26}$$

has a unique solution p such that

$$h(\pi(\omega)) = \kappa(p(\omega)), \tag{2.27}$$

where π is the (unique) solution of equation (2.23).

[5] Note that for non-affine systems the impulse response, and its derivatives, may not be well-defined functions, see [79, Chapter 10].

In other words, we have to determine the mappings ϕ, κ and p such that equations (2.26) and (2.27) hold. We introduce the following assumption to simplify the problem.

Assumption 3. There exist mappings κ and p such that $\kappa(0) = 0$, $p(0) = 0$, p is locally continuously differentiable, equation (2.27) holds and $\det \left.\frac{\partial p(\omega)}{\partial \omega}\right|_{\omega=0} \neq 0$, *i.e.* the mapping p possesses a local inverse p^{-1}.

Remark 8. [3] Similar to the linear case, Assumption 3 holds selecting $p(\omega) = \omega$ and $\kappa(\omega) = h(\pi(\omega))$. ■

Finally, as shown in [3], the system described by the equations

$$\dot{\xi} = s(\xi) - \delta(\xi)l(\xi) + \delta(\xi)u, \qquad \psi = h(\pi(\xi)), \tag{2.28}$$

where δ is any mapping such that the equation

$$\frac{\partial p}{\partial \omega} s(\omega) = s(p(\omega)) - \delta(p(\omega))l(p(\omega)) + \delta(p(\omega))l(\omega), \tag{2.29}$$

has the unique solution $p(\omega) = \omega$, is a family of *reduced order models of (2.20) at* (s, l).

2.5.4 Model reduction by moment matching with additional properties

Similarly to the linear case we can determine the conditions on the mapping δ such that the reduced order model satisfies additional properties. The proofs are omitted and can be found in [3]. Note also that therein other problems, such as matching with prescribed zero dynamics, matching with a passivity constraint and matching with L_2-gain, are discussed and solved.

Matching with asymptotic stability [3]

Consider the problem of determining a reduced order model (2.28) which has an asymptotically stable zero equilibrium. This problem can be solved if it is possible to select the mapping δ such that the zero equilibrium of the system $\dot{\xi} = s(\xi) - \delta(\xi)l(\xi)$ is locally asymptotically stable. To this end, for instance, it is sufficient that the pair $\left(\left.\frac{\partial l(\xi)}{\partial \xi}\right|_{\xi=0}, \left.\frac{\partial s(\xi)}{\partial \xi}\right|_{\xi=0} \right)$ is observable.

Matching with prescribed relative degree

The problem of constructing a reduced order model which has a given relative degree $r \in [1, \nu]$ at some point $\bar{\xi}$ can be solved selecting δ as follows.

Theorem 4. [3] For all $r \in [1, \nu]$ there exists a δ such that system (2.28) has relative degree r at $\bar{\xi}$ if and only if the codistribution

$$d\mathcal{O}_\nu = \text{span}\{dh(\pi(\xi)), \cdots, dL_s^{\nu-1}h(\pi(\xi))\} \tag{2.30}$$

has dimension ν at $\bar{\xi}$.

Matching and nonlinear systems at $S\omega$ [3]

Finally, in [3] the problem of model reduction for nonlinear systems when the signal generator is a linear system has been considered. This problem, which is relevant for one of the results developed in Chapter 3, is important because the resulting reduced order models have a very simple description, namely

$$\begin{aligned} \dot{\xi} &= (S - \delta(\xi)L)\xi + \delta(\xi)u, \\ \psi &= h(\pi(\xi)), \end{aligned}$$

with δ a free mapping. A further simplification can be achieved with the selection $\delta(\xi) = G$, for some constant matrix G. In fact, in this case the family of reduced order models is described by a linear differential equation with a nonlinear output map. The two obvious advantages of this result are that the matrix G can be selected to achieve the additional properties presented in Section 2.4.2 and that the determination of the reduced order model reduces to the computation of the mapping $h \circ \pi$. This can be determined as suggested in [80, Section 4.2 and 4.3] or as shown in Chapter 4.

2.6 Conclusion

In this chapter we have reviewed the model reduction technique by moment matching. We have recalled the classical interpolation theory and we have then introduced the steady-state-based notion of moment. Exploiting this description of moment the solution of the problem of model reduction by moment matching for linear and nonlinear systems achiev-

ing additional properties has been given and an enhancement of the notion of frequency response for nonlinear systems has been presented. These basic results are instrumental for the developments in the rest of the Thesis.

Chapter 3

Model reduction of neutral systems with discrete and distributed delays

3.1 Introduction

In this chapter we extend the model reduction techniques based on moment matching to a general class of linear and nonlinear differential time-delay systems. Time-delay systems are a class of infinite dimensional systems extensively studied in the literature, *e.g.* see the monographs [81–87] and the survey papers [6, 88, 89]. From a practical point of view every dynamical system presents delays of some extent, see for instance [85, 90], in which several examples from biology, chemistry, physics and engineering are discussed. Delays in closed-loop systems can generate unexpected behaviors: for instance "small" delays may be destabilizing [89, 91–95], while "large" delays may be stabilizing [96–99].

The problem of model reduction of time-delay systems is a classic topic in control theory. The optimal reduction (in the sense of some norm) is listed as an unsolved problem in systems theory in [100] and several results have been given using rational interpolations, *e.g.* [101–103], see also [104–110]. Recent results include model order reduction techniques for linear time-delay systems, see *e.g.* [111–113], and for infinite dimensional systems, see *e.g.* [114, 115] in which operators are used to provide reduced order models for linear systems.

In this chapter the model reduction theory introduced in Chapter 2 is extended to linear and nonlinear time-delay systems. For linear systems, it is shown that the moments of the system are fully characterized by the solution of a Sylvester-like equation. Although Sylvester equations have been widely studied (see for instance [116, 117]), some care is needed to extend the classical results to the particular Sylvester-like equation that arises in the chapter. The results are then generalized to the class of linear neutral differential time-delay systems with discrete and distributed delays. A family of systems that achieve moment matching is characterized and connections with the results of Chapter 2 are drawn. As noted in [118] a reduced order model with time-delays may lead to improvements in the approximation. Accordingly, the possibility to maintain the delay in the reduced order model is discussed and, in addition, it is shown that the introduction of delays can be used to improve the approximation, interpolating at a larger number of points. For nonlinear systems, the description of moment is based on the center manifold theory for time-delay systems and is derived using the steady-state response of the system. The conditions and properties of the center manifold hold as for finite dimensional systems, see *e.g.* [81, 119]. Exploiting this interpretation, a parameterized family of models described by differential time-delay equations is characterized. In addition, the notion of moment at infinity for linear and nonlinear time-delay systems is investigated and a brief discussion for nonlinear neutral differential time-delay systems is given. Finally, the problem of obtaining a reduced order model of the open-loop system given the closed-loop system is discussed and solved.

The rest of the chapter is organized as follows. In Section 3.2 the notion of moment is extended to linear time-delay systems (Section 3.2.1), the solution of the resulting Sylvester-like equation is discussed (Section 3.2.2) and a revisitation of this description is given using the notion of steady-state response (Section 3.2.3). The results are then extended to neutral systems and systems with distributed-delays (Section 3.2.4). A family of systems achieving moment matching is presented (Sections 3.2.5 and 3.2.6) and a simple example based on the model of a LC transmission line is given (Section 3.2.7). The possibility of interpolating a larger number of points maintaining the same "number of equations" is investigated (Section 3.2.8). Two examples, one inspired by [4] (Section 3.2.9) and one

inspired by the problem of the automatic control of a platoon of vehicles, illustrate the results (Section 3.2.10). In Section 3.3 the center manifold theory for time-delay systems is used to extend the definition of moment to nonlinear time-delay systems (Section 3.3.1) and a family of systems achieving moment matching for nonlinear time-delay systems is given (Sections 3.3.2 and 3.3.3). The possibility of interpolating multiple moments is investigated (Section 3.3.4) and the problem of "open-loop" reduced order models is discussed (Section 3.3.5). The theory is illustrated by the reduction of a nonlinear neutral differential time-delay equation describing the torsional dynamics of an oilwell drillstring (Section 3.3.6). In Section 3.4 the characterization of the moments at infinity for linear and nonlinear time-delay systems is presented. Finally Section 3.5 contains some concluding remarks.

All the results of this chapter are original contributions developed in fulfillment of my PhD course and they have been published in the conference papers [8], [9], in the book chapter [7] and in the journal paper [10].

3.2 Linear time-delay systems

In this section we derive a model reduction theory for linear differential time-delay systems. To keep the notation as simple as possible we begin with the class of systems with discrete-delays and we then discuss the extension of the results to more general types of delays and representations in Section 3.2.4.

3.2.1 Definition of Π: linear time-delay systems

Consider a linear, single-input, single-output, continuous-time, time-delay system with constant discrete-delays described by the equations[1]

$$\begin{aligned} \dot{x} &= \sum_{j=0}^{\varsigma} A_j x_{\tau_j} + \sum_{j=\varsigma+1}^{\mu} B_j u_{\tau_j}, \qquad & y &= \sum_{j=0}^{\varsigma} C_j x_{\tau_j}, \\ x(\theta) &= \phi(\theta), & -T &\leq \theta \leq 0, \end{aligned} \tag{3.1}$$

[1]The results can be extended to multi-input, multi-output (MIMO) systems straightforwardly. The problem in the MIMO case is called tangential interpolation, see [63, 120], and it is extensively discussed in Chapter 5.

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $\phi \in \mathfrak{R}_T^n$, $A_j \in \mathbb{R}^{n\times n}$ and $C_j \in \mathbb{R}^{1\times n}$ with[2] $j = 0, \ldots, \varsigma$, $B_j \in \mathbb{R}^{n\times 1}$ with $j = \varsigma+1, \ldots, \mu$, $\tau_0 = 0$ and $\tau_j \in \mathbb{R}_{>0}$ with $j = 1, \ldots, \mu$. Let

$$W(s) = \bar{C}(s)(sI - \bar{A}(s))^{-1}\bar{B}(s), \tag{3.2}$$

with

$$\bar{A}(s) = \sum_{j=0}^{\varsigma} A_j e^{-s\tau_j}, \quad \bar{B}(s) = \sum_{j=\varsigma+1}^{\mu} B_j e^{-s\tau_j}, \quad \bar{C}(s) = \sum_{j=0}^{\varsigma} C_j e^{-s\tau_j}, \tag{3.3}$$

be the associated transfer function and assume (3.1) is minimal[3], *i.e.* controllable and observable. We begin defining the moments of system (3.1) at some $s_i \in \mathbb{C}$ and showing that there exists a one-to-one relation between the moments and the (unique) solution of a Sylvester-like equation.

Definition 6. Let $s_i \in \mathbb{C} \setminus \sigma(\bar{A}(s))$. The 0-*moment of system (3.1) at* s_i is the complex number

$$\eta_0(s_i) = \bar{C}(s_i)(s_i I - \bar{A}(s_i))^{-1}\bar{B}(s_i).$$

The k-*moment of system (3.1) at* s_i is the complex number

$$\eta_k(s_i) = \frac{(-1)^k}{k!}\left[\frac{d^k}{ds^k}\bar{C}(s)(sI - \bar{A}(s))^{-1}\bar{B}(s)\right]_{s=s_i},$$

with $k \geq 1$ integer.

This definition of moment is justified by the fact that, as already stated, the k-moment of a linear system at s_i is defined as the k-th coefficient of the Laurent series expansion of the transfer function $W(s)$ at $s_i \in \mathbb{C}$ (see Chapter 2), provided it exists.

Lemma 6. Consider system (3.1) and let $s_i \in \mathbb{C} \setminus \sigma(\bar{A}(s))$. Assume that $\widetilde{\Pi} \in \mathbb{R}^{n\times\nu}$ is the unique solution of the Sylvester-like equation

$$\sum_{j=0}^{\varsigma} A_j \widetilde{\Pi} e^{-\Sigma_k \tau_j} + \sum_{j=\varsigma+1}^{\mu} B_j L_k e^{-\Sigma_k \tau_j} = \widetilde{\Pi}\Sigma_k, \tag{3.4}$$

[2]The delays of A_j and C_j are taken, without loss of generality, equal to ease the notation.

[3]See [121] for a in-depth discussion of the different characterizations of minimality for time-delay systems. Herein, we mean that there are no common zeros between the numerator and denominator of the transfer function.

with $L_k = [1\ 0\ \dots\ 0] \in \mathbb{R}^{(k+1)}$ and

$$\Sigma_k = \begin{bmatrix} s_i & 1 & 0 & \dots & 0 \\ 0 & s_i & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & s_i & 1 \\ 0 & \dots & \dots & 0 & s_i \end{bmatrix} \in \mathbb{R}^{(k+1)\times(k+1)}.$$

Then

$$\begin{bmatrix} \eta_0(s_i) & \dots & \eta_k(s_i) \end{bmatrix} = \sum_{j=0}^{\varsigma} C_j \widetilde{\Pi} e^{-\Sigma_k \tau_j} \Psi_k,$$

where $\Psi_k = \mathrm{diag}(1, -1, 1, \dots, (-1)^k) \in \mathbb{R}^{(k+1)\times(k+1)}$.

Proof. Let $\widetilde{\Pi} = [\widetilde{\Pi}_0\ \widetilde{\Pi}_1\ \dots\ \widetilde{\Pi}_k]$. Since Σ_k is in Jordan form then

$$e^{-\Sigma_k \tau_j} = e^{-s_i \tau_j} \begin{bmatrix} 1 & -\tau_j & \frac{(-\tau_j)^2}{2} & \dots & \frac{(-\tau_j)^{k-1}}{(k-1)!} \\ 0 & 1 & -\tau_j & \dots & \frac{(-\tau_j)^{k-2}}{(k-2)!} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -\tau_j \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix}.$$

Thus, the first column of equation (3.4) can be rewritten as

$$\sum_{j=0}^{\varsigma} A_j \widetilde{\Pi}_0 e^{-s_i \tau_j} + \sum_{j=\varsigma+1}^{\mu} B_j e^{-s_i \tau_j} = \widetilde{\Pi}_0 s_i, \tag{3.5}$$

the second column can be rewritten as

$$\sum_{j=0}^{\varsigma} A_j e^{-s_i \tau_j} \widetilde{\Pi}_1 + \sum_{j=0}^{\varsigma} -\tau_j A_j e^{-s_i \tau_j} \widetilde{\Pi}_0 - \sum_{j=\varsigma+1}^{\mu} \tau_j B_j e^{-s_i \tau_j} = \widetilde{\Pi}_1 s_i + \widetilde{\Pi}_0, \tag{3.6}$$

and so on until the last column

$$\sum_{l=k}^{0} \sum_{j=0}^{\varsigma} A_j \widetilde{\Pi}_{k-l} \frac{(-\tau_j)^l}{l!} e^{-s_i \tau_j} + \sum_{j=\varsigma+1}^{\mu} B_j \frac{(-\tau_j)^k}{k!} e^{-s_i \tau_j} = \widetilde{\Pi}_k s_i + \widetilde{\Pi}_{k-1}. \tag{3.7}$$

As a result, $\widetilde{\Pi}_0$ can be determined from equation (3.5) as

$$\widetilde{\Pi}_0 = \left(s_i I - \sum_{j=0}^{\varsigma} A_j e^{-s_i \tau_j} \right)^{-1} \sum_{j=\varsigma+1}^{\mu} B_j e^{-s_i \tau_j} = \left(s_i I - \bar{A}(s_i) \right)^{-1} \bar{B}(s_i),$$

$\widetilde{\Pi}_1$ from equation (3.6) and $\widetilde{\Pi}_0$ as

$$\begin{aligned} \widetilde{\Pi}_1 =& - \left(s_i I - \sum_{j=0}^{\varsigma} A_j e^{-s_i \tau_j} \right)^{-1} \left(I + \sum_{j=1}^{\varsigma} \tau_j A_j e^{-s_i \tau_j} \right) \left(s_i I - \sum_{j=0}^{\varsigma} A_j e^{-s_i \tau_j} \right)^{-1} \sum_{j=\varsigma+1}^{\mu} B_j e^{-s_i \tau_j} \\ & - \left(s_i I - \sum_{j=0}^{\varsigma} A_j e^{-s_i \tau_j} \right)^{-1} \sum_{j=\varsigma+1}^{\mu} \tau_j B_j e^{-s_i \tau_j} = \left[\frac{d}{ds} \left((sI - \bar{A}(s))^{-1} \bar{B}(s) \right) \right]_{s=s_i}. \end{aligned}$$

Iterating for all k, yields

$$\widetilde{\Pi}_k = \frac{1}{k!} \left[\frac{d^k}{ds^k} \left((sI - \bar{A}(s))^{-1} \bar{B}(s) \right) \right]_{s=s_i}.$$

Finally, exploiting the columns of $\widetilde{\Pi}$, the moments can be written as

$$\begin{aligned} \sum_{j=0}^{\varsigma} C_j \widetilde{\Pi} e^{-\Sigma_k \tau_j} &= \left[\begin{array}{ccc} \sum_{j=0}^{\varsigma} C_j \widetilde{\Pi}_0 e^{-s_i \tau_j} & \dots & \sum_{l=k}^{0} \sum_{j=0}^{\varsigma} C_j \widetilde{\Pi}_{k-l} \frac{(-\tau_j)^l}{l!} e^{-s_i \tau_j} \end{array} \right] \\ &= \left[\begin{array}{ccc} \bar{C}(s_i) \widetilde{\Pi}_0 & \dots & \sum_{l=k}^{0} \frac{1}{l!} \frac{d^l}{ds^l} \left[\bar{C}(s) \right]_{s=s_i} \widetilde{\Pi}_{k-l} \end{array} \right] \\ &= \left[\begin{array}{ccc} \eta_0(s_i) & \dots & (-1)^k \eta_k(s_i) \end{array} \right], \end{aligned}$$

which proves the claim. □

Equation (3.4) can be written eliminating the fact that Σ_k and L_k have a special structure. As a result the following holds.

Theorem 5. Consider system (3.1) and let s_i, with $i = 1, \dots, \eta$, be a set of numbers such that $s_i \in \mathbb{C} \setminus \sigma(\bar{A}(s))$. Let $S \in \mathbb{R}^{\nu \times \nu}$ be any non-derogatory matrix with characteristic polynomial

$$p(s) = \prod_{i=1}^{\eta} (s - s_i)^{k_i}, \tag{3.8}$$

where $\nu = \sum_{i=1}^{\eta} k_i$, and L be such that the pair (L, S) is observable. Assume that Π is the

unique solution of the Sylvester-like equation

$$\sum_{j=0}^{\varsigma} A_j \Pi e^{-S\tau_j} - \Pi S = -\sum_{j=\varsigma+1}^{\mu} B_j L e^{-S\tau_j}. \tag{3.9}$$

Then there exists a one-to-one relation between the moments $\eta_0(s_1)$, ..., $\eta_{k_1-1}(s_1)$, ..., $\eta_0(s_\eta)$, ..., $\eta_{k_\eta-1}(s_\eta)$ and the matrix $\sum_{j=0}^{\varsigma} C_j \Pi e^{-S\tau_j}$.

Proof. Note that it is sufficient to prove the claim for $\eta = 1$. By observability of the pair (L, S) there is a unique invertible matrix T such that $S = T^{-1}\Sigma_k T$ and $L = L_k T$. Then equation (3.9) becomes

$$\sum_{j=0}^{\varsigma} A_j \Pi e^{-(T^{-1}\Sigma_k T)\tau_j} - \Pi T^{-1}\Sigma_k T = -\sum_{j=\varsigma+1}^{\mu} B_j L e^{-(T^{-1}\Sigma_k T)\tau_j}.$$

The claim follows defining $\widetilde{\Pi} = \Pi T^{-1}$, recalling that $e^{T^{-1}XT} = T^{-1}e^X T$ and that the moments are coordinates invariant. □

3.2.2 Solution of the Sylvester-like equation

Equation (3.9) is a Sylvester equation only if $\varsigma = 0$. Nevertheless, it is a linear equation in Π and it can be solved with the use of the vectorization operator and the Kronecker product. To this end, it is necessary to determine when the equation admits a unique solution. In this section we solve this problem in the general case and for two special cases.

Lemma 7. Equation (3.9) has a unique solution if and only if $s_i \in \mathbb{C} \setminus \sigma(\bar{A}(s))$ for all $i = 1, \ldots, \eta$.

Proof. Suppose, without loss of generality, that the matrix S is in complex Jordan form. Then the matrices $S^\top$ and $e^{-S^\top\tau_j}$ are lower triangular and their i-th eigenvalue is s_i and $e^{-s_i\tau_j}$, respectively. We recall that the eigenvalues of the sum of lower triangular matrices is the sum of the eigenvalues. The claim follows since equation (3.9) has a unique solution (see [116, 117]) if and only if

$$\det\left(\sum_{j=0}^{\varsigma}\left(e^{-S^\top\tau_j}\otimes A_j\right) - S^\top \otimes I\right) \neq 0,$$

which holds if and only if

$$\prod_{i=1}^{\eta} \det\left(\sum_{j=0}^{\varsigma} A_j e^{-s_i\tau} - s_i I\right) \neq 0,$$

hence the claim. □

Corollary 1. Equation (3.9) has a unique solution if the following holds.

- $A_0 = 0$, $A_1 \neq 0$, $\mu = \varsigma = 1$, and $\sigma(A_1) \cap \sigma(Se^{S\tau}) = \emptyset$.
- The matrices A_j for $j = 0, 1, \ldots, \varsigma$ commute and $\sum_{j=0}^{\mu} e^{-s_l\tau_j}\lambda_{ji} \neq s_l$ for $i = 1, ..., n$ and $l = 1, ..., \eta$, with λ_{ji} and s_l eigenvalues of A_j and S, respectively.

Proof: The claim is a direct consequence of the use of the vectorization operator (see [116, 117]). □

3.2.3 Definition of Π: linear time-delay systems - Revisited

To prepare the ground for the study of nonlinear time-delay systems, in this section we revisit the interpolation-based description of moment developed just now and give the equivalent steady-state-based description using the center manifold theory. The center manifold theory for time-delay systems has been widely studied. The results in [119] establish that the theory for finite dimensional systems can be extended to infinite dimensional systems (and thus to time-delay systems). In particular, as for finite dimensional systems, if the linearized system has q eigenvalues on the imaginary axis then there exists a q-dimensional local integral manifold (referred to as center manifold) for the original system. In addition, the well-defined restriction of the dynamics of the system to the manifold is finite dimensional. An overview on the center manifold theory for time-delay systems has been given in [81] and references therein.

Theorem 6. Let $S \in \mathbb{R}^{\nu\times\nu}$ be any non-derogatory matrix with characteristic polynomial (3.8). Consider system (3.1) and assume $s_i \in \mathbb{C} \setminus \sigma(\bar{A}(s))$, with $i = 1, \ldots, \eta$, and $\sigma(\bar{A}(s)) \subset \mathbb{C}_{<0}$. Consider the interconnection of system (3.1) with the system

$$\dot{\omega} = S\omega, \qquad u = L\omega, \tag{3.10}$$

with L and $\omega(0)$ such that the triple $(L, S, \omega(0))$ is minimal. Then there exists a one-to-one relation between the moments $\eta_0(s_1)$, ..., $\eta_{k_1-1}(s_1)$, ..., $\eta_0(s_\eta)$, ..., $\eta_{k_\eta-1}(s_\eta)$ and the steady-state response of the output of such interconnected system.

Proof. Consider the interconnection of system (3.1) with system (3.10). By the assumptions on $\sigma(\bar{A}(s))$ and $\sigma(S)$, the interconnected system has a globally well-defined invariant manifold given by $\mathcal{M} = \{(x, \omega) \in \mathbb{R}^{n+\nu} \,:\, x = \Pi\omega\}$, with Π the unique solution of the Sylvester-like equation (3.9). We prove now that $\mathcal{M}$ is attractive. Consider the equation

$$\dot{\overbrace{x - \Pi\omega}} = \sum_{j=0}^{\varsigma} A_j x_{\tau_j} + \sum_{j=\varsigma+1}^{\mu} B_j L \omega_{\tau_j} - \Pi S \omega = \sum_{j=0}^{\varsigma} A_j x_{\tau_j} + \left(\sum_{j=\varsigma+1}^{\mu} B_j L e^{-S\tau_j} - \Pi S \right) \omega$$

in which we used the fact that $\omega(t - \tau_j) = e^{-S\tau_j}\omega(t)$. Substituting (3.9) in the right-hand side of the last equation, yields

$$\dot{\overbrace{x - \Pi\omega}} = \sum_{j=0}^{\varsigma} A_j (x_{\tau_j} - \Pi\omega_{\tau_j}).$$

Computing the Laplace transform on both sides yields

$$s(X(s) - \Pi\Omega(s)) - (x(0) - \Pi\omega(0)) = \left(\sum_{j=0}^{\varsigma} A_j e^{-s\tau_j} \right) (X(s) - \Pi\Omega(s))$$

and, by the assumptions on $\sigma(\bar{A}(s))$, we have

$$X(s) - \Pi\Omega(s) = (sI - \bar{A}(s))^{-1}(x(0) - \Pi\omega(0)).$$

Finally, computing the inverse Laplace transform, yields

$$x(t) - \Pi\omega(t) = \mathcal{L}^{-1}\{(sI - \bar{A}(s))^{-1}(x(0) - \Pi\omega(0))\}.$$

Since $\sigma(\bar{A}(s)) \subset \mathbb{C}_{<0}$, by [81, Chapter 1, Theorem 6.2], $\mathcal{M}$ is attractive. As a result

$$\begin{aligned} y(t) &= \sum_{j=0}^{\varsigma} C_j (x_{\tau_j} - \Pi\omega_{\tau_j}) + \sum_{j=0}^{\varsigma} C_j \Pi \omega_{\tau_j} \\ &= \sum_{j=0}^{\varsigma} C_j \Pi \omega_{\tau_j} + \sum_{j=0}^{\varsigma} C_j \mathcal{L}^{-1}\{(sI - \bar{A}(s))^{-1}(x(0) - \Pi\omega(0))\} e^{-S\tau_j} = \sum_{j=0}^{\varsigma} C_j \Pi e^{-S\tau_j}\omega + \varepsilon(t), \end{aligned}$$

where $\sum_{j=0}^{\varsigma} C_j \Pi e^{-S\tau_j}\omega(t)$ describes the steady-state response, whereas

$$\varepsilon(t) = \sum_{j=0}^{\varsigma} C_j \mathcal{L}^{-1}\{(sI - \bar{A}(s))^{-1}(x(0) - \Pi\omega(0))\}e^{-S\tau_j},$$

describes the transient response which vanishes exponentially. This proves the claim. □

Remark 9. Exploiting Theorem 6, in Chapter 4 the problem of model reduction for linear time-delay systems from input/output data is addressed. Therein, an algorithm that, given the signal ω and the output y, retrieves the moments of a system for which the matrices A_j, B_j and C_j are not known is devised. ■

Note that the importance of Theorem 6 goes beyond the simple computation of the moments because it highlights the relation between the steady-state response and the moments.

3.2.4 A general class of linear time-delay systems

All the results presented for discrete-delays can be generalized to linear neutral differential time-delay systems with distributed-delays. Consider a linear, single-input, single-output, continuous-time, neutral time-delay system with discrete-delays and distributed-delays described by the equations

$$\begin{aligned} \dot{x} &= \sum_{j=1}^{q} D_j \dot{x}_{c_j} + \sum_{j=0}^{\varsigma} A_j x_{\tau_j} + \sum_{j=\varsigma+1}^{\mu} B_j u_{\tau_j} + \sum_{j=1}^{r} \int_{t-h_j}^{t} (G_j x(\theta) + H_j u(\theta)) d\theta, \\ y &= \sum_{j=0}^{\varsigma} C_j x_{\tau_j}, \\ x(\theta) &= \phi(\theta), \qquad\qquad -T \leq \theta \leq 0, \end{aligned} \tag{3.11}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $\phi \in \mathfrak{R}_T^n$, $A_j \in \mathbb{R}^{n\times n}$ and $C_j \in \mathbb{R}^{1\times n}$ with $j = 0, \ldots, \varsigma$, $B_j \in \mathbb{R}^{n\times 1}$ with $j = \varsigma+1, \ldots, \mu$, $D_j \in \mathbb{R}^{n\times n}$ with $j = 1, \ldots, q$, $G_j \in \mathbb{R}^{n\times n}$ and $H_j \in \mathbb{R}^{n\times 1}$ with $j = 1, \ldots, r$, $\tau_0 = 0$, $\tau_j \in \mathbb{R}_{>0}$ with $j = 1, \ldots, \mu$, $c_j \in \mathbb{R}_{>0}$ with $j = 1, \ldots, q$ and

$h_j \in \mathbb{R}_{>0}$ with $j = 1, \dots, r$. The transfer function $W(s)$ is defined by (3.2) with

$$\begin{aligned}
\bar{A}(s) &= \sum_{j=1}^{q} D_j s e^{-sc_j} + \sum_{j=0}^{\varsigma} A_j e^{-s\tau_j} + \sum_{j=1}^{r} G_j \frac{1 - e^{-sh_j}}{s}, \\
\bar{B}(s) &= \sum_{j=\varsigma+1}^{\mu} B_j e^{-s\tau_j} + \sum_{j=1}^{r} H_j \frac{1 - e^{-sh_j}}{s}, \\
\bar{C}(s) &= \sum_{j=0}^{\varsigma} C_j e^{-s\tau_j}.
\end{aligned} \tag{3.12}$$

Theorem 7. Assume $0 \notin \sigma(S)$. Theorem 5 holds, with the same assumptions, for system (3.11) replacing equation (3.9) with

$$\begin{aligned}
\sum_{j=0}^{\varsigma} A_j \Pi e^{-S\tau_j} + \sum_{j=1}^{r} G_j \Pi S^{-1}(I - e^{-Sh_j}) + \sum_{j=1}^{q} D_j \Pi S e^{-Sc_j} - \Pi S = \\
= - \sum_{j=\varsigma+1}^{\mu} B_j L e^{-S\tau_j} - \sum_{j=1}^{r} H_j L S^{-1}(I - e^{-Sh_j}).
\end{aligned} \tag{3.13}$$

Proof. The proof is similar to that of Theorem 5, hence it is omitted. □

We introduce the following stability condition (*i.e.* a "formal stability" condition, see [122], [6]).

Assumption 4. Assume the difference equation

$$x(t) + \sum_{j=1}^{q} D_j x(t - c_j) = 0$$

is asymptotically stable.

Theorem 8. Assume $0 \notin \sigma(S)$ and Assumption 4 holds. Theorem 6 holds, with the same assumptions, for system (3.11).

Proof. The claim can be proved noting that $\sigma(\bar{A}(s)) \subset \mathbb{C}_{<0}$, with the definitions in (3.12), and Assumption 4 guarantee asymptotic stability of system (3.11) [123], [6]. The additional assumption that S is invertible is necessary because in equation (3.13) the distributed-delays generate terms in S^{-1}. □

Remark 10. As noted in [124], many systems can be described by the equations (3.11) with, in most cases, a single neutral delay, *i.e.* $q = 1$. In this case Assumption 4 holds if $\sigma(D_1) \subset \mathbb{D}_{<1}$, see [81] and [125]. ■

Remark 11. Since hyperbolic partial differential equations can be locally expressed as neutral time-delay systems and, conversely, any time-delay $y(t) = u(t-\tau)$ can be represented by a classical transport equation (see [126], [83], [6]), the techniques presented in this chapter can be used to establish a model reduction theory for some classes of partial differential equations: the example we study in Section 3.3.6 is, in fact, a neutral system originated from a partial differential equation. A similar remark applies for other relations, such as the ones established between time-delay systems and fractional derivation equations [127]. ■

To keep the notation light only the discrete-delay case is considered in the remaining of the chapter. However, the extension of the following results to system (3.11) is straightforward.

3.2.5 Reduced order model for linear time-delay systems

In this and the following sections a family of systems achieving moment matching is presented and the possibility of interpolating a larger number of points maintaining the same "number of equations" is investigated.

Definition 7. Consider the signal generator (3.10). The system described by the equations

$$\dot{\xi} = \sum_{j=0}^{\varrho} F_j \xi_{\chi_j} + \sum_{j=\varrho+1}^{\rho} G_j u_{\chi_j}, \qquad \psi = \sum_{j=0}^{d} H_j \xi_{\chi_j}, \tag{3.14}$$

with $\xi(t) \in \mathbb{R}^\nu$, $\psi(t) \in \mathbb{R}$, $F_j \in \mathbb{R}^{\nu\times\nu}$ for $j = 0, \dots, \varrho$, $G_j \in \mathbb{R}^{\nu\times 1}$ for $j = \varrho+1, \dots, \rho$, $H_j \in \mathbb{R}^{1\times\nu}$ for $j = 0, \dots, k$, $\chi_0 = 0$ and $\chi_j \in \mathbb{R}_{>0}$ for $j = 1, \dots, \max\{\rho, d\}$, is a *model of system (3.1) at* S, if system (3.14) has the same moments at S as (3.1).

Lemma 8. Consider system (3.1) and the signal generator (3.10). Suppose $\sigma(S) \cap \sigma(\bar{A}(s)) = \emptyset$. Then the system (3.14) is a model of system (3.1) at S, if there exists a unique solution P of the equation

$$\sum_{j=0}^{\varrho} F_j P e^{-S\chi_j} - PS = -\sum_{j=\varrho+1}^{\rho} G_j L e^{-S\chi_j}, \tag{3.15}$$

such that

$$\sum_{j=0}^{\varsigma} C_j \Pi e^{-S\tau_j} = \sum_{j=0}^{d} H_j P e^{-S\chi_j}, \tag{3.16}$$

where Π is the unique solution of (3.9). System (3.14) is a *reduced order model of system (3.1) at* S if $\nu < n$, or if $\varrho < \varsigma$, or if $\rho < \mu$, or if $d < \varsigma$.

Proof. The claim is a consequence of Definition 7 and the definition of moment. In fact, note that equation (3.15) defines the moments of the reduced order model, whereas equation (3.16) gives the matching condition between the moments of the system to be reduced and the moments of the reduced order model. □

The following lemma assures that the solution P exists, is unique and that equation (3.16) have a solution.

Lemma 9. Let $\bar{F}(s) = \sum_{j=0}^{\varrho} F_j e^{-s\chi_j}$. Equation (3.15) has a unique solution P and equation (3.16) has a solution uniquely determined by the matrices H_j if and only if $s_i \notin \sigma(\bar{F}(s))$ for all $i = 1, \ldots, \eta$ and the pair (L, S) is observable.

Proof. The uniqueness of the solution P of equation (3.15) follows from the arguments given in Section 3.2.2. The observability of the pair (L, S) guarantees that P is full rank and, as a consequence, (3.16) can always be solved. □

3.2.6 Reduced order model with free F_j

To construct a family of models that achieves moment matching at ν points one could select

$$\begin{aligned} F_0 &= S - \sum_{j=\varrho+1}^{\rho} G_j L e^{-S\chi_j} - \sum_{j=1}^{\varrho} F_j e^{-S\chi_j} \\ H_0 &= \sum_{j=0}^{\varsigma} C_j \Pi e^{-S\tau_j} - \sum_{j=1}^{d} H_j e^{-S\chi_j}, \end{aligned} \tag{3.17}$$

and note that this selection solves equations (3.15), (3.16) for $P = I$. This yields the family of reduced order models described by the equations

$$\begin{aligned} \dot{\xi} &= \left(S - \sum_{j=\varrho+1}^{\rho} G_j L e^{-S\chi_j} - \sum_{j=1}^{\varrho} F_j e^{-S\chi_j} \right) \xi + \sum_{j=1}^{\varrho} F_j \xi_{\chi_j} + \sum_{j=\varrho+1}^{\rho} G_j u_{\chi_j}, \\ \psi &= \left(\sum_{j=0}^{\varsigma} C_j \Pi e^{-S\tau_j} - \sum_{j=1}^{d} H_j e^{-S\chi_j} \right) \xi + \sum_{j=1}^{d} H_j \xi_{\chi_j}, \end{aligned} \tag{3.18}$$

with G_j and F_j any matrices such that $s_i \notin \sigma(\bar{F}(s))$, for all $i = 1, \ldots, \eta$.

The proposed model has several free design parameters, namely G_j, F_j, H_j, χ_j,

ϱ, ρ and d. We note that selecting $\varrho = 0$, $\rho = 1$, $d = 0$ and $\chi_1 = 0$ (in this case we define $G = G_1$), yields a reduced order model with no delays. In other words, we reduce an infinite dimensional system to a finite dimensional one, of dimension ν. This reduced order model coincides with the one in Chapter 2 (equation (2.14)) and all results therein are directly applicable: the parameter G can be selected to achieve matching with prescribed eigenvalues, matching with prescribed relative degree, etc.

Remark 12. The problem of reducing an infinite dimensional system to a finite dimensional one is not new in literature, see *e.g.* [114, 115, 128], and is how the model reduction has been traditionally intended for time-delay systems, see *e.g.* [100–110, 129], in which the problem of reducing the transfer function of a linear time-delay system to a rational function is studied. A variety of methods are used, *e.g.* Padé approximation, Taylor expansions, spline approximations and Hankel operator.

On the other hand, the choice of eliminating the delays is likely to destroy some underlying dynamics of the model and, as shown in [96–99, 118], delays are not always detrimental (for example to stability). With this in mind, a possible choice is to keep F_j, G_j and H_j free with $\varrho = d = \varsigma$ and $\rho = \mu$. In this case we can use the matrices F_j, G_j and H_j, with $\tau_j = \chi_j$, to maintain some important physical properties of the delay structure of the system. ■

3.2.7 Example: model of a LC transmission line

To illustrate the above idea consider the example in Section 2.5 of [85] in which a model of a LC transmission line in considered. The linear neutral differential time-delay system is described by the equations

$$\begin{aligned} \dot{x} &= Ax + D\dot{x}_\tau + Bu, \\ y &= Cx, \end{aligned} \tag{3.19}$$

with

$$A = -\frac{1}{C_1}\begin{bmatrix} \frac{1}{R_1} + \sqrt{\frac{C_0}{L_0}} & 0 \\ -C_1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & -\frac{2}{C_1}\sqrt{\frac{C_0}{L_0}}\alpha \\ 0 & \alpha \end{bmatrix},$$

$$B = \begin{bmatrix} b_1 & b_2 \end{bmatrix}^\top, \quad C = \begin{bmatrix} c_1 & c_2 \end{bmatrix},$$

$$\alpha = \frac{1 - R_0\sqrt{\frac{C_0}{L_0}}}{1 + R_0\sqrt{\frac{C_0}{L_0}}}, \quad \tau = 2\sqrt{L_0 C_0},$$

in which $C_1 \in \mathbb{R}_{>0}$, $R_1 \in \mathbb{R}_{>0}$, $C_0 \in \mathbb{R}_{>0}$, $L_0 \in \mathbb{R}_{>0}$, $R_0 \in \mathbb{R}_{>0}$, $b_1 \in \mathbb{R}$, $b_2 \in \mathbb{R}$, $c_1 \in \mathbb{R}$ and $c_2 \in \mathbb{R}$. The system is such that if $R_0\sqrt{C_0/L_0} = 1$ the delay part of the system disappears (a phenomenon called line-matching) and the model can be described by a system of ordinary differential equations. In the reduced model it is desirable to maintain this property to preserve the physical structure of the system. To simplify the example, suppose $S = 1$ and $L = 1$. Then the vector Π can be computed from equation (3.13), that in this case is

$$A\Pi + D\Pi e^{-\tau} - \Pi = -B,$$

which has a unique solution if

$$-\frac{1}{C_1}\left(\frac{1}{R_1} + \sqrt{\frac{C_0}{L_0}}\right) \neq 1, \quad e^{-\tau}\frac{1 - R_0\sqrt{\frac{C_0}{L_0}}}{1 + R_0\sqrt{\frac{C_0}{L_0}}} \neq 1.$$

Hence, a family of reduced order models, parameterized in G, is described by the equations

$$\begin{aligned} \dot{\xi} &= (1 - e^{-\tau}\alpha - G)\xi + \alpha\dot{\xi}_\tau + Gu, \\ \psi &= C\Pi\xi. \end{aligned} \tag{3.20}$$

Both equations (3.19) and (3.20) describe linear neutral differential time-delay systems when $R_0\sqrt{C_0/L_0} \neq 1$ and linear delay-free systems otherwise.

3.2.8 Reduced order model interpolating at $(\varrho + 1)\nu$ points

The matrices F_j and H_j in (3.18) are design parameters. In this section we show how to exploit them to achieve moment matching at more than ν points, still maintaining the same dimension ν of the matrix F_0. We analyze the case in which $\varrho = 1$, $\rho = 3$ and $d = 1$ (F_1, G_2, G_3 and H_1 are the free parameters), for ease of notation. We further assume without loss of generality that there are no delays in the equation of the output y of system (3.1). The general case can be analyzed in a similar way.

Proposition 4. Let $S_a \in \mathbb{R}^{\nu\times\nu}$ and $S_b \in \mathbb{R}^{\nu\times\nu}$ be two non-derogatory matrices such that

$\sigma(S_a) \cap \sigma(S_b) = \emptyset$ and let L_a and L_b be such that the pairs (L_a, S_a) and (L_b, S_b) are observable. Let $\Pi_a = \Pi$ be the unique solution of (3.9), with $L = L_a$ and $S = S_a$, and let $\Pi_b = \Pi$ be the unique solution of (3.9), with $L = L_b$ and $S = S_b$. Consider F_0 and H_0 as in (3.17) with $\chi_2 = 0$, $S = S_a$ and $L = L_a$.

- If $d = 1$ and $L = L_a = L_b$, system (3.14) with the selection

$$\begin{aligned} F_1 &= (S_b - S_a - G_3(e^{-S_b\chi_3} - e^{-S_a\chi_3}))(e^{-S_b\chi_1} - e^{-S_a\chi_1})^{-1}, \\ F_0 &= S_a - G_2 L - G_3 L e^{-S_a\chi_3} - F_1 e^{-S_a\chi_1}, \\ H_1 &= (C\Pi_b - C\Pi_a)(e^{-S_b\chi_1} - e^{-S_a\chi_1})^{-1}, \\ H_0 &= C\Pi_a - H_1 e^{-S_a\chi_1}, \end{aligned} \tag{3.21}$$

  belongs to the family (3.18) and is a reduced order model of system (3.1) achieving moment matching at S_a and S_b, for any G_2 and G_3 such that $s_i \notin \sigma(\bar{F}(s))$, for all $i = 1, \ldots, \eta$.

- If $d = 0$, the family (3.18) with

$$\begin{aligned} F_1 = (&P_b S_b - S_a P_b + G_2 L_a P_b + G_3 L_a e^{-S_b\chi_3} P_b \\ &- G_2 L_b - G_3 L_b e^{-S_b\chi_3})(P_b e^{-S_b\chi_1} - e^{-S_a\chi_1} P_b)^{-1}, \end{aligned} \tag{3.22}$$

  is, for some P_b such that $C\Pi_a P_b = C\Pi_b$, a reduced order model of system (3.1) achieving moment matching at S_a and S_b, for any G_2 and G_3 such that $s_i \notin \sigma(\bar{F}(s))$, for all $i = 1, \ldots, \eta$.

Proof. We begin with the case $d = 1$. Easy computations show that

$$\begin{aligned} F_0 &= S_a - G_2 L_a - G_3 L_a e^{-S_a\chi_3} - F_1 e^{-S_a\chi_1}, \\ H_0 &= C\Pi_a - H_1 e^{-S_a\chi_1}, \end{aligned} \tag{3.23}$$

defined in (3.17), solve

$$\begin{aligned} F_0 P_a + F_1 P_a e^{-S_a\chi_1} - P_a S_a &= -G_2 L_a - G_3 L_a e^{-S_a\chi_3}, \\ C\Pi_a &= H_0 P_a + H_1 P_a e^{-S_a\chi_1}, \end{aligned} \tag{3.24}$$

with $P_a = I$. F_1 given in (3.22) solves the equation

$$F_0 P_b + F_1 P_b e^{-S_b \chi_1} - P_b S_b = -G_2 L_b - G_3 L_b e^{-S_b \chi_3},$$

for any invertible P_b. Substituting H_0 in

$$C\Pi_b = H_0 P_b + H_1 P_b e^{-S_b \chi_1}, \tag{3.25}$$

yields

$$H_1 = (C\Pi_b - C\Pi_a P_b)(P_b e^{-S_b \chi_1} - e^{-S_a \chi_1} P_b)^{-1}.$$

The matrices F_0, F_1, H_0, H_1 are such that the resulting reduced order model achieves moment matching at S_a and S_b and selecting $L = L_a = L_b$ and $P_b = I$ they yield (3.21). If $d = 0$, equation (3.25) reduces to

$$C\Pi_b = H_0 P_b$$

for some P_b. We then have to prove that there always exists a P_b such that $C\Pi_a P_b = C\Pi_b$ and F_1 is well-defined. To prove the first claim note that the condition consists in finding ν^2 parameters to solve ν equations. If $C\Pi_a \neq 0$ there exist always such a P_b, full rank and upper triangular (possibly after a change of coordinates). Finally note that by the hypotheses on the system and the signal generator there exists at least a component of $C\Pi_a$ which is not zero. To prove the second claim we have to show that

$$\operatorname{rank}\left\{P_b e^{-S_b \chi_1} - e^{-S_a \chi_1} P_b\right\} = \nu. \tag{3.26}$$

Note now that selecting S_a and S_b in complex Jordan form implies that the matrices in equation (3.26) are all upper triangular. Condition (3.26) can be rewritten as

$$\pi_{bi}\left(e^{-s_{bi}\chi_1} - e^{-s_{ai}\chi_1}\right) \neq 0, \qquad \forall\, i = 1, \ldots, \nu,$$

with π_{bi}, s_{ai} and s_{bi} the eigenvalues of P_b, S_a and S_b, respectively. Since $\sigma(S_a) \cap \sigma(S_b) = \emptyset$ then $\sigma(e^{-S_a \chi_1}) \cap \sigma(e^{-S_b \chi_1}) = \emptyset$, hence the claim follows. □

The family of systems characterized in Proposition 4 achieves moment matching at

2ν interpolation points. Note that the matrices G_2 and G_3 remain free parameters and they can be used to achieve the properties discussed in Chapter 2. For instance G_2 and G_3 can be used to set both the eigenvalues of F_0 and F_1. In addition, note that G_j has exactly ν free parameters. Hence, for instance, to assign the eigenvalues of j F_j matrices, j G_j matrices are needed. In [62], as already hinted in [3], it has been shown how to select G (our G_2) to achieve the two-sided interpolation, *i.e.* how to exploit the free parameter G to achieve interpolation at 2ν points. The two techniques may be combined and, in the case of Proposition 4, G_2 and G_3 may be selected to achieve interpolation at 4ν points.

Remark 13. The result can be generalized to $\varrho > 1$ delays, obtaining a reduced order model that interpolates at $(\varrho + 1)\nu$ points. This result can be used also when the system to be reduced is not a time-delay system. In other words, a system described by ordinary differential equations can be reduced to a system described by time-delay differential equations with an arbitrary number of delays ϱ achieving moment matching at $(\varrho + 1)\nu$ points. This property of interpolating an arbitrary large number of points comes to the cost that the reduced order model becomes an infinite dimensional system. However, as noted in [96–99, 118], a reduced model with time delays may have better properties than one without delays. ■

3.2.9 Example: a system of order $n = 1006$

To illustrate the idea of approximating delay-free systems with time-delay systems exploiting the additional degrees of freedom to increase the number of interpolation points, we consider an example inspired by [4] (see also [5]). The example is a single-input, single-output system of order $n = 1006$, which has a Bode plot with three peaks, described by the equations

$$\dot{x} = Ax + Bu, \qquad y = Cx,$$

where $A = \operatorname{diag}(A_1, A_2, A_3, A_4)$, with

$$A_1 = \begin{bmatrix} -1 & 10 \\ -10 & -1 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} -1 & 20 \\ -20 & -1 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} -1 & 40 \\ -40 & -1 \end{bmatrix},$$

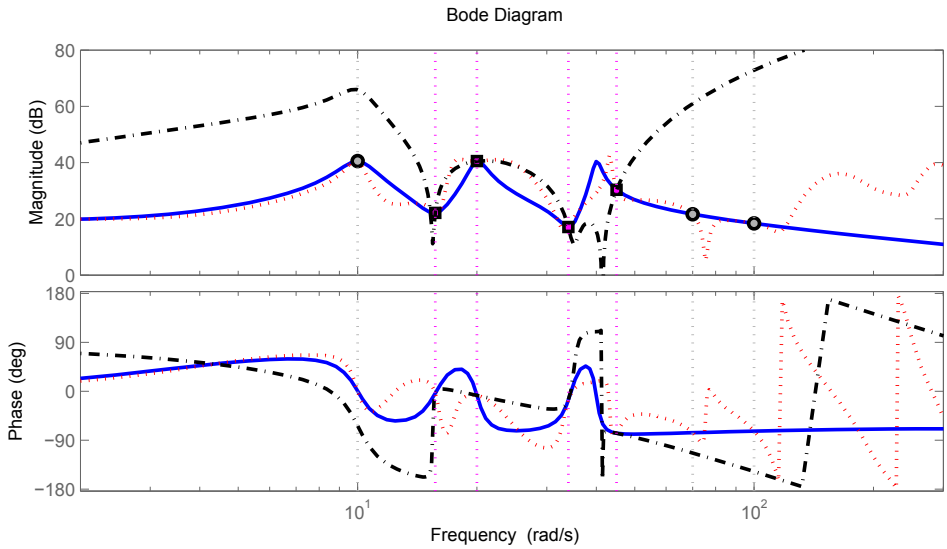

Figure 3.1: Bode plot of the system (solid line), of the delay-free reduced order model (dash-dotted line) and of the time-delay reduced order model (dotted line). The squares indicate the first set of interpolation points, whereas the circles indicate the second set.

and

$$A_4 = \mathrm{diag}(-1, -2, \ldots, -1000), \qquad B^\top = C = [\underbrace{10 \ldots 10}_{6 \text{ times}} \ \underbrace{1 \ldots 1}_{1000 \text{ times}}].$$

We start with obtaining a linear delay-free reduced order model of order $\nu = 8$. The matrices of the signal generator (3.10) have been selected as $S = S_a = \mathrm{diag}(S_2, S_3, S_4, S_5)$, with $S_2 = 1.57\, S_1$, $S_3 = 2\, S_1$, $S_4 = 3.4\, S_1$, $S_5 = 4.5\, S_1$, where $S_1 = A_1 + I$, and L randomly generated, to interpolate the moments close to the three peaks. The delay-free model (3.18) has been constructed with the technique presented in Chapter 2 assigning the eigenvalues of F_0 such that $\sigma(F_0) \subset \sigma(A)$. Fig. 3.1 shows the Bode plot of the system to be reduced (solid line) and of this reduced order model (dash-dotted line). The interpolation points are indicated by four squares. Note that the reduced order model approximates poorly the system because few interpolation points have been used (for comparison in [5] the order of the reduced order model is $\nu = 11$, while in [11] $\nu = 13$). We apply the technique presented in Proposition 4. The matrix S_b has been selected as $S_b = \mathrm{diag}(S_6, S_7, S_8, S_1)$,

with $S_7 = 7\,S_1$, $S_8 = 10\,S_1$ and

$$S_6 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Selecting $\chi = \chi_1 = \chi_3 = 0.05$, yields

$$F_0 = S_a - (S_b - S_a)(e^{-S_b\chi} - e^{-S_a\chi})^{-1}e^{-S_a\chi} - G_2L,$$

and

$$F_1 = (S_b - S_a)(e^{-S_b\chi} - e^{-S_a\chi})^{-1} - G_3L.$$

Note that, because of the selection $\chi_1 = \chi_3$, F_0 does not depend upon G_3. Thus, the eigenvalues of both F_0 and F_1 have been assigned such that $\sigma(F_0) = \sigma(F_1) \subset \sigma(A)$. In Fig. 3.1, the Bode plot of this reduced order model is represented by the dotted line. Three additional interpolation points are indicated with the circles. In addition the plot shows clearly that the model interpolates the point at zero, which is confirmed by a direct computation. Thus, the addition of one delay improved the quality of the approximation of the system without increasing the size of the matrices. However, note that a delay-free model with $\nu = 16$ would be a better approximation because the introduction of the delay is, at the same time, detrimental (in particular at high frequencies).

Remark 14. Although it is possible to interpolate at several different points s_i maintaining the same dimension ν, the order of interpolation at s_i cannot exceed ν because it is limited, by definition, by the dimension of the matrix S_j. ■

3.2.10 Example: reduction of a platoon of vehicles

Consider a controlled platoon of vehicles as presented in [130,131]. The platooning problem consists in controlling a group of vehicles tightly spaced following a leader, all moving in longitudinal direction. The advantages of the automatic cruise control is twofold. First, the use of automatic control to replace human drivers and their low-predictable reaction time with respect to traffic problems (spacing of around $30\,m$ at $60\,km/h$) can reduce the spacing distance between vehicles, consequently decreasing the traffic congestion. Second, the automatic control reduces the human error factor and then increases safety. In recent

years successful experiments involving autonomous vehicles have been carried out (*e.g.* the Google driver-less cars), and the use of this technology may be possible in the immediate future. However, when a large number of vehicles is considered, to study the dynamics of the whole platoon to guarantee individual vehicle stability and avoid slinky-type effect (*i.e.* the amplification of the spacing errors between subsequent vehicles as the vehicle "index" increases) can be computationally demanding (see [132]).

In what follows we use a model well-studied (see [85, 130, 131]), for which the solution of the platooning problem is known, to illustrate the results of the chapter. In particular, we are interested in reducing the number of vehicles to only a leader and a following car. Let $x_i(t)$ be the position of the i-th vehicle with respect to some well-defined reference, $v_i(t)$ its speed, $a_i(t)$ its acceleration and denote with $e_i = x_{i+1} - x_i - \ell_i$ the spacing error, with $\ell_i > 0$ the minimum separation distance. The resulting model is described by the equations

$$\begin{aligned} \dot{e}_i(t) &= v_{i+1}(t) - v_i(t), \\ \dot{v}_i(t) &= a_i(t), \\ \dot{a}_i(t) &= -\frac{a_i(t)}{c} + \frac{1}{c}[k_s e_i(t-\tau) + k_v(v_{i+1}(t-\tau) - v_i(t-\tau))], \end{aligned} \tag{3.27}$$

where $c > 0$ is the engine time constant, $\tau > 0$ is the total delay (including fueling and transport, etc.) for each vehicle, and k_s and k_v are the transmission gains between the vehicles. To this platoon we add a leader car with dynamics described by the equations

$$\begin{aligned} \dot{v}_n(t) &= a_n(t), \\ \dot{a}_n(t) &= -\frac{a_n(t)}{c} + \frac{1}{c} k_v \left(u(t) - v_n(t)\right), \end{aligned} \tag{3.28}$$

where $u(t)$ is a desired velocity imposed on the leader with no delay. We select as output of the system the sum of all the spacing errors, namely the distance between the first and the last vehicle. We rewrite the system in compact form as

$$\begin{aligned} \dot{x}(t) &= A_0 x(t) + A_1 x(t-\tau) + Bu(t), \\ y(t) &= Cx(t), \end{aligned} \tag{3.29}$$

with

$$A_0 = \frac{1}{c}\begin{bmatrix} A_0^1 & A_0^2 & 0 & \dots & 0 & 0 \\ 0 & A_0^1 & A_0^2 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & \ddots & 0 & A_0^1 & A_0^2 & 0 \\ 0 & \dots & \dots & 0 & A_0^1 & A_0^3 \\ 0 & \dots & \dots & \dots & 0 & A_0^4 \end{bmatrix}, \qquad A_1 = \frac{1}{c}\begin{bmatrix} A_1^1 & A_1^2 & 0 & \dots & 0 & 0 \\ 0 & A_1^1 & A_1^2 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & \ddots & 0 & A_1^1 & A_1^2 & 0 \\ 0 & \dots & \dots & 0 & A_1^1 & A_1^3 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 0 & 0 & | & 0 & 0 & 0 & | & \dots & | & 0 & 0 & 0 & | & 0 & \frac{k_v}{c} \end{bmatrix}^T,$$

$$C = \begin{bmatrix} 1 & 0 & 0 & | & 1 & 0 & 0 & | & \dots & | & 1 & 0 & 0 & | & 0 & 0 \end{bmatrix},$$

where

$$A_0^1 = \begin{bmatrix} 0 & -c & 0 \\ 0 & 0 & c \\ 0 & 0 & -1 \end{bmatrix},\ A_0^2 = \begin{bmatrix} 0 & c & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ A_0^3 = \begin{bmatrix} c & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},\ A_0^4 = \begin{bmatrix} 0 & c \\ -k_v & -1 \end{bmatrix},$$

$$A_1^1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ k_s & -k_v & 0 \end{bmatrix},\ A_1^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & k_v & 0 \end{bmatrix},\ A_1^3 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ k_v & 0 \end{bmatrix}.$$

Simulations

We consider $n = 8$ identical vehicles with $c = 0.25\,s$, $k_s = 0.875\,s^{-2}$, $k_v = 2.5\,s^{-1}$ and $\tau = 0.005\,s$.

We propose two reduced order models that match the 0-moments at $s_1 = 0$, $s_{2,3} = \pm\pi/5$, $s_{4,5} = \pm\pi/30$, with $u(t) = L\omega(t)$, $\dot{\omega}(t) = S\omega(t)$, $L = [1\ 0\ 1\ 0\ 1]'$, and described by the equations

$$\begin{aligned} \dot{\xi}(t) &= F_0\xi(t) + F_1\xi(t-\tau) + Gu(t), \\ \psi(t) &= C\Pi\xi(t), \end{aligned} \tag{3.30}$$

with F_0 defined as in (3.17) and F_1 free. Note that the number of equations decreases from $3n-1$ to ν. We denote with ψ_I the output of the system (3.30) when F_1 is defined

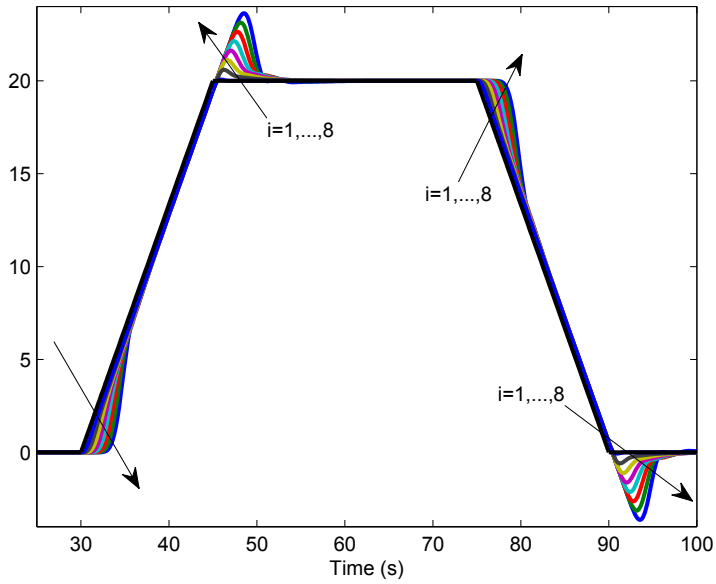

Figure 3.2: Speed of the eight vehicles.

as

$$F_1 = \frac{1}{c}\begin{bmatrix} A_1^1 & A_1^3 \\ 0 & 0 \end{bmatrix}. \tag{3.31}$$

Note that F_1 has the same structure of A_1. We denote with ψ_0 the output of the system (3.30) when $F_1 = 0$. In the latter case all the eigenvalues of the matrix F_0 have been placed at $-\frac{1}{2}$. The input given to the system consists of a speed increase from 0 to $20\,m/s = 72\,km/h$ in $15\,s$, a constant speed of $20\,m/s$ for $30\,s$ and a deceleration to $0\,m/s$ in $15\,s$. The speed of the vehicles are shown in Fig. 3.2. Fig. 3.3 shows the time histories of the output signals $y(t)$ (solid line), $\psi_I(t)$ (dashed line), $\psi_0(t)$ (dotted line). Fig. 3.4 shows the absolute errors between $y(t)$ and $\psi_I(t)$ (dashed line), and between $y(t)$ and $\psi_0(t)$ (dotted line). We see that the output is similar in the three cases and that the reduced order model with delays is tighter to the system, *i.e.* the ratio between the area under the error curve of the model with delays and the area under the error curve of the model with no delays is 0.799.

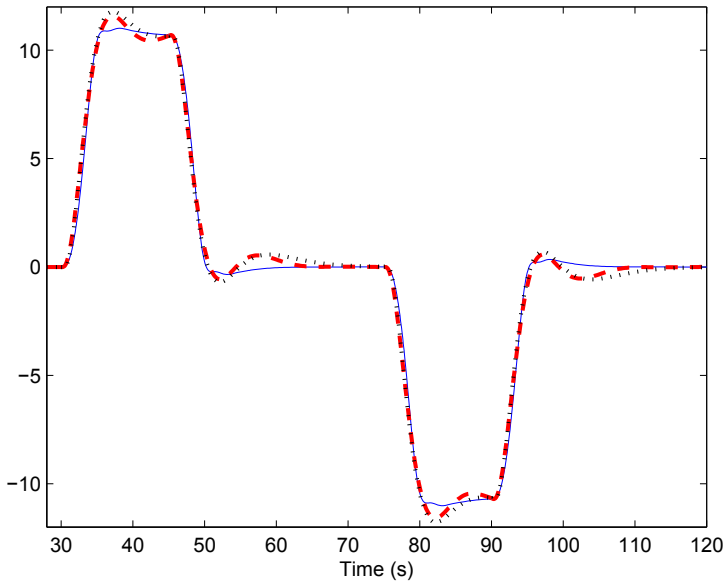

Figure 3.3: Output signals $y(t)$ (solid line), $\psi_I(t)$ (dashed line), and $\psi_0(t)$ (dotted line).

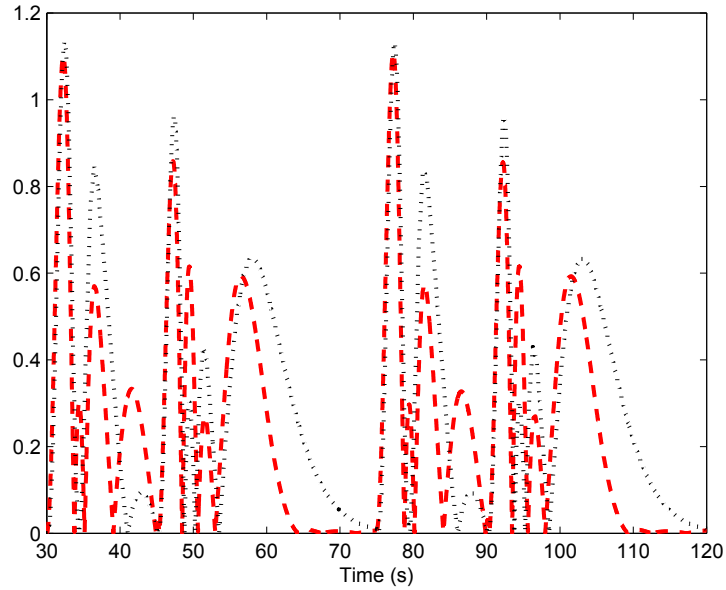

Figure 3.4: Absolute errors between $y(t)$ and $\psi_I(t)$ (dashed line), and between $y(t)$ and $\psi_0(t)$ (dotted line).

3.3 Model reduction for nonlinear time-delay systems

In this section we derive an extension of the model reduction method for nonlinear differential time-delay systems. To keep the notation simple we consider, without loss of generality, only delays (discrete or distributed) in the state and in the input, *i.e.* the output is delay-free. The neutral case is briefly discussed at the end of the section.

3.3.1 Definition of π: nonlinear time-delay systems

Consider a nonlinear, single-input, single-output, continuous-time, time-delay system described by the equations

$$\begin{aligned} &\dot{x} = f(x_{\tau_0}, \ldots, x_{\tau_\varsigma}, u_{\tau_{\varsigma+1}}, \ldots, u_{\tau_\mu}), \qquad y = h(x), \\ &x(\theta) = \phi(\theta), \qquad\qquad -T \leq \theta \leq 0, \end{aligned} \tag{3.32}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $\phi \in \mathfrak{R}^n_T$, $\tau_0 = 0$, $\tau_j \in \mathbb{R}_{>0}$ with $j = 1, \ldots, \mu$ and f and h smooth mappings. Consider a signal generator described by the equations

$$\dot{\omega} = s(\omega), \qquad\qquad u = l(\omega), \tag{3.33}$$

with $\omega(t) \in \mathbb{R}^\nu$, s and l smooth mappings, and the interconnected system

$$\begin{aligned} &\dot{\omega} = s(\omega), \\ &\dot{x} = f(x_{\tau_0}, \ldots, x_{\tau_\varsigma}, l(\omega_{\tau_{\varsigma+1}}), \ldots, l(\omega_{\tau_\mu})), \\ &y = h(x). \end{aligned} \tag{3.34}$$

Suppose that $f(0, \ldots, 0, 0, \ldots, 0) = 0$, $s(0) = 0$, $l(0) = 0$ and $h(0) = 0$.

Assumption 5. The signal generator (3.33) is observable and Poisson stable with $\omega(0) \neq 0$.

Assumption 6. Assume the zero equilibrium of the system $\dot{x} = f(x_{\tau_0}, \ldots, x_{\tau_\varsigma}, 0, \ldots, 0)$ is locally exponentially stable.

Lemma 10. Consider system (3.32) and the signal generator (3.33). Suppose Assumptions 5 and 6 hold. Then there exists a unique mapping $\pi(\omega)$, locally defined in a neigh-

borhood of $\omega = 0$, which solves the partial differential equation

$$\frac{\partial \pi}{\partial \omega} s(\omega) = f(\pi(\bar{\omega}_{\tau_0}), \ldots, \pi(\bar{\omega}_{\tau_\varsigma}), l(\bar{\omega}_{\tau_{\varsigma+1}}), \ldots, l(\bar{\omega}_{\tau_\mu})), \tag{3.35}$$

where $\bar{\omega}_{\tau_i} = \Phi^s_{\tau_i}(\omega)$, with $i = 0, \ldots, \mu$, is the flow of the vector field s at $-\tau_i$, see [78].

Remark 15. Lemma 10 implies that the interconnected system (3.34) possesses an invariant manifold, described by the equation $x = \pi(\omega)$. Note that the partial differential equation (3.35) is independent of time (as equation (3.9) and the correspondent equations given in Chapter 2), *e.g.* if $s(\omega) = S\omega$ then $\bar{\omega}_{\tau_i} = e^{-S\tau_i}\omega$. ■

Definition 8. Consider system (3.32) and the signal generator (3.33). Suppose Assumption 5 holds. The function $h \circ \pi$, with π solution of equation (3.35), is the *moment of system (3.32) at* (s, l).

Theorem 9. Consider system (3.32) and the signal generator (3.33). Suppose Assumptions 5 and 6 hold. Then the moment of system (3.32) at (s, l) coincides with the steady-state response of the output of the interconnected system (3.34).

Proof. Under the stated assumptions there exist a well-defined center manifold described by $x = \pi(\omega)$. In addition, the assumptions on the signal generator guarantee that the steady-state response of the output is locally well-defined and it is described by $h(\pi(\omega))$ (see [133]). □

3.3.2 Reduced order model for nonlinear time-delay systems

In this section a family of systems achieving moment matching is given.

Definition 9. Consider system (3.32) and the signal generator (3.33). Suppose Assumption 5 and 6 hold. Then the system

$$\dot{\xi} = \phi(\xi_{\chi_0}, \ldots, \xi_{\chi_\varrho}, u_{\chi_{\varrho+1}}, \ldots, u_{\chi_\rho}), \qquad \psi = \kappa(\xi), \tag{3.36}$$

with $\xi(t) \in \mathbb{R}^\nu$, $u(t) \in \mathbb{R}$, $\psi(t) \in \mathbb{R}$, $\chi_0 = 0$, $\chi_j \in \mathbb{R}_{>0}$ with $j = 1, \ldots, \rho$, and ϕ and κ smooth mappings, is a *model of system (3.32) at* (s, l) if system (3.36) has the same moment at (s, l) as system (3.32).

Lemma 11. Consider system (3.32) and the signal generator (3.33). Suppose Assump-

tion 5 and 6 hold. Then system (3.36) is a model of system (3.32) at (s, l) if the equation

$$\frac{\partial p}{\partial \omega} s(\omega) = \phi(p(\bar{\omega}_{\chi_0}), \ldots, p(\bar{\omega}_{\chi_\varrho}), l(\bar{\omega}_{\chi_{\varrho+1}}), \ldots, l(\bar{\omega}_{\chi_\rho})), \tag{3.37}$$

where $\bar{\omega}_{\chi_i} = \Phi^s_{\chi_i}(\omega)$, with $i = 0, \ldots, \rho$, has a unique solution p such that

$$h(\pi(\omega)) = \kappa(p(\omega)), \tag{3.38}$$

where π is the unique solution of (3.35). System (3.36) is a *reduced order model of system (3.32) at* (s, l) if $\nu < n$, or if $\varrho < \varsigma$, or if $\rho < \mu$.

Proof. The claim follows from Definition 9 and the definition of moment. □

3.3.3 The identity family of models

We now identify a simple family of models.

Assumption 7. There exist mappings κ and p such that $\kappa(0) = 0$, $p(0) = 0$, p is locally continuously differentiable, equation (3.38) holds and p has a local inverse p^{-1}.

Consistently with Lemma 11, a family of models that achieves moment matching at (s, l) is described by

$$\begin{aligned} \dot{\xi} &= \Phi(\xi, \bar{\xi}_{\chi_1}, \ldots, \bar{\xi}_{\chi_\varrho}) + \frac{\partial p(\omega)}{\partial \omega} \gamma(\xi_{\chi_1}, \ldots, \xi_{\chi_\varrho}) + \frac{\partial p(\omega)}{\partial \omega} \sum_{j=\varrho+1}^{\rho} \delta_j(\xi) u_{\chi_j}, \\ \psi &= \kappa(\xi), \end{aligned} \tag{3.39}$$

with

$$\Phi(\xi, \bar{\xi}_{\chi_1}, \ldots, \bar{\xi}_{\chi_\varrho}) = \left[\frac{\partial p(\omega)}{\partial \omega} (s(\omega) - \gamma(p(\bar{\omega}_{\chi_1}), \ldots, p(\bar{\omega}_{\chi_\varrho})) - \sum_{j=\varrho+1}^{\rho} \delta_j(p(\omega)) l(\bar{\omega}_{\chi_j})) \right]_{\omega = p^{-1}(\xi)},$$

where $\bar{\xi}_{\chi_j} = \left[\bar{\omega}_{\chi_j}\right]_{\omega = p^{-1}(\xi)}$, κ and p are such that Assumption 7 holds, p is the unique solution of (3.37) and δ_j and γ are free mappings.

Assumption 7 holds with the selection $p(\omega) = \omega$ and $\kappa(\omega) = h(\pi(\omega))$. This yields a family

of models described by the equations

$$\begin{aligned}\dot{\xi} &= s(\xi) - \sum_{j=\varrho+1}^{\rho} \delta_j(\xi) l(\bar{\xi}_{\chi_j}) - \gamma(\bar{\xi}_{\chi_1}, \ldots, \bar{\xi}_{\chi_\varrho}) + \gamma(\xi_{\chi_1}, \ldots, \xi_{\chi_\varrho}) + \sum_{j=\varrho+1}^{\rho} \delta_j(\xi) u_{\chi_j}, \\ \psi &= h(\pi(\xi)),\end{aligned} \tag{3.40}$$

where δ_j and γ are arbitrary mappings such that equation (3.37), namely

$$\begin{aligned}\frac{\partial p}{\partial \omega} s(\omega) &= s(p(\omega)) - \sum_{j=\varrho+1}^{\rho} \delta_j(p(\omega)) l(p(\bar{\omega}_{\chi_l})) - \gamma(p(\bar{\omega}_{\chi_1}), \ldots, p(\bar{\omega}_{\chi_\varrho})) \\ &\quad + \gamma(p(\omega_{\chi_1}), \ldots, p(\omega_{\chi_\varrho})) + \sum_{j=\varrho+1}^{\rho} \delta_j(p(\omega)) l(\omega_{\chi_j}),\end{aligned}$$

has the unique solution $p(\omega) = \omega$.

The nonlinear model (3.40) is the direct counterpart of the linear model (3.18). The model has several free design parameters, namely δ_j, γ, χ_j, ϱ and ρ. We note that selecting $\gamma \equiv 0$, $\varrho = 0$, $\rho = 1$ and $\chi_1 = 0$ (in this case we define $\delta = \delta_1$), yields a reduced order model with no delays. This reduced order model coincides with the one in Chapter 2 (equation (2.28)) and all results therein are directly applicable: the mapping δ can be selected to achieve matching with asymptotic stability, matching with prescribed relative degree, etc. However, as stressed previously in the chapter, the choice of eliminating the delays is likely to destroy some important dynamics of the model.

Remark 16. As in the case of linear time-delay systems the results of this section can be extended to more general classes of time-delay systems provided that, for such systems, the center manifold theory applies. In particular, one can consider the class of neutral differential time-delay systems described by equations of the form

$$\begin{aligned}d(\dot{x}_{\tau_0}, \ldots, \dot{x}_{\tau_{\varsigma_1}}) &= f(x_{\tau_{\varsigma_1+1}}, \ldots, x_{\tau_{\varsigma_2}}, u_{\tau_{\varsigma_2+1}}, \ldots, u_{\tau_\mu}), \\ y &= h(x),\end{aligned} \tag{3.41}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $\tau_0 = 0$, $\tau_j \in \mathbb{R}_{>0}$ with $j = 1, \ldots, \mu$ and d, f and h smooth mappings. The center manifold theory does not hold for this class of systems for a general mapping d. Specific cases have to be considered and we refer the reader

to [81], [125], [122] and references therein. Note, however, that for the simple case

$$\begin{aligned} &\dot{x} + D\dot{x}_{\tau_1} = f(x_{\tau_2}, \ldots, x_{\tau_{\varsigma_1}}, u_{\tau_{\varsigma_1+1}}, \ldots, u_{\tau_\mu}), \\ &y = h(x), \end{aligned} \tag{3.42}$$

with $D \in \mathbb{R}^{n\times n}$, the center manifold theory holds as for standard time-delay systems if the matrix D is such that $\sigma(D) \subset \mathbb{D}_{<1}$. ■

3.3.4 Matching at $h \circ \pi_a$ and $h \circ \pi_b$

In this section we present a nonlinear version of the result of Section 3.2.8, namely we show how to exploit the free parameters to achieve moment matching at two moments $h \circ \pi_a$ and $h \circ \pi_b$ maintaining the same number of equations describing the reduced order model. Consider system (3.32) and, to simplify the exposition, the signal generators described by the linear equation

$$\dot{\omega} = S_a\omega, \qquad u = L_{ab}\omega.$$

As highlighted in Chapter 2, considering the model reduction problem for nonlinear systems when the signal generator is a linear system is of particular interest since the reduced order models have a very simple description. This observation holds true also in the case of time-delay systems, namely a nonlinear time-delay system can be approximated by a linear time-delay equation with a nonlinear output map. Hence, a reduced order model of system (3.32) at (S_a, L_{ab}) is given by the family

$$\begin{aligned} \dot{\xi} &= F_0\xi + F_1\xi_\chi + G_2u + G_3u_\chi, \\ \psi &= \kappa_0(\xi) + \kappa_1(\xi_\chi), \end{aligned} \tag{3.43}$$

with κ_0 and κ_1 smooth mappings, if there exists a unique matrix P_a such that

$$\begin{aligned} &F_0P_a + F_1P_ae^{-S_a\chi} - P_aS_a = -G_2L_{ab} - G_3L_{ab}e^{-S_a\chi}, \\ &h(\pi_a(\omega)) = \kappa_0(P_a\omega) + \kappa_1(P_ae^{-S_a\chi}\omega), \end{aligned}$$

Consider now another signal generator described by the linear equation

$$\dot{\omega} = S_b\omega, \qquad u = L_{ab}\omega,$$

and the problem of selecting F_0, F_1, G_2, G_3, κ_0 and κ_1 such that the reduced order model (3.43) matches the moments of system (3.32) at (S_a, L_{ab}) and (S_b, L_{ab}).

Proposition 5. Let $S_a \in \mathbb{R}^{\nu\times\nu}$ and $S_b \in \mathbb{R}^{\nu\times\nu}$ be two non-derogatory matrices such that $\sigma(S_a) \cap \sigma(S_b) = \emptyset$ and let L_{ab} be such that the pairs (L_{ab}, S_a) and (L_{ab}, S_b) are observable. Let $\pi_a(\omega) = \pi(\omega)$ be the unique solution of (3.35), with $L = L_{ab}$ and $S = S_a$, and let $\pi_b(\omega) = \pi(\omega)$ be the unique solution of (3.35), with $L = L_{ab}$ and $S = S_b$. Then system (3.43) with the selection

$$\begin{aligned}
F_1 &= (S_b - S_a - G_3(e^{-S_b\chi} - e^{-S_a\chi_3}))(e^{-S_b\chi} - e^{-S_a\chi})^{-1},\\
F_0 &= S_a - G_2L_{ab} - G_3L_{ab}e^{-S_a\chi} - F_1e^{-S_a\chi},\\
\kappa_0(\omega) &= h(\pi_a(\omega)) - \kappa_1(e^{-S_a\chi}\omega),
\end{aligned}$$

and κ_1 a mapping such that

$$\kappa_1\left(e^{-S_b\chi}\omega\right) - \kappa_1\left(e^{-S_a\chi}\omega\right) = h(\pi_b(\omega)) - h(\pi_a(\omega)),$$

is a reduced order model of the nonlinear time-delay system (3.32) achieving moment matching at (S_a, L_{ab}) and (S_b, L_{ab}), for any G_2 and G_3 such that $s_i \notin \sigma(F_0 + F_1e^{-s\chi})$, for all $s_i \in \sigma(S_a)$ and $s_i \in \sigma(S_b)$.

Proof. As showed in the proof of Proposition 4, F_0 and F_1 solve the two Sylvester equations

$$\begin{aligned}
F_0P_a + F_1P_ae^{-S_a\chi} - P_aS_a &= -G_2L_{ab} - G_3L_{ab}e^{-S_a\chi},\\
F_0P_b + F_1P_be^{-S_b\chi} - P_bS_b &= -G_2L_{ab} - G_3L_{ab}e^{-S_b\chi},
\end{aligned}$$

with $P_a = P_b = I$. It remains to determine the mappings κ_0 and κ_1 that solve the matching conditions

$$\begin{aligned}
h(\pi_a(\omega)) &= \kappa_0(\omega) + \kappa_1\left(e^{-S_a\chi}\omega\right),\\
h(\pi_b(\omega)) &= \kappa_0(\omega) + \kappa_1\left(e^{-S_b\chi}\omega\right).
\end{aligned}$$

Solving the first equation with respect to κ_0 and substituting the resulting expression in the second yields

$$\kappa_1\left(e^{-S_b\chi}\omega\right) - \kappa_1\left(e^{-S_a\chi}\omega\right) = h(\pi_b(\omega)) - h(\pi_a(\omega)),$$

from which the claim follows. □

The family of linear time-delay systems with nonlinear output mapping characterized in Proposition 5 matches the moments $h \circ \pi_a$ and $h \circ \pi_b$ of the nonlinear system (3.32). Note that the matrices G_2 and G_3 remain free parameters and they can be used to achieve the properties discussed in Section 2.5.4.

Remark 17. Proposition 5 can be generalized to $\hat{\varrho} > 1$ delays, obtaining a reduced order model that match $(\hat{\varrho}+1)\nu$ moments. The result can also be generalized to nonlinear generators $s_i(\omega)$ assuming that the flow $\Phi_{\chi_i}^{s_i}(\omega)$ is known for all the delays χ_i and that $\gamma(\xi_{\chi_1}, \ldots, \xi_{\chi_{\hat{\varrho}}})$ in (3.40) is replaced by $\hat{\gamma}_1(\xi_{\chi_1}) + \cdots + \hat{\gamma}_{\hat{\varrho}}(\xi_{\chi_{\hat{\varrho}}})$. ■

Remark 18. Similarly to Proposition 4, the number of delays in (3.32) does not play a role in Proposition 5. Thus, this result can be applied to reduce a system with an arbitrary number of delays always obtaining a reduced order model with, for example, two delays. ■

3.3.5 Open-loop reduced order model

We consider now the problem of obtaining a reduced order model of an open-loop system from the closed-loop system. This problem may arise when the system to be reduced is not stable and we have to apply a feedback to use the reduction techniques proposed, yet we are interested in the reduced order model of the uncontrolled system. For ease of notation we assume that there are no delays on the input u.

Consider a closed-loop, nonlinear, single-input, single-output, continuous-time, time-delay system described by the equations

$$\dot{x} = f(x_{\tau_0}, \ldots, x_{\tau_{\mu_1}}, u), \qquad u = g(x_{\epsilon_0}, \ldots, x_{\epsilon_{\mu_2-1}}) + v_{\tau_{\mu_2}}, \qquad y = h(x), \tag{3.44}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $v(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $\tau_0 = 0$, $\tau_j \in \mathbb{R}_{>0}$ with $j = 1, \ldots, \mu_1$, $\epsilon_0 = 0$, $\epsilon_j \in \mathbb{R}_{>0}$ with $j = 1, \ldots, \mu_2$ and f, g and h smooth mappings. Consider the signal generator (3.33) and the interconnected system

$$\dot{\omega} = s(\omega), \qquad \dot{x} = f(x_{\tau_0}, \ldots, x_{\tau_{\mu_1}}, g(x_{\epsilon_0}, \ldots, x_{\epsilon_{\mu_2-1}}) + l(\omega_{\tau_{\mu_2}})), \qquad y = h(x). \tag{3.45}$$

Suppose that $f(0, \ldots, 0, 0) = 0$, $g(0, \ldots, 0, 0) = 0$, $s(0) = 0$, $l(0) = 0$ and $h(0) = 0$.

Assumption 8. Assume the zero equilibrium of the system $\dot{x} = f(x_{\tau_0}, \ldots, x_{\tau_{\mu_1}}, g(x_{\epsilon_0}, \ldots, x_{\epsilon_{\mu_2-1}}))$ is locally exponentially stable.

Lemma 12. Consider system (3.44) and the signal generator (3.33). Suppose Assumptions 5 and 8 hold. Then there exists a unique mapping $\pi(\omega)$, locally defined in a neighborhood of $\omega = 0$, which solves the partial differential equation

$$\frac{\partial \pi}{\partial \omega} s(\omega) = f(\pi(\bar{\omega}_{\tau_0}), \ldots, \pi(\bar{\omega}_{\tau_\mu}), g(\pi(\bar{\omega}_{\epsilon_0}), \ldots, \pi(\bar{\omega}_{\epsilon_{\mu_2-1}})) + l(\bar{\omega}_{\tau_{\mu_2}})), \tag{3.46}$$

where $\bar{\omega}_{\tau_i} = \Phi^s_{\tau_i}(\omega)$, with $i = 0, \ldots, \mu_1$, and $\bar{\omega}_{\epsilon_i} = \Phi^s_{\epsilon_i}(\omega)$, with $i = 0, \ldots, \mu_2$.

Definition 10. Consider system (3.44) and the signal generator (3.33). Suppose Assumption 5 and 8 hold. Then the system

$$\dot{\xi} = \phi(\xi_{\chi_0}, \ldots, \xi_{\chi_\rho}, u), \qquad \psi = \kappa(\xi), \tag{3.47}$$

with $\xi(t) \in \mathbb{R}^\nu$, $u(t) \in \mathbb{R}$, $\chi_0 = 0$, $\chi_j \in \mathbb{R}_{>0}$, with $j = 1, \ldots, \rho$, and ϕ and κ smooth mappings, is an *open-loop model of system (3.44) at* (s, l) if the system

$$\dot{\xi} = \phi(\xi_{\chi_0}, \ldots, \xi_{\chi_\rho}, u), \qquad u = g(\pi(p^{-1}(\xi_{\epsilon_0})), \ldots, \pi(p^{-1}(\xi_{\epsilon_{\mu_2-1}}))) + v_{\chi_{\mu_2}}, \qquad \psi = \kappa(\xi), \tag{3.48}$$

with $v(t) \in \mathbb{R}$, π the unique solution of (3.46) and p invertible and the unique solution of the equation

$$\frac{\partial p}{\partial \omega} s(\omega) = \phi(p(\bar{\omega}_{\chi_0}), \ldots, p(\bar{\omega}_{\chi_\rho}), l(\bar{\omega}_{\chi_{\mu_2}})), \tag{3.49}$$

where $\bar{\omega}_{\chi_i} = \Phi^s_{\chi_i}(\omega)$, with $i = 0, \ldots, \rho$, such that

$$h(\pi(\omega)) = \kappa(p(\omega)), \tag{3.50}$$

is a model of the (closed-loop) system (3.45) at (s, l).

Obtaining a reduced order model of an open-loop system given the closed-loop system solves the problem of the reduction of nonlinear systems when their zero equilibrium is not locally exponentially stable, extending the model reduction technique by moment matching to a larger class of systems. We illustrate this point with the next example.

3.3.6 Example: model of an oilwell drillstring

Consider a neutral type model of the torsional dynamics of an oilwell drillstring as presented in [134], [135], and [136]. In recent years, there has been increasing interest on the modeling and analysis of oilwell drilling vibrations because of its economic consequences (see [137] for an overview). In drilling operations the nonlinear interaction between the bit of the drillstring and the bottom of the hole originates the stick-slip phenomenon. This consists in the undesired event that a constant rotational velocity applied on the top of the string does not translate to a steady speed at the bottom of the hole. In particular the bit undergoes intervals where it is completely blocked and intervals where the accumulated energy is released and the rotational speed becomes larger than the prescribed value. Among several models of the dynamics of the system which have been presented, we consider the model presented in [134] (see [138] for an alternative). This model is described by the neutral differential time-delay system

$$\begin{aligned} \dot{x} &= \bar{\Upsilon}\dot{x}_{\tau_1} - \Psi x - \Psi\bar{\Upsilon}x_{\tau_1} - \frac{1}{I_B}T(x) + \frac{1}{I_B}\bar{\Upsilon}T(x_{\tau_1}) + 2\frac{\Psi c_a}{\Lambda}\Omega_{\tau_2}, \\ y &= x, \end{aligned} \tag{3.51}$$

with

$$\bar{\Upsilon} = \frac{c_a - \sqrt{IG_sJ}}{c_a + \sqrt{IG_sJ}}, \quad \Psi = \frac{\sqrt{IG_J}}{I_B}, \quad \Lambda = c_a - \sqrt{IG_sJ},$$

where $x(t)$ is the angular velocity at the bottom of the string, $y(t)$ is the output of the system, $\Omega(t)$ is the input variable, I is the inertia, J is the geometrical moment of inertia, G_s is the shear modulus, I_B is the lumped inertia representing the block at the bottom, c_a is a constant related to the local torsion of the drillstring, $\tau_2 = \Gamma$, $\tau_1 = 2\Gamma$, $\Gamma = L_s\sqrt{\dfrac{I}{G_sJ}}$ and L_s is the length of the string. The nonlinear function T describes the bit-rock interaction. Several models of this function have been proposed (for an overview see [137]). One model is given by the equation [134, 139],

$$T(x) = c_b x + W_{ob}R_b\left[\mu_{cb} + (\mu_{sb} - \mu_{cb})e^{-\frac{\gamma_b}{v_f}|x|}\right]\operatorname{sign}(x), \tag{3.52}$$

Table 3.1: Parameters of the model of the oilwell drillstring

$G_s = 79.3 \cdot 10^9\ N/m^2$,	$I = 0.095\ kg \cdot m$,	$L_s = 1172\ m$,
$J = 1.19 \cdot 10^{-5}\ m^4$,	$R_b = 0.155575$,	$v_f = 1$,
$W_{ob} = 97347\ N$,	$I_B = 89\ kg \cdot m^2$,	$\mu_{cb} = 20\ rad/s$,
$c_a = 2000\ N \cdot m \cdot s$,	$\mu_{sb} = 0.8$,	$\gamma_b = 0.9$,
$c_b = 0.03\ N \cdot m \cdot s/rad$,	$t_g = 10.$	$\zeta_1 = 6.96 \cdot 10^3$
$\zeta_2 = 0.09$,		

where μ_{sb}, μ_{cb} are, respectively, the static and Coulomb friction coefficients, W_{ob} is the weight on the bit, R_b is the bit radius, γ_b is a positive constant defining the decaying velocity of the exponential, v_f is a constant velocity introduced to have appropriate units and c_b is the viscous damping coefficient. This function has the disadvantages, for the sake of illustrating the results of this chapter, of being discontinuous and hard to simulate with sufficient precision.

The function

$$T(x) = \frac{\zeta_1 x}{x^2 + \zeta_2}, \tag{3.53}$$

where ζ_1, ζ_2 are positive parameters, has been proposed in [140]. This function has the advantage to be continuous, however, it does not approximate (3.52) sufficiently well far from $x = 0$. We propose, as a continuous approximation of (3.52), the function

$$T(x) = c_b x + W_{ob} R_b \left[\mu_{cb} + (\mu_{sb} - \mu_{cb}) e^{-\frac{\gamma_b}{v_f}|x|} \right] \tanh(t_g x), \tag{3.54}$$

where t_g is the gain of the hyperbolic tangent. Fig. 3.5 shows the graphs of the functions (3.52), (3.53), (3.54) and the error between (3.52) and (3.53) and between (3.52) and (3.54), as a function of the angular velocity x.

Simulations

The parameters have been selected as in [135] and are listed in Table 3.1. System (3.51), with (3.54), has been simulated using the Matlab solver *ddensd* with relative and absolute tolerances of 10^{-5} (the smallest presently supported) and 10^{-6}, respectively. Fig. 3.6 shows the state of system (3.51), with (3.54) and $\Omega(t) = r(t) = \text{const}$, for different values

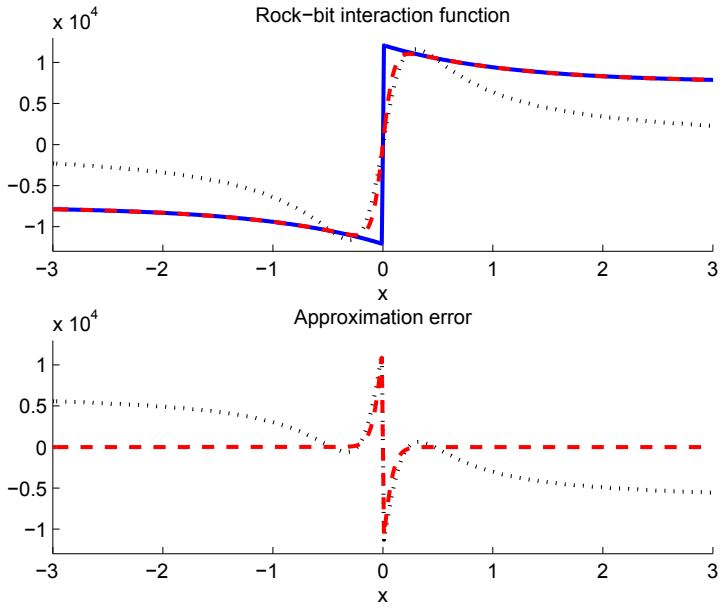

Figure 3.5: Top: graphs of the function (3.52) (solid line), (3.53) (dotted line) and (3.54) (dashed line). Bottom: graphs of the error between (3.52) and (3.53) (dotted line) and between (3.52) and (3.54) (dashed line).

of the desired angular velocity r. The typical behavior of the interaction between bit and rock can be seen, *i.e.* for high values of r the steady-state is constant whereas for low values of r the stick-slip phenomenon is present.

System (3.51), with (3.54), is a nonlinear neutral differential time-delay system for which the origin is not exponentially asymptotically stable and Lemma 11 cannot be directly applied. However, several closed-loop feedbacks have been proposed to asymptotically stabilize the origin of the system. We apply the feedback control law proposed in [135], namely

$$\Omega(t) = k_1 \dot{x}(t-\tau_2) + k_2 x(t-\tau_2) + r(t), \tag{3.55}$$

where $k_1 = -0.05$ and $k_2 = 0.36$. For the closed-loop system (3.51), with (3.54) and (3.55), we compute numerically the solution of equation (3.35). The function π is approximated by the piecewise continuous function

$$\pi(\omega) = \begin{cases} 0, & \omega \leq 5.7, \\ 1.5633\omega - 5.9250, & \omega > 5.7. \end{cases}$$

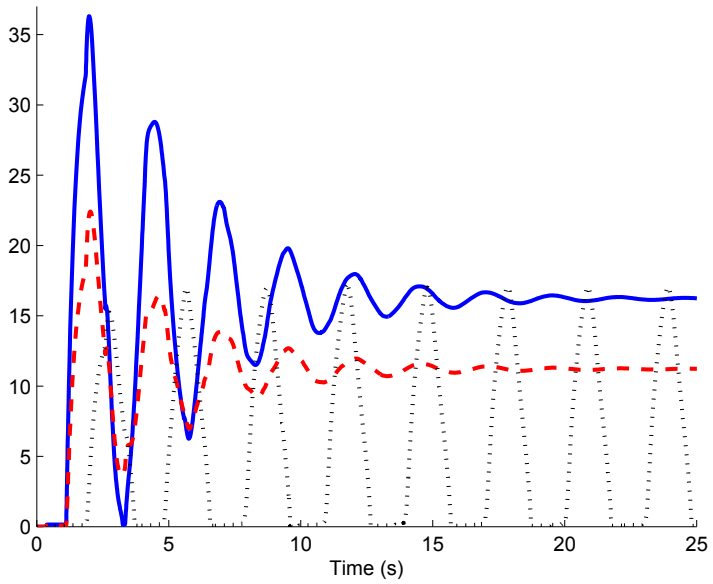

Figure 3.6: Angular speed of the system (3.51), with (3.54), for different values of r: 20 rad/s (solid line), 15 rad/s (dashed line) and 10 rad/s (dotted line). Note the stick-slip phenomenon for $r(t) = 10$ rad/s.

A simple reduced order model achieving moment matching at $\dot{\omega} = s(\omega) = 0$ and belonging to the family of models (3.40) is described by the equations

$$\dot{\xi} = -\delta(\xi)\left[\xi - r_{\tau_2}\right], \quad \psi = \pi(\xi). \tag{3.56}$$

Fig. 3.7 shows a comparison between the output of system (3.51), with (3.54) and (3.55), and the output of the reduced order model (3.56), with $\delta = 2$, for various desired angular velocities.

We are now interested, using system (3.56), in obtaining a model of the *open-loop system* (3.51) with (3.54). An open-loop model of system (3.51), with (3.54), achieving moment matching at $\dot{\omega} = 0$ is given by

$$\dot{\xi} = -\delta(\xi)\left[\xi - \mu_{\tau_2}\right], \quad \psi = \pi(\xi), \tag{3.57}$$

with $\mu = -k_1\pi(\dot{\xi}_{\tau_2}) - k_2\pi(\xi_{\tau_2}) + r$.

Fig. 3.8 shows the time histories of the output of the open-loop system (3.51), with (3.54),

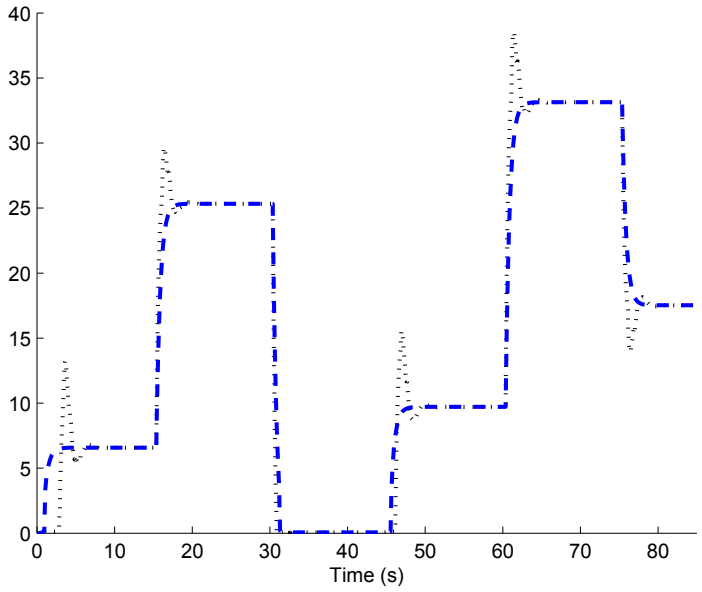

Figure 3.7: Time histories of the output of system (3.51), with (3.54) and (3.55), (dotted line) and the output ψ of the reduced order model (3.56), with $\delta = 2$, (dashed line) for various desired velocities.

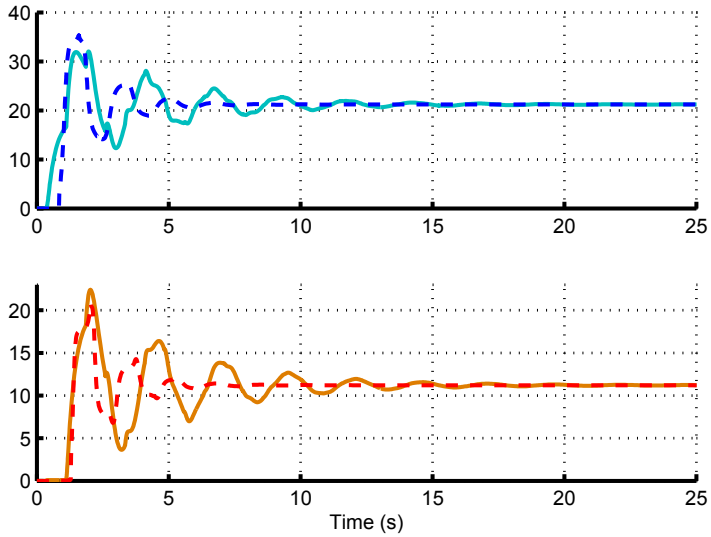

Figure 3.8: Time histories of the output of system (3.51), with (3.54), (solid line) and of system (3.57) (dashed line) with $\delta(z) = qz^2 + \varepsilon$ for $r = 25$, $q = 0.0333$, $\varepsilon = 0.3$ (top) and $r = 15$, $q = 0.125$, $\varepsilon = 0.01$ (bottom).

and of the model (3.57), with $\delta(z) = qz^2 + \varepsilon$, for $r = 25$, $q = 0.0333$, $\varepsilon = 0.3$ (top) and $r = 15$, $q = 0.125$, $\varepsilon = 0.01$ (bottom). We can see that the model (3.57) and the open-loop system (3.51), with (3.54), have the same steady-state value and that, using the free mapping δ, the transient behavior can be partially recovered.

3.4 Moment at infinity

We conclude the chapter with the characterization of the moments at infinity for linear and nonlinear time-delay systems. The description of moment given in Definitions 6 and 8 does not characterize the moments at $s_i = +\infty$. Since, as already pointed out, the k moments of linear delay-free systems at s_i are defined as the first k coefficients of the Laurent series expansion of the transfer function $W(s)$ at $s_i \in \mathbb{C}$, in a similar way the k moments at infinity can be computed by evaluating the expansion at infinity of the transfer function. In addition, by using the final value theorem (see e.g. [5] and [141]), the moments from 1 to $k+1$ correspond to the $j = 0, \ldots, k$ coefficients of the expansion at $t = 0^+$ of the impulse response (with $\eta_0(\infty) = 0$).

Note that for differential time-delay systems, the transfer function $W(s)$ is transcendental. This implies that the computation of the limit at $s = +\infty$ has to be done with care. As noted in [141] there is widespread carelessness in the definition of the Laplace transform and its properties. In our case, it is fundamental to determine what is the meaning of the limit at infinity because the result (if well-defined) would depend upon which direction at infinity is considered. As explained in [141], in this context the limit has to be taken along the positive real axis. Then we have the following results[4].

Theorem 10. Let Υ be the set of values of $j = \varsigma + 1, \ldots, \mu$ such that $\tau_j = 0$.

Consider system (3.1).

- If $\Upsilon \neq \emptyset$ then the k moments at infinity are $\eta_k(\infty) = \sum_{j \in \Upsilon} C A_0^{k-1} B_j$, with $\eta_0(\infty) = 0$.
- If $\Upsilon = \emptyset$ then all the moments at infinity are identically zero.

Consider system (3.32).

- If $\Upsilon \neq \emptyset$ then the k moments at infinity are[5] $\eta_k(\infty) = y_I^{(k)}(0^+)$, with $\eta_0(\infty) = 0$.

[4] To be coherent with equation (3.32) we ignore the delays in the equation of the output y of the linear system (3.1).

[5] We remind that $y_I^{(k)}$ denotes the k-th derivative of the impulse response of the system.

- If $\Upsilon = \emptyset$ then all the moments at infinity are identically zero.

Proof. By the equivalence between the moments at infinity and the impulse response at $t = 0^+$, it follows that if $\tau_j = 0$ for some $j = \varsigma + 1, \ldots, \mu$ the behavior of the systems at 0^+ is the same as the corresponding delay-free system (because no delay on the state has "kicked in" at $t = 0^+$). If $\tau_j \neq 0$ for all $j = \varsigma + 1, \ldots, \mu$ then the impulse response is delayed and it follows that the response at 0^+ is identically zero. Once established that the behavior is as for delay-free systems, the proof is as in [3]. □

Remark 19. From Theorem 10 it appears that a finite dimensional system is sufficient to characterize the moments at infinity and we can use it to match the moments at infinity as described in Chapter 2. However, note that some properties of the transfer function are lost with a finite dimensional system. In fact if, for instance, $\tau_j \neq 0$ for some $j = 1, \ldots, \varsigma$ and $\tau_j = 0$ for some $j = \varsigma + 1, \ldots, \mu$ the expansion at infinity along the *negative* real axis is identically zero, while for finite dimensional systems the expansion is the same as along the positive real axis. Or, if $\tau_1 = \tau_{\varsigma+1} > 0$ and $\tau_j = 0$ for $i = 2, \ldots, \varsigma, \varsigma+2, \ldots, \mu$, then the first coefficient of the expansion at infinity along the negative real axis is $-CB_1$ for time-delay systems, while zero (*i.e.* the same as along the positive axis) for finite dimensional systems. This suggests that a finite dimensional model that matches the moments of the system at infinity may not be a good approximation of the dynamics of the system far from $t = 0$ and that to preserve the properties of the transcendental transfer function it is necessary to choose a model with the same delay structure as the original system. ■

3.5 Conclusion

The model reduction theory based on moment matching has been extended to linear and nonlinear differential time-delay systems. The model reduction problem has been solved first for linear time-delay systems. The description of moment has been given in terms of the unique solution of a Sylvester-like equation. This has been subsequently extended to neutral differential time-delay systems with distributed and discrete delays. A family of systems achieving moment matching has been proposed and the problem of interpolating a larger number of points maintaining the same number of equations has been studied and solved. Then the definition of moment developed for linear time-delay systems has been extended to nonlinear systems by means of the center manifold theory.

The moments at infinity have been characterized for both linear and nonlinear time-delay systems and a parameterized family of models achieving moment matching has been given. The solution to the problem of obtaining a reduced order model of an unstable system has been given and connections with the delay-free framework, classical results in the literature and further developments have been drawn throughout the chapter. Finally, the theory has been illustrated by means of several examples.

Chapter 4

Data-driven model reduction for linear and nonlinear, possibly time-delay, systems

4.1 Introduction

The model reduction techniques that we have presented, as well as the majority of any model reduction method presented in the literature, assume the knowledge of a state-space model of the system to be reduced. However, in practice this model is not always available. In this chapter, inspired by the learning algorithm given in [142–144] to solve a model-free adaptive dynamic programming problem (see also the references therein, *e.g.* [145, 146]), we propose data-driven on-line algorithms for the model reduction of linear and non-linear, possibly time-delay, systems. Collecting at a given sequence of time instants t_k *time-snapshots* (which resemble the ones used to compute a proper orthogonal decomposition (POD), see *e.g.* [55, 147–149]) of the input and output of the system, an algorithm is devised to define a family of reduced order models (in the framework introduced in Chapter 2) at each instant of the iteration t_k. The reduced order model asymptotically matches the moments of the unknown system to be reduced. This algorithm has several advantages with respect to an identification plus reduction technique: there is no need to identify the system, which is expensive both in terms of computational power and storage memory; since the reduced order model matches the moments of the unknown system, it

is not just the result of a low-order identification but it actually retains some properties of the larger system; since the proposed algorithm determines directly the moment of a nonlinear system from the input and output data, it does not involve the computation of the solution of a partial differential equation, which is usually difficult.

The rest of the chapter is organized as follows. In Section 4.2.1 we give a preliminary analysis to compute on-line estimates of the moments of a linear system. In Section 4.2.2 (4.3.1) approximations which converge asymptotically to the moments of the linear (nonlinear) system are given. Therein, a discussion on the computational complexity associated to the evaluation of these approximations is presented, a recursive least-square formula is given and a moment estimation algorithm is provided. In Section 4.3.2 we present a simple experiment in which we show how to estimate the moment of a nonlinear system using the measured average heartbeat under "stress". In Sections 4.2.3 (4.3.3) and 4.2.4 (4.3.4) we give a family of reduced order models for linear (nonlinear) systems and linear (nonlinear) time-delay systems, respectively. In Section 4.2.5 we discuss how several properties, *e.g.* matching with prescribed eigenvalues or zeros, can be enforced in the present scenario. In Section 4.2.6 a linear reduced order model computed with the method proposed in the chapter is estimated for a system of order $n = 1010$ [4, 5]. In Section 4.3.5 a nonlinear reduced order model constructed using an approximation of the moment of the DC-to-DC Ćuk converter provides a further example. Finally Section 4.4 contains some concluding remarks.

All the results of this chapter are original contributions developed in fulfillment of my PhD course and they have been published in the conference papers [11], [12] and in the journal paper [13].

4.2 Linear (time-delay) systems

For the sake of convenience we report here the equations of system (2.1), namely

$$\dot{x} = Ax + Bu, \qquad y = Cx, \tag{4.1}$$

and the Sylvester equation (2.8), namely

$$A\Pi + BL = \Pi S, \tag{4.2}$$

which will be used through the chapter. The quantities appearing in the equations retain their original meaning.

4.2.1 A preliminary analysis

In this section we provide a preliminary analysis assuming to know the matrices A, B, C and the state $x(0)$ in equation (4.1). This analysis is used in the following section for the development of an estimation algorithm which, this time, does not use the matrices A, B, C and the state $x(0)$. To this end we make the following assumptions.

Assumption 9. The input u of system (4.1) is described by the equations

$$\dot{\omega} = S\omega, \qquad u = L\omega, \tag{4.3}$$

with S such that $\sigma(S) \subset \mathbb{C}_0$. In addition, assume that the triple $(L, S, \omega(0))$ is minimal.

Assumption 10. System (4.1) is asymptotically stable, *i.e.* $\sigma(A) \subset \mathbb{C}_{<0}$, and minimal.

Assumption 9 has a series of implications. The hypothesis on the eigenvalues of S is reasonable since the contribution of the negative eigenvalues of S to the response of the system decays to zero. The minimality of the triple $(L, S, \omega(0))$, which in turn implies the observability of the pair (L, S), guarantees that all the modes of S are present in the signals ω and u and it can be seen as a condition of persistency of excitation of order ν, see [150]. The choice of the particular structure (4.3) for the input u is limiting in applications in which the input cannot be arbitrarily chosen. We suggest in Section 4.2.2 a possible way to deal with alternative input signals. Note that the two assumptions imply that $\sigma(A) \cap \sigma(S) = \emptyset$, which in turn implies that equation (4.2) has a unique solution or, equivalently, that the steady-state in Theorem 1 is well-defined. Hence, recalling from Section 2, Assumptions 9 and 10 yield for all $t \in \mathbb{R}$

$$x(t) = \Pi\omega(t) + e^{At}(x(0) - \Pi\omega(0)). \tag{4.4}$$

Theorem 11. Let the time-snapshots $Q_k \in \mathbb{R}^{np\times n\nu}$ and $\chi_k \in \mathbb{R}^{np}$ be defined as

$$Q_k = \begin{bmatrix} \omega(t_{k-p+1})^\top \otimes I - \omega(0)^\top \otimes e^{At_{k-p+1}} \\ \vdots \\ \omega(t_{k-1})^\top \otimes I - \omega(0)^\top \otimes e^{At_{k-1}} \\ \omega(t_k)^\top \otimes I - \omega(0)^\top \otimes e^{At_k} \end{bmatrix},$$

and

$$\chi_k = \begin{bmatrix} x(t_{k-p+1}) - e^{At_{k-p+1}}x(0) \\ \vdots \\ x(t_{k-1}) - e^{At_{k-1}}x(0) \\ x(t_k) - e^{At_k}x(0) \end{bmatrix},$$

respectively, where $0 \le t_0 < t_1 < \cdots < t_{k-p} < \cdots < t_k < \cdots < t_q$, with $p > 0$ and $q \ge p$, and define $T_k^p = \{t_{k-p+1}, \ldots, t_{k-1}, t_k\}$. Assume the matrix Q_k has full rank, then

$$\operatorname{vec}(\Pi) = (Q_k^\top Q_k)^{-1} Q_k^\top \chi_k. \tag{4.5}$$

Proof. Equation (4.4) can be rewritten as

$$\Pi\omega(t) - e^{At}\Pi\omega(0) = x(t) - e^{At}x(0). \tag{4.6}$$

Using the vectorization operator and the Kronecker product on equation (4.6) yields

$$\operatorname{vec}(\Pi\omega(t)) - \operatorname{vec}(e^{At}\Pi\omega(0)) = \operatorname{vec}(x(t) - e^{At}x(0)),$$

and

$$(\omega(t)^\top \otimes I - \omega(0)^\top \otimes e^{At}) \operatorname{vec}(\Pi) = \operatorname{vec}(x(t) - e^{At}x(0)). \tag{4.7}$$

Computing equation (4.7) at all elements of T_k^p yields

$$Q_k \operatorname{vec}(\Pi) = \chi_k. \tag{4.8}$$

If the matrix Q_k has full rank, we can compute Π from the last equation yielding equation (4.5). □

Note that the selection of the set T_k^p can affect the *quality* of the data and the

rank of the matrix Q_k. Thus, to assure that T_k^p is non-pathological [151] we introduce the following technical assumption.

Assumption 11. The elements of T_k^ν are such that rank $\left(\left[\begin{array}{ccc} \omega(t_{k-\nu+1}) & \dots & \omega(t_k) \end{array}\right]\right) = \nu$ for all k.

Remark 20. If Assumption 9 holds, *i.e.* $(S, \omega(0))$ is controllable, it is always possible to choose the elements of T_k^ν such that Assumption 11 holds. *Vice versa* Assumption 11 implies the controllability of $(S, \omega(0))$. See [26] for more details. ■

Lemma 13. Suppose Assumptions 9, 10 and 11 hold. If $p = \nu$, Q_k is full rank.

Proof. Consider the i-block element of the matrix Q_k, namely

$$\omega(t_i)^\top \otimes I - \omega(0)^\top \otimes e^{At_i}.$$

Note that the properties of the Kronecker product yield

$$\begin{aligned} &\omega(t_i)^\top \otimes I - \omega(0)^\top \otimes e^{At_i} = \omega(0)^\top e^{S^\top t_i} \otimes II - \omega(0)^\top I \otimes I e^{At_i} = \\ &\quad = (\omega(0)^\top \otimes I)(e^{S^\top t_i} \otimes I) - (\omega(0)^\top \otimes I)(I \otimes e^{At_i}) = (\omega(0)^\top \otimes I)(e^{S^\top t_i} \otimes I - I \otimes e^{At_i}). \end{aligned}$$

Since $\sigma(A) \subset \mathbb{C}_{<0}$ and $\sigma(S) \subset \mathbb{C}_0$, the controllability of $(S, \omega(0))$ implies that the i-block element of the matrix Q_k is a $n \times n\nu$ matrix of rank n. Assumption 11 implies that ν of these blocks are linearly independent for any $t_i > 0$. As a result Q_k is a square full rank matrix.

□

Remark 21. Since real data are affected by noise the assumptions of Lemma 13 may not hold. In this case p can be taken larger than $n\nu$ and, as well-known from linear algebra and remarked in [142] and [150], the solution of equation (4.5) is the least squares solution of (4.8). ■

The discussion carried out so far has the drawback that requires information on the state of the system. In practice, this is usually not the case and only the output y is available.

Theorem 12. Let the time-snapshots $R_k \in \mathbb{R}^{w \times n\nu}$ and $\Upsilon_k \in \mathbb{R}^w$ be defined as

$$R_k = \begin{bmatrix} (\omega(0)^\top \otimes C)(e^{S^\top t_{k-w+1}} \otimes I - I \otimes e^{At_{k-w+1}}) \\ \vdots \\ (\omega(0)^\top \otimes C)(e^{S^\top t_{k-1}} \otimes I - I \otimes e^{At_{k-1}}) \\ (\omega(0)^\top \otimes C)(e^{S^\top t_k} \otimes I - I \otimes e^{At_k}) \end{bmatrix},$$

and

$$\Upsilon_k = \begin{bmatrix} y(t_{k-w+1}) - Ce^{At_{k-w+1}}x(0) \\ \vdots \\ y(t_{k-1}) - Ce^{At_{k-1}}x(0) \\ y(t_k) - Ce^{At_k}x(0) \end{bmatrix},$$

respectively. Assume the matrix R_k has full rank, then

$$\mathrm{vec}(\Pi) = (R_k^\top R_k)^{-1} R_k^\top \Upsilon_k. \tag{4.9}$$

Proof. The result can be proved following the same steps used to obtain equation (4.5). □

Lemma 14. Suppose Assumptions 9, 10 and 11 hold. If $w = n\nu$, R_k is full rank.

Proof. The proof is similar to the one of Lemma 13, although this time also the observability of (C, A) is used. □

4.2.2 On-line moment estimation from data

Equation (4.5) contains terms that depend upon the matrix A and the initial states $x(0)$ and $\omega(0)$. However, we note that given the stability hypothesis of the system, these terms are exponentially decaying functions.

Definition 11. Let the time-snapshots $\widetilde{Q}_k \in \mathbb{R}^{np \times n\nu}$ and $\widetilde{\chi}_k \in \mathbb{R}^{np}$ be

$$\widetilde{Q}_k = \begin{bmatrix} \omega(t_{k-p+1}) \otimes I & \dots & \omega(t_{k-1}) \otimes I & \omega(t_k) \otimes I \end{bmatrix}^\top$$

and

$$\widetilde{\chi}_k = \begin{bmatrix} x(t_{k-p+1})^\top & \dots & x(t_{k-1})^\top & x(t_k)^\top \end{bmatrix}^\top.$$

Assume the matrix $\widetilde{Q}_k$ has full rank, then following the same steps used to obtain equa-

tion (4.5), we define

$$\mathrm{vec}(\widetilde{\Pi}_k) = (\widetilde{Q}_k^\top \widetilde{Q}_k)^{-1} \widetilde{Q}_k^\top \widetilde{\chi}_k. \tag{4.10}$$

Note that if Assumption 11 holds and $p = \nu$, $\widetilde{Q}_k$ is full rank.

Lemma 15. Suppose Assumptions 9, 10 and 11 hold. There exists a matrix $\bar{\Pi}$ such that $\lim_{t_k \to \infty} \widetilde{\Pi}_k = \bar{\Pi}$.

Proof. By Assumption 9 and 10 there exists a matrix $\bar{\Pi}$ such that the steady-state response $x^{ss}(t)$ of the interconnection of system (4.1) and the generator (4.3) is described by the equation $x^{ss}(t) = \bar{\Pi}\omega(t)$. Then substituting $\widetilde{\chi}_k^{ss} = \widetilde{Q}_k \,\mathrm{vec}(\bar{\Pi})$ in equation (4.10), yields

$$\lim_{t_k \to \infty} \mathrm{vec}(\widetilde{\Pi}_k) = (\widetilde{Q}_k^\top \widetilde{Q}_k)^{-1} \widetilde{Q}_k^\top \widetilde{\chi}_k^{ss} = \mathrm{vec}(\bar{\Pi}),$$

which is well-defined by Assumption 11 if $p = \nu$. □

Theorem 13. Let Π be the solution of equation (4.2). Suppose Assumptions 9, 10 and 11 hold. There exists a sequence $\{t_k\}$ such that

$$\lim_{k \to \infty} \widetilde{\Pi}_k = \Pi.$$

Proof. The matrix $\widetilde{\Pi}_k$ defined in equation (4.10) is such that

$$x(t_k) = \widetilde{\Pi}_k \omega(t_k), \tag{4.11}$$

whereas Π is such that

$$\dot{x}(t)|_{t=t_k} = \Pi S \omega(t_k) + A e^{A t_k} (x(0) - \Pi\omega(0)) \tag{4.12}$$

hold. Consider the first equation of system (4.1) computed at t_k, namely

$$\dot{x}(t)|_{t=t_k} = A x(t_k) + BL\omega(t_k). \tag{4.13}$$

Substituting equations (4.11) and (4.12) in equation (4.13) yields

$$\Pi S \omega(t_k) + A e^{A t_k} (x(0) - \Pi\omega(0)) = A\widetilde{\Pi}_k \omega(t_k) + BL\omega(t_k)$$

and

$$\left(A\widetilde{\Pi}_k + BL - \Pi S\right)\omega(t_k) = Ae^{At_k}(x(0) - \Pi\omega(0)),$$

from which, using equation (4.2) and Assumption 10, the equation

$$\left(\widetilde{\Pi}_k - \Pi\right)\omega(t_k) = e^{At_k}(x(0) - \Pi\omega(0))$$

follows. By Assumption 9 there exists a sequence $\{t_k\}$, with $\lim_{k\to\infty} t_k = \infty$, such that for any $t_i \in \{t_k\}$, $\omega(t_i) \neq 0$ and Assumption 11 holds. By Assumption 10

$$\lim_{k\to\infty}\left(\widetilde{\Pi}_k - \Pi\right)\omega(t_k) = \lim_{k\to\infty} e^{At_k}(x(0) - \Pi\omega(0)) = 0,$$

and by Assumptions 11 and Lemma 15, $\lim_{k\to\infty}\left(\widetilde{\Pi}_k - \Pi\right) = \lim_{k\to\infty}\left(\bar{\Pi} - \Pi\right) = 0$. It follows that $\widetilde{\Pi}_k$ converges asymptotically to Π. □

Remark 22. Equation (4.11) is reminiscent of the POD of the collection $\{x(t_i)\}$. However, the two concepts are quite different. In fact, the POD of $\{x(t_i)\}$ is

$$\begin{bmatrix} x(t_0) & \dots & x(t_q) \end{bmatrix} = \underbrace{\begin{bmatrix} u(t_0) & \dots & u(t_q) \end{bmatrix}}_{U}\Delta$$

with $\Delta \in \mathbb{R}^{q\times q}$ and $U^*U = I$, where the superscript $*$ indicates the complex conjugate transpose. The dimensions of Δ are related to the number of samples, whereas the dimensions of Π are related to the ones of the system to be reduced and of the signal generator. In fact, the POD is a decomposition of the entire cloud of data $\{x(t_i)\}$ along the vectors $u(t_i)$, called principal directions of $\{x(t_i)\}$ [5]. Instead, in the technique proposed in this chapter the oldest data are discarded as soon as new data satisfying Assumption 11 is collected. As a consequence, while Δ is built to describe the entire dynamics of $\{x(t_i)\}$, Π is built to describe only the steady-state response of the system to be reduced. The result is that the POD is usually used with the Petrov-Galerkin projection for a SVD-based approximation [148], [55], whereas this technique is a moment matching method. ■

A similar discussion can be carried out for equation (4.9) that contains also terms which depend upon the matrix C. In this case note that equation (4.4) can be written as

$$y(t) = C\Pi\omega(t) + \varepsilon(t),$$

with $\varepsilon(t) = Ce^{At}(x(0) - \Pi\omega(0))$ an exponentially decaying signal.

Theorem 14. Define the time-snapshots $\widetilde{R}_k \in \mathbb{R}^{w\times\nu}$ and $\widetilde{\Upsilon}_k \in \mathbb{R}^w$ as

$$\widetilde{R}_k = \begin{bmatrix} \omega(t_{k-w+1}) & \dots & \omega(t_{k-1}) & \omega(t_k) \end{bmatrix}^\top$$

and

$$\widetilde{\Upsilon}_k = \begin{bmatrix} y(t_{k-w+1}) & \dots & y(t_{k-1}) & y(t_k) \end{bmatrix}^\top.$$

Assume the matrix $\widetilde{R}_k$ has full rank, then

$$\operatorname{vec}(\widetilde{C\Pi}_k) = (\widetilde{R}_k^\top \widetilde{R}_k)^{-1}\widetilde{R}_k^\top \widetilde{\Upsilon}_k, \tag{4.14}$$

is an approximation of the on-line estimate $C\Pi_k$, namely there exists a sequence $\{t_k\}$ such that

$$\lim_{k\to\infty} \widetilde{C\Pi}_k = C\Pi.$$

Proof. Equation (4.14) can be derived following the same steps used to obtain equation (4.5). We delay the proof of the convergence of the limit to $C\Pi$ to Chapter 5 (Theorem 18). □

Note that if Assumption 11 holds and $w = \nu$, $\widetilde{R}_k$ is full rank.

Remark 23. The matrix $\widetilde{R}_k$ is considerably smaller than R_k since it is not obtained from Kronecker products. ■

With equation (4.14) we are not able to retrieve the matrix $\widetilde{\Pi}_k$, but only $\widetilde{C\Pi}_k$. However, as shown in equation (2.14), we only need $C\Pi$ to compute the reduced order model, *i.e.* Π is not explicitly required. Equation (4.14) is a classic least-square estimation formula. The following result holds.

Theorem 15. Assume that $\Phi_k = (\widetilde{R}_k^\top \widetilde{R}_k)^{-1}$ and $\Psi_k = (\widetilde{R}_{k-1}^\top \widetilde{R}_{k-1} + \omega(t_k)\omega(t_k)^\top)^{-1}$ are full rank for all $t \geq t_r$ with $t_r \geq t_w$. Given $\operatorname{vec}(\widetilde{C\Pi}_r)$, Φ_r and Ψ_r, the recursive least-square formula

$$\begin{aligned}\operatorname{vec}(\widetilde{C\Pi}_k) = \operatorname{vec}(\widetilde{C\Pi}_{k-1}) &+ \Phi_k\omega(t_k)\Big(y(t_k) - \omega(t_k)^\top \operatorname{vec}(\widetilde{C\Pi}_{k-1})\Big) \\ &-\Phi_k\omega(t_{k-w})\Big(y(t_{k-w}) - \omega(t_{k-w})^\top\operatorname{vec}(\widetilde{C\Pi}_{k-1})\Big),\end{aligned} \tag{4.15}$$

with

$$\Phi_k = \Psi_k - \Psi_k \omega(t_{k-w}) \left(I + \omega(t_{k-w})^\top \Psi_k \omega(t_{k-w}) \right)^{-1} \omega(t_{k-w})^\top \Psi_k \tag{4.16}$$

and

$$\Psi_k = \Phi_{k-1} - \Phi_{k-1} \omega(t_k) \left(I + \omega(t_k)^\top \Phi_{k-1} \omega(t_k) \right)^{-1} \omega(t_k)^\top \Phi_{k-1}. \tag{4.17}$$

holds for all $t \geq t_r$.

Proof. The formula is obtained adapting the results in [150] (see also [152–154]) to the present scenario, in which at each step we acquire a new measure and we discard an old measure: for completeness we provide the details of the proof. Note that

$$\begin{aligned} \Phi_k^{-1} = \widetilde{R}_k^\top \widetilde{R}_k = \sum_{i=k-w+1}^{k} \omega(t_i)\omega(t_i)^\top &= \sum_{i=k-w}^{k-1} \omega(t_i)\omega(t_i)^\top + \omega(t_k)\omega(t_k)^\top - \omega(t_{k-w})\omega(t_{k-w})^\top \\ &= \Phi_{k-1}^{-1} + \omega(t_k)\omega(t_k)^\top - \omega(t_{k-w})\omega(t_{k-w})^\top \end{aligned}$$

and that rewriting equation (4.14) for $k-1$ yields

$$\widetilde{R}_{k-1}^\top \widetilde{\Upsilon}_{k-1} = \sum_{i=k-w}^{k-1} \omega(t_i) y(t_i) = \Phi_{k-1}^{-1} \operatorname{vec}(\widetilde{C\Pi}_{k-1}).$$

Substituting the first equation in the second we obtain

$$\sum_{i=k-w}^{k-1} \omega(t_i) y(t_i) = \Phi_k^{-1} \operatorname{vec}(\widetilde{C\Pi}_{k-1}) - \omega(t_k)\omega(t_k)^\top \operatorname{vec}(\widetilde{C\Pi}_{k-1}) + \omega(t_{k-w})\omega(t_{k-w})^\top \operatorname{vec}(\widetilde{C\Pi}_{k-1}),$$

which substituted in (4.14), namely

$$\operatorname{vec}(\widetilde{C\Pi}_k) = \Phi_k \left(\sum_{i=k-w}^{k-1} \omega(t_i) y(t_i) + \omega(t_k) y(t_k) - \omega(t_{k-w}) y(t_{k-w}) \right)$$

yields equation (4.15). Finally equations (4.16) and (4.17) are obtained applying recursively the matrix inversion lemma [150] to

$$\Phi_k = \left(\Psi_k^{-1} - \omega(t_{k-w})\omega(t_{k-w})^\top \right)^{-1}$$

and

$$\Psi_k = \left(\Phi_{k-1}^{-1} + \omega(t_k)\omega(t_k)^\top\right)^{-1}.$$

□

Remark 24. The construction of the initial values $\mathrm{vec}(\widetilde{C\Pi}_r)$, Φ_r and Ψ_r needed to start the recursion can be done in two ways: the first consists in using equation (4.14) to build $\mathrm{vec}(\widetilde{C\Pi}_r)$, Φ_r and Ψ_r and then updating the estimate with the equations in Theorem 15. However, this method has the drawback of requiring the inversion of $(\widetilde{R}_k^\top \widetilde{R}_k)^{-1}$. The second method consists in starting with dummy initial values $\mathrm{vec}(\widetilde{C\Pi}_r)$, Φ_r and Ψ_r. Since the formulas "forget" the oldest measurements, after a sufficient number of iterations all the dummy measurements are forgotten. ■

Remark 25. For single-input, single-output systems the two matrix inversions in the definition of Φ_k and Ψ_k are two divisions. Equations (4.15)-(4.16)-(4.17) can be used to compute a fast, on-line, estimate of $\widetilde{C\Pi}_k$, since the computational complexity of updating (4.15) is $\mathcal{O}(1)$. Thus, the implementation of equations (4.15)-(4.16)-(4.17) to update $\widetilde{C\Pi}_k$ is preferred to equation (4.14), which has a computational complexity, when $w = \nu$, of[1] $\mathcal{O}(\nu^{2.373})$ at each iteration k. ■

Remark 26. In comparison, the Arnoldi or Lanczos procedure for the model reduction by moment matching have a computational complexity of $\mathcal{O}(\nu n^2)$ [5, Section 14.1] (or $\mathcal{O}(\alpha\nu n)$ for a sparse matrix A, with α the average number of non-zero elements per row/column of A). In addition, note that these procedures require a model to be reduced and thus further expensive computation has to be considered for the identification of the original system. ■

The approximations $\widetilde{\Pi}_k$ and $\widetilde{C\Pi}_k$ can be computed with the following algorithm.

Algorithm 1. Let k be a sufficiently large integer. Select $\eta > 0$ sufficiently small. Select $w \geq n\nu$ ($w \geq \nu$, respectively).

1: Construct the matrices $\widetilde{Q}_k$ and $\widetilde{\chi}_k$ ($\widetilde{R}_k$ and $\widetilde{\Upsilon}_k$, respectively).

2: **If** $\mathrm{rank}\left(\widetilde{Q}_k\right) = n\nu$ ($\mathrm{rank}\left(\widetilde{R}_k\right) = \nu$, respectively) **then** compute $\widetilde{\Pi}_k$ ($\widetilde{C\Pi}_k$, respectively) solving equation (4.10) ((4.14), or (4.15), respectively).

Else increase w.

[1] This is the computational complexity of the fastest algorithm [155] for the inversion and multiplication of matrices.

If $k - w < 0$ **then** restart the algorithm selecting a larger initial k.

3: **If**

$$\left|\left|\widetilde{\Pi}_k - \widetilde{\Pi}_{k-1}\right|\right| > \frac{\eta}{t_k - t_{k-1}}, \qquad \left(\left|\left|\widetilde{C\Pi}_k - \widetilde{C\Pi}_{k-1}\right|\right| > \frac{\eta}{t_k - t_{k-1}}, \text{ respectively}\right) \tag{4.18}$$

then $k = k + 1$ **go to** 1.

4: **Stop**.

Remark 27. It is not always possible to arbitrarily select the input of the system to be reduced. For instance the input signal may be composed by several unwanted frequencies. Instead of system (4.3), consider the input described by the equations

$$\dot{\omega} = S\omega, \qquad u = L\omega + v,$$

with $v(t) \in \mathbb{R}^n$ an unknown signal. In this case the output response of system (4.1) is

$$y(t) = C\Pi\omega(t) + Ce^{At}(x(0) - \Pi\omega(0)) + \int_0^t e^{A(t-\tau)}Bv(\tau)d\tau,$$

which can be written as

$$y(t) = C\Pi\omega(t) + \varepsilon(t) + \mathfrak{v}(t),$$

with $\mathfrak{v}(t) = \int_0^t e^{A(t-\tau)}Bv(\tau)d\tau$ and $\varepsilon(t) = Ce^{At}(x(0) - \Pi\omega(0))$. One can then apply the filtering techniques given in [150, Chapter 11]: we filter out $\mathfrak{v}$ from y and u with a band-pass filter and apply the results of the chapter to the filtered y_f and u_f. ■

4.2.3 Families of reduced order models

Using the approximations given by Algorithm 1 a reduced order model of system (4.1) can be defined at each instant of time t_k.

Definition 12. Consider system (4.1) and the signal generator (4.3). Suppose Assumptions 9, 10 and 11 hold. Then the system

$$\dot{\xi} = F_k\xi + G_k u, \qquad \psi = H_k\xi, \tag{4.19}$$

with $\xi(t) \in \mathbb{R}^\nu$, $F_k \in \mathbb{R}^{\nu\times\nu}$, $G_k \in \mathbb{R}^{\nu\times 1}$, $H_k \in \mathbb{R}^{1\times\nu}$, is a *model of system (4.1) at S at*

time t_k, if there exists a unique solution P_k of the equation

$$F_k P_k + G_k L = P_k S, \tag{4.20}$$

such that

$$\widetilde{C\Pi}_k = H_k P_k, \tag{4.21}$$

where $\widetilde{C\Pi}_k$ is the solution of (4.14).

Proposition 6. Select $P_k = I$, for all $k \geq 0$. If $\sigma(F_k) \cap \sigma(S) = \emptyset$ for all $k \geq 0$, then the model

$$\dot{\xi} = (S - G_k L)\xi + G_k u, \qquad \psi = \widetilde{C\Pi}_k \xi, \tag{4.22}$$

is a model of system (4.1) at S for all t_k.

The result is a consequence of Definition 12 and Chapter 2.

4.2.4 Linear time-delay systems

The results developed so far can be easily extended to linear time-delay systems renaming the moments of system (3.1) in Chapter 3 as $\overline{C\Pi} = \sum_{j=0}^{\varsigma} C_j \Pi e^{-S\tau_j}$. Theorem 13 holds for the linear time-delay system (3.1) replacing $\widetilde{C\Pi}_k$ with $\widetilde{\overline{C\Pi}}_k$ which is an approximation of $\overline{C\Pi}$ (the proof is a simple exercise).

Definition 13. Consider system (3.1) and the signal generator (4.3). Assume $\sigma(\bar{A}(s)) \subset \mathbb{C}_{<0}$, system (3.1) is minimal and suppose Assumptions 9 and 11 hold. Then the system

$$\dot{\xi} = \sum_{j=0}^{\varrho} F_{j,k} \xi_{\chi_j} + \sum_{j=\varrho+1}^{\rho} G_{j,k} u_{\chi_j}, \qquad \psi = \sum_{j=0}^{d} H_{j,k} \xi_{\chi_j}, \tag{4.23}$$

with $\xi(t) \in \mathbb{R}^\nu$, $\psi(t) \in \mathbb{R}$, $F_{j,k} \in \mathbb{R}^{\nu \times \nu}$ for $j = 0, \ldots, \varrho$, $G_{j,k} \in \mathbb{R}^{\nu \times 1}$ for $j = \varrho + 1, \ldots, \rho$, $H_{j,k} \in \mathbb{R}^{1 \times \nu}$ for $j = 0, \ldots, k$, $\chi_0 = 0$ and $\chi_j \in \mathbb{R}_{>0}$ for $j = 1, \ldots, \max\{\rho, d\}$, is a *model of system (3.1) at* S *at time* t_k, if there exists a unique solution P_k of the equation

$$\sum_{j=0}^{\varrho} F_{j,k} P_k e^{-S\chi_j} - P_k S = -\sum_{j=\varrho+1}^{\rho} G_{j,k} L e^{-S\chi_j}, \tag{4.24}$$

such that

$$\widetilde{\overline{C\Pi}}_k = \sum_{j=0}^{d} H_{j,k} P_k e^{-S\chi_j}, \tag{4.25}$$

where $\widetilde{C\Pi}_k$ is the solution of (4.14).

Proposition 7. Let $\bar{F}_k(s) = \sum_{j=0}^{\varrho} F_{j,k} e^{-s\chi_j}$ and select $P_k = I$ for all $k \geq 0$. If $\sigma(\bar{F}_k(s)) \cap \sigma(S) = \emptyset$ for all $k \geq 0$, then the model

$$\begin{aligned} \dot{\xi} &= \left(S - \sum_{j=\varrho+1}^{\rho} G_{j,k} L e^{-S\chi_j} - \sum_{j=1}^{\varrho} F_{j,k} e^{-S\chi_j} \right) \xi + \sum_{j=1}^{\varrho} F_{j,k} \xi_{\chi_j} + \sum_{j=\varrho+1}^{\rho} G_{j,k} u_{\chi_j}, \\ \psi &= \left(\widetilde{C\Pi}_k - \sum_{j=1}^{d} H_{j,k} e^{-S\chi_j} \right) \xi + \sum_{j=1}^{d} H_{j,k} \xi_{\chi_j}, \end{aligned} \tag{4.26}$$

is a model of system (3.1) at S for all t_k.

The result is a consequence of Definition 13 and Chapter 3.

4.2.5 Properties of the exponentially converging models

In Chapters 2 and 3, see also [3, 10], we have studied the problem of enforcing additional properties and constraints on the reduced order model. In this section we go through these properties to determine if, and under which conditions, they hold for the models (4.22) and (4.26).

Matching with prescribed eigenvalues

Consider system (4.22) and the problem of determining at every k the matrix G_k such that $\sigma(F_k) = \{\lambda_{1,k}, \ldots, \lambda_{\nu,k}\}$ for some prescribed values $\lambda_{i,k}$. The solution of this problem is well-known and consists in selecting G_k such that

$$\sigma(S - G_k L) = \sigma(F_k).$$

This is possible for every k and for all $\lambda_{i,k} \notin \sigma(S)$ and note that G_k is independent from the estimate $\widetilde{C\Pi}_k$. Note also that by observability of (L, S), G_k is unique at every k.

Matching with interpolation at 2ν points

Let $S_a \in \mathbb{R}^{\nu \times \nu}$ and $S_b \in \mathbb{R}^{\nu \times \nu}$ be two non-derogatory matrices such that $\sigma(S_a) \cap \sigma(S_b) = \emptyset$ and let L_{ab} be such that the pairs (L_{ab}, S_a) and (L_{ab}, S_b) are observable. Let $\widetilde{C\Pi}_{a,k} = \widetilde{C\Pi}_k$ computed with (4.14) with $L = L_{ab}$ and $S = S_a$, and let $\widetilde{C\Pi}_{b,k} = \widetilde{C\Pi}_k$ computed with (4.14) with $L = L_{ab}$ and $S = S_b$. Consider system (4.26) with $\varrho = 1$, $\rho = 3$, $d = 1$,

$\chi_2 = 0$ and the problem of determining $F_{1,k}$ and $H_{1,k}$ such that system (4.26) is a model of system (3.1) at S_a and S_b at time t_k. This problem is solved for every k by the selection (see Proposition 4)

$$\begin{aligned} F_{1,k} &= (S_b - S_a - G_{3,k}(e^{-S_b\chi_3} - e^{-S_a\chi_3}))(e^{-S_b\chi_1} - e^{-S_a\chi_1})^{-1}, \\ F_{0,k} &= S_a - G_{2,k}L_{ab} - G_{3,k}L_{ab}e^{-S_a\chi_3} - F_{1,k}e^{-S_a\chi_1}, \\ H_{1,k} &= (\widetilde{C\Pi}_{b,k} - \widetilde{C\Pi}_{a,k})(e^{-S_b\chi_1} - e^{-S_a\chi_1})^{-1}, \\ H_{0,k} &= \widetilde{C\Pi}_{a,k} - H_1 e^{-S_a\chi_1}, \end{aligned} \tag{4.27}$$

which belongs to the family (4.26) for any $G_{2,k}$ and $G_{3,k}$ such that $\sigma(\bar{F}_k(s)) \cap \sigma(S) = \emptyset$.

Matching with prescribed relative degree, matching with prescribed zeros, matching with compartmental constraints

These problems can be solved at each k as detailed in Chapter 2 if and only if

$$\operatorname{rank} \begin{bmatrix} sI - S \\ \widetilde{C\Pi}_k \end{bmatrix} = n, \tag{4.28}$$

for all $s \in \sigma(S)$ at k. Even though the asymptotic value of $\widetilde{C\Pi}_k$ satisfies this condition there is no guarantee that the condition holds for all k. However, if the condition holds for the asymptotic value, there exists $\bar{k} \gg 0$ such that for all $k \geq \bar{k}$ equation (4.28) holds.

4.2.6 Example: A system of order $n = 1010$

In this section we apply Algorithm 1 to a system similar to the one used in Section 3.2.9 (see also [4,5]). The system has order $n = 1010$ and it has Bode plot with five peaks. The state space matrices of system (4.1) are given by $A = \operatorname{diag}(A_1, A_2, A_3, A_4, A_5, \bar{A})$, with

$$A_i = \begin{bmatrix} -1 & a_i \\ -a_i & -1 \end{bmatrix}, \quad \bar{A} = \operatorname{diag}(-1, -2, \ldots, -1000),$$

$a_1 = 50$, $a_2 = 100$, $a_3 = 150$, $a_4 = 200$, $a_5 = 400$, and $B^\top = C = [\underbrace{10 \ldots 10}_{10 \text{ times}} \ \underbrace{1 \ldots 1}_{1000 \text{ times}}]$. The matrices of the signal generator (4.3) have been selected as $S = \operatorname{diag}(0, S_1, S_2, S_3, S_4, S_5, \bar{S})$, with $S_1 = A_1 + I$, $S_2 = 0.5\,S_1$, $S_3 = 1.5\,S_1$, $S_4 = 2\,S_1$,

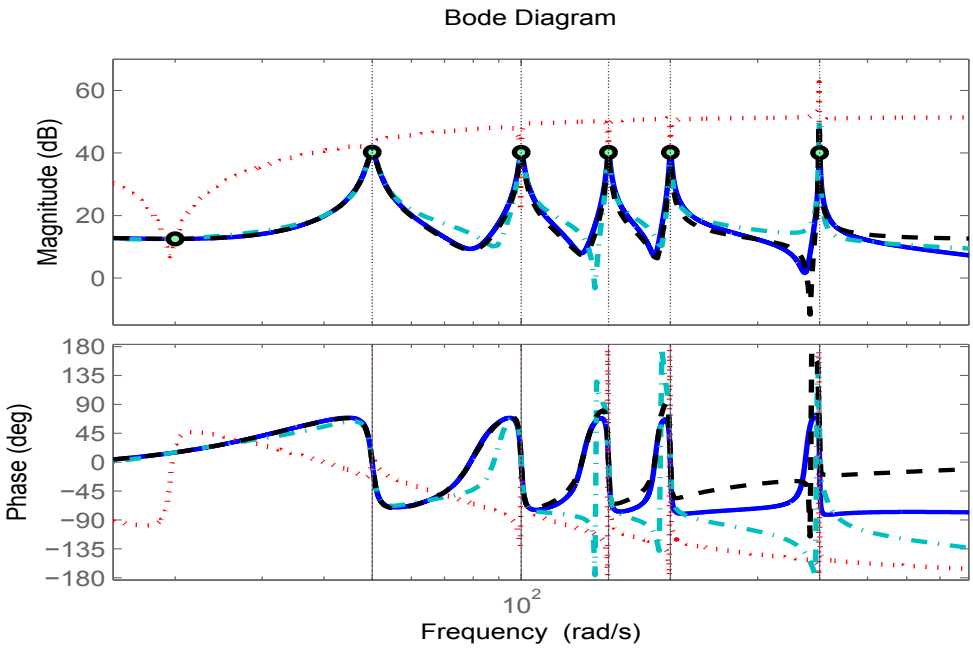

Figure 4.1: Bode plot of the system (solid/blue line), of the reduced order model at $t_{12803} = 8s$ (dotted/red line), of the reduced order model at $t_{22334} = 14s$ (dash-dotted/cyan line) and of the reduced order model at $t_{39873} = 25s$ (dashed/black line). The circles indicate the interpolation points.

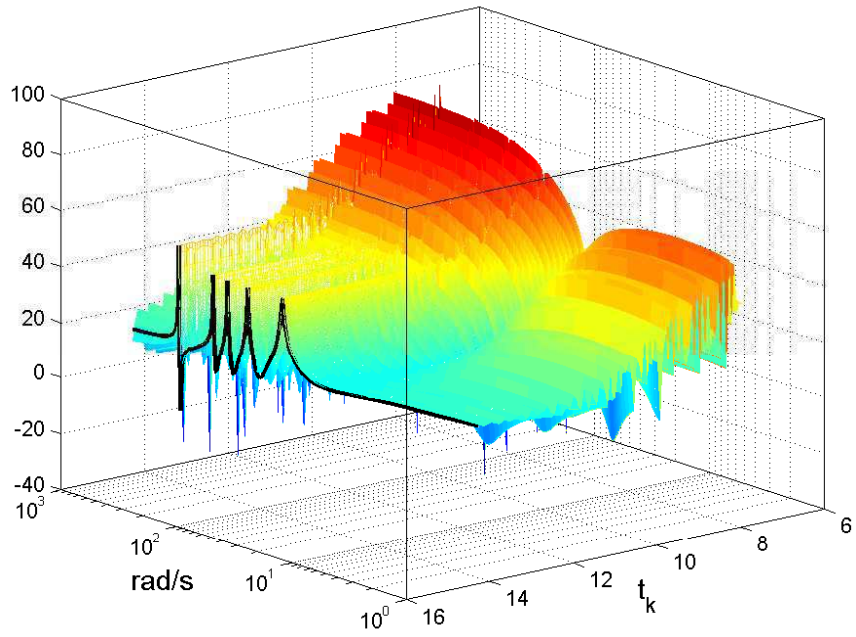

Figure 4.2: The colored mesh represents the magnitude of the transfer function of the reduced order model as a function of t_k, with $8 \leq t_k \leq 14\,s$. The solid/black line indicates the magnitude of the transfer function of the reduced order model for the exact moments $C\Pi$.

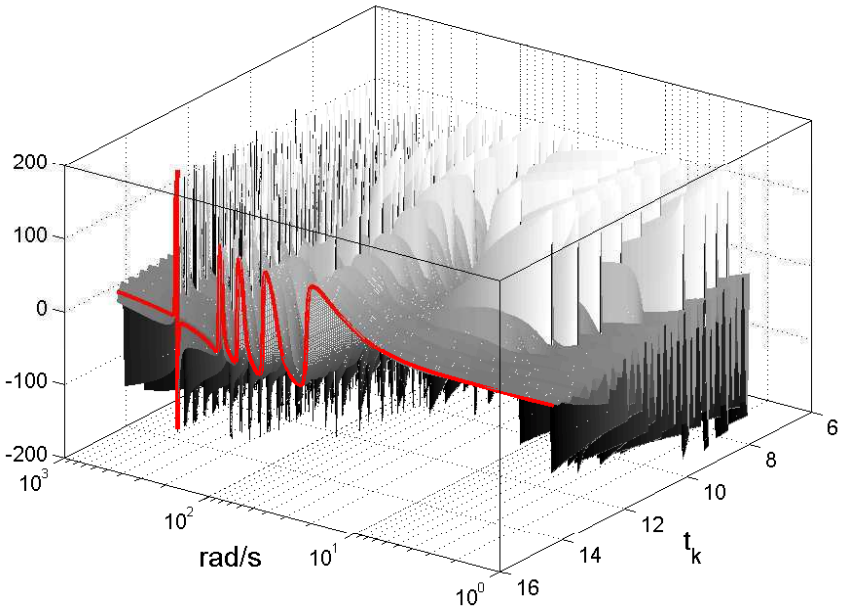

Figure 4.3: The grey mesh represents the phase of the transfer function of the reduced order model as a function of t_k, with $8 \leq t_k \leq 14\,s$. The solid/red line indicates the phase of the transfer function of the reduced order model for the exact moments $C\Pi$.

$S_5 = 4\,S_1$ and

$$\bar{S} = \begin{bmatrix} \frac{1}{5}S_1 & I \\ 0 & \frac{1}{5}S_1 \end{bmatrix},$$

to interpolate the moments at 0 and close to the five frequency peaks.

A reduced order model (4.22) at time t_k has been constructed assigning the eigenvalues of F_k. Fig. 4.1 shows the Bode plot of the system (solid/blue line), of the reduced order model at $t_{12803} = 8s$ (dotted/red line), of the reduced order model at $t_{22334} = 14s$ (dash-dotted/cyan line) and of the reduced order model at $t_{39873} = 25s$ (dashed/black line). Note that the frequencies of interest, indicated with circles, are interpolated already at $t_{12803} = 8s$. We see also that the error between the reduced order model and the system decreases as t_k increases and that at $t_{39873} = 25s$ the frequency responses of the reduced order model and the system match over a wide range of frequencies. The surface in Fig. 4.2 (4.3, respectively) represents the magnitude (phase, respectively) of the transfer function of the reduced order model as a function of t_k, with $8 \leq t_k \leq 14\,s$. The solid/black (solid/red, respectively) line indicates the magnitude (phase, respectively) of the transfer

function of the reduced order model for the exact moments $C\Pi$. The figures show how the approximated magnitude and phase of model (4.22) at S of system (4.1) evolve over time and approach the respective quantities of the exact reduced order model as $t_k \to \infty$. The approximated and actual Bode plot are indistinguishable at $t_k = 25\,s$, but the figures are sliced at $t_k = 14$ to show the detail of the initial convergence, when the Bode plot changes more swiftly. Note also that we have tested different color schemes to ease the individuation of the evolution of the two graphs. This resulted in the two different codings used in Fig. 4.2 and 4.3.

4.3 Nonlinear (time-delay) systems

For the sake of convenience we report here the equations of system (2.20), namely

$$\dot{x} = f(x, u), \qquad y = h(x), \tag{4.29}$$

of the signal generator (2.21), namely

$$\dot{\omega} = s(\omega), \qquad u = l(\omega), \tag{4.30}$$

and the partial differential equation (2.23), namely

$$\frac{\partial \pi}{\partial \omega} s(\omega) = f(\pi(\omega), l(\omega)). \tag{4.31}$$

which will be used through the chapter. The quantities appearing in the equations retain their original meaning.

4.3.1 On-line moment estimation from data

Solving equation (4.31) with respect to the mapping π is a difficult task even when there is perfect knowledge of the dynamics of the system, *i.e.* the mapping f. When f is not known equation (4.31) may be solved numerically requiring information on the state of the system. In practice, this is usually not the case and only the output y is available, with the consequence that also the mapping h has to be known.

Note that given the exponential stability hypothesis on the system and Theorem 3, the

equation

$$y(t) = h(\pi(\omega(t))) + \varepsilon(t), \tag{4.32}$$

where $\varepsilon(t)$ is an exponentially decaying signal, holds. We introduce the following assumption.

Assumption 12. The mapping $h \circ \pi$ belongs to the function space identified by the family of continuous basis functions $\varphi_j : \mathbb{R}^\nu \to \mathbb{R}$, with $j = 1, \ldots, M$ (M may be ∞), *i.e.* there exist $\pi_j \in \mathbb{R}$, with $j = 1, \ldots, M$, such that

$$h(\pi(\omega)) = \sum_{j=1}^{M} \pi_j \varphi_j(\omega),$$

for any ω.

Let

$$\begin{aligned} \Gamma &= \begin{bmatrix} \pi_1 & \pi_2 & \ldots & \pi_N \end{bmatrix}, \\ \Omega(\omega(t)) &= \begin{bmatrix} \varphi_1(\omega(t)) & \varphi_2(\omega(t)) & \ldots & \varphi_N(\omega(t)) \end{bmatrix}^\top, \end{aligned}$$

with $N \leq M$. Using a weighted sum of basis functions, equation (4.32) can be written as

$$y(t) = \sum_{j=1}^{N} \pi_j \varphi_j(\omega(t)) + e(t) + \varepsilon(t) = \Gamma\Omega(\omega(t)) + e(t) + \varepsilon(t), \tag{4.33}$$

where $e(t) = \sum_{N+1}^{M} \pi_j \varphi_j(\omega(t))$ is the error resulting by stopping the summation at N. Consider now the approximation

$$y(t) \approx \sum_{j=1}^{N} \widetilde{\pi}_j \varphi_j(\omega(t)) = \widetilde{\Gamma}\Omega(\omega(t)), \tag{4.34}$$

which neglects the approximation error $e(t)$ and the transient error $\epsilon(t)$. Let Γ_k be an on-line estimate of the matrix Γ computed at T_k^w, namely computed at the time t_k using the last w instants of time t_i assuming that $e(t)$ and $\epsilon(t)$ are known. Since this is not the case in practice, define $\widetilde{\Gamma}_k = \begin{bmatrix} \widetilde{\pi}_1 & \widetilde{\pi}_2 & \ldots & \widetilde{\pi}_N \end{bmatrix}$ as the approximation, in the sense of (4.34), of the estimate Γ_k. Finally we can compute this approximation as follows.

Theorem 16. Define the time-snapshots $\widetilde{U}_k \in \mathbb{R}^{w\times N}$ and $\widetilde{\Upsilon}_k \in \mathbb{R}^w$ as

$$\widetilde{U}_k = \begin{bmatrix} \Omega(\omega(t_{k-w+1})) & \dots & \Omega(\omega(t_{k-1})) & \Omega(\omega(t_k)) \end{bmatrix}^\top$$

and

$$\widetilde{\Upsilon}_k = \begin{bmatrix} y(t_{k-w+1}) & \dots & y(t_{k-1}) & y(t_k) \end{bmatrix}^\top.$$

If $\widetilde{U}_k$ is full rank, then

$$\mathrm{vec}(\widetilde{\Gamma}_k) = (\widetilde{U}_k^\top \widetilde{U}_k)^{-1}\widetilde{U}_k^\top \widetilde{\Upsilon}_k, \tag{4.35}$$

is an approximation of the estimate Γ_k

To ensure that the approximation is well-defined for all k, we give an assumption in the spirit of persistency of excitation.

Assumption 13. For any $k \geq 0$, there exist $\bar{K} > 0$ and $\alpha > 0$ such that the elements of T_k^K, with $K > \bar{K}$, are such that

$$\frac{1}{K}\widetilde{U}_k^\top \widetilde{U}_k \geq \alpha I.$$

Note that if Assumption 13 holds (see [143] for a similar argument), $\widetilde{U}_k^\top \widetilde{U}_k$ is full rank. To ease the notation we introduce the following definition.

Definition 14. The *estimated moment* of system (4.29) is defined as

$$\widetilde{h\circ\pi}_{N,k} = \widetilde{\Gamma}_k \widetilde{U}_k, \tag{4.36}$$

with $\widetilde{\Gamma}_k$ computed with (4.35).

An equivalent of Theorem 15 can be formulated in the nonlinear framework. The recursive least-square algorithm is obtained with $\Omega(\omega(t_k))$ playing the role of $\omega(t_k)$, $\widetilde{U}_k$ playing the role of $\widetilde{R}_k$ and $\widetilde{\Gamma}_k$ playing the role of $\widetilde{C\Pi}_k$, namely

$$\begin{aligned}\mathrm{vec}(\widetilde{\Gamma}_k) = \mathrm{vec}(\widetilde{\Gamma}_{k-1}) &+ \Phi_k\Omega(\omega(t_k))\Big(y(t_k) - \Omega(\omega(t_k))^\top \mathrm{vec}(\widetilde{\Gamma}_{k-1})\Big) \\ &-\Phi_k\Omega(\omega(t_{k-w}))\Big(y(t_{k-w}) - \Omega(\omega(t_{k-w}))^\top \mathrm{vec}(\widetilde{C\Pi}_{k-1})\Big),\end{aligned} \tag{4.37}$$

with $\Phi_k = (\widetilde{U}_k^\top \widetilde{U}_k)^{-1}$ and $\Psi_k = (\widetilde{U}_{k-1}^\top \widetilde{U}_{k-1} + \Omega(\omega(t_k))\Omega(\omega(t_k))^\top)^{-1}$.

Remark 28. Similarly, Algorithm 1 can be adapted to the present scenario, with straightforward variations, to determine the approximation $\widetilde{h\circ\pi}_{N,k}$. ■

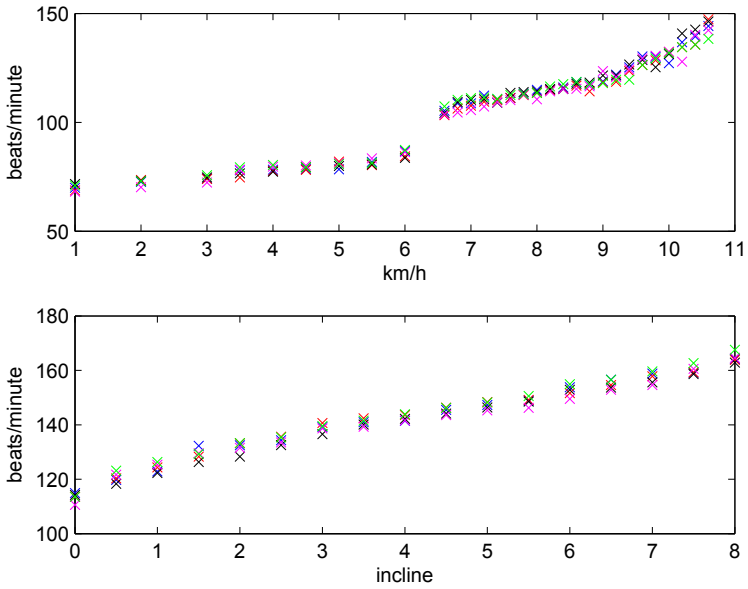

Figure 4.4: Steady-state average heartbeat for five iteration of the experiment for velocities from 1 to 10.5 km/h (top graph) and inclines from 0 to 8 (bottom graph).

Theorem 17. Suppose Assumptions 1, 2, 12 and 13 hold. Then

$$\lim_{t\to\infty}\left(h(\pi(\omega(t))) - \lim_{N\to M}\widetilde{h\circ\pi}_{N,k}(\omega(t))\right) = 0.$$

Proof. Assumption 13 guarantees that the approximation $\widetilde{\Gamma}_k$ is well-defined for all k, whereas Assumptions 1 and 2 guarantee that Theorem 3 holds and thus that $h\circ\pi$ is well-defined. The quantity $||\varepsilon(t_k)||$ vanishes exponentially to zero by Assumption 2. Hence, by Assumption 12, $\lim_{N\to M}\widetilde{h\circ\pi}_{N,k}(\omega(t))$ converges to $h(\pi(\omega(t)))$. □

Remark 29. The framework presented here to approximate the moment of nonlinear systems exploits the estimation techniques based on basis functions. However, note that the results of the following sections hold also when the estimation techniques used to approximate $h\circ\pi$ are different. In fact, the approach based on basis functions should be considered as an example used to formalize the results of the chapter. ■

4.3.2 Experiment: nonlinear moment of the average heartbeat model

An experiment has been carried out to illustrate how to estimate the moment of a nonlinear system in a practical example. A test subject ran on a treadmill starting from a resting

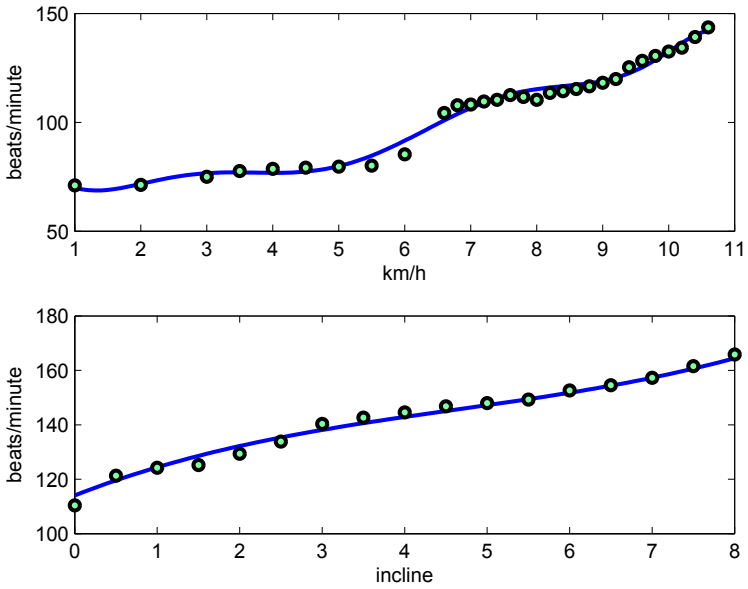

Figure 4.5: Estimated moment of the average heartbeat (solid line) and validation dataset (circles) for the velocity experiment (top graph) and for the incline experiment (bottom graph).

position. In the first experiment a certain constant speed (from 1 to 10.5 km/h) has been set and the subject ran at that speed for a period of time sufficient to make his average heartbeat constant. This value corresponds to the steady-state response of his average heartbeat for a constant input speed, *i.e.* $\dot{\omega} = 0$. The experiment has been iterated five times and the measured steady-state average heartbeat values for different constant velocities are shown in the top graph of Fig. 4.4. In the second experiment, the incline (from option 0 to 8) of the treadmill has been considered as a constant input (for all the incline options the speed has been kept constant to 8 km/h). The steady-state average values of the heartbeat are shown in the bottom graph of Fig. 4.4. Also this experiment has been iterated five times. A polynomial fitting of the data gives the estimated moment

$$\widetilde{h\circ\pi}_{9,\eta}(\omega) = -0.0003578\omega^8 + 0.01431\omega^7 - 0.2233\omega^6 + 1.682\omega^5 - 5.852\omega^4 + 4.429\omega^3 + 24.44\omega^2 - 54.66\omega + 100.2,$$

for the velocity experiment and

$$\widetilde{h\circ\pi}_{4,\eta}(\omega) = 0.1169\omega^3 - 1.63\omega^2 + 11.86\omega + 114.1,$$

for the incline experiment. Note that the basis functions for the first experiment are the polynomials of order eight, whereas the basis functions for the second experiment are the polynomials of order three. Instead of using (4.35), the fitting has been performed with the function *fit* of MATLAB$^{\text{TM}}$ directly using the steady-state values. Fig. 4.5 shows the estimated moment in blu/solid line and a validation set of data, indicated with circles, taken from a sixth iteration of the experiment for the velocity experiment (top graph) and the incline experiment (bottom graph). Then the estimated moment can be used to obtain a reduced order model as defined in the following section.

4.3.3 Families of reduced order models

Using the approximation given by (4.36) a reduced order model of system (4.29) can be defined at each instant of time t_k.

Definition 15. Consider system (4.29) and the signal generator (4.30). Suppose Assumptions 1, 2, 12 and 13 hold. Then the system

$$\dot{\xi} = \phi_k(\xi, u), \qquad \psi = \kappa_k(\xi), \tag{4.38}$$

with $\xi(t) \in \mathbb{R}^\nu$, $u(t) \in \mathbb{R}$, $\psi(t) \in \mathbb{R}$ and ϕ_k and κ_k smooth mappings, is a *model of system (4.29) at* (s, l) *at time* t_k, *i.e.* system (4.38) has the same moment of system (4.29) at (s, l), if the equation

$$\frac{\partial p_k}{\partial \omega} s(\omega) = \phi_k(p_k(\omega), l(\omega)), \tag{4.39}$$

has a unique solution p_k such that

$$\widetilde{h\circ\pi}_{N,k}(\omega) = \kappa_k(p_k(\omega)), \tag{4.40}$$

where $\widetilde{h\circ\pi}_{N,k}(\omega)$ is obtained by (4.35).

Proposition 8. Let δ_k be such that, for all $k \geq 0$,

$$\frac{\partial p_k}{\partial \omega} s(\omega) = s(p_k(\omega)) - \delta_k(p_k(\omega))l(p_k(\omega)) + \delta_k(p_k(\omega))l(\omega),$$

has the unique solution $p_k(\omega) = \omega$ and select $\kappa_k(\omega) = \widetilde{h \circ \pi}_{N,k}(\omega)$. Then the system described by the equations

$$\begin{aligned} \dot{\xi} &= s(\xi) - \delta_k(\xi)l(\xi) + \delta_k(\xi)u, \\ \psi &= \widetilde{h \circ \pi}_{N,k}(\xi), \end{aligned} \tag{4.41}$$

is a model of system (4.29) at (s, l) for all t_k.

The result is a consequence of Definition 15 and Chapter 2.

Remark 30. Similarly to the linear case (see Section 4.2.5), the conditions given in Chapter 2 to enforce additional properties upon the reduced order model can be adapted to hold in the present scenario. ■

4.3.4 Nonlinear time-delay systems

The results developed so far can be extended to nonlinear time-delay systems. In fact, since the output $y(t)$ of system (3.32) can be described by the equation (4.33), Theorem 17 holds for the nonlinear time-delay system (3.32). For ease of notation we consider only one delay in the state and one delay in the input.

Definition 16. Consider system (3.32) and the signal generator (4.30). Suppose Assumptions 1, 6, 12 and 13 hold. Then the system

$$\dot{\xi} = \phi_k(\xi, \xi_{\chi_1}, u, u_{\chi_2}), \qquad \psi = \kappa_k(\xi), \tag{4.42}$$

with $\xi(t) \in \mathbb{R}^\nu$, $u(t) \in \mathbb{R}$, $\psi(t) \in \mathbb{R}$, $\chi_j \in \mathbb{R}_{>0}$ with $j = 1, 2$ and ϕ_k and κ_k smooth mappings, is a *model of system (3.32) at* (s, l) *at time* t_k, *i.e.* system (4.42) has the same moment of system (3.32) at (s, l), if the equation

$$\frac{\partial p_k}{\partial \omega} s(\omega) = \phi_k(p_k(\omega), p_k(\bar{\omega}_{\chi_1}), l(\omega), l(\bar{\omega}_{\chi_2})), \tag{4.43}$$

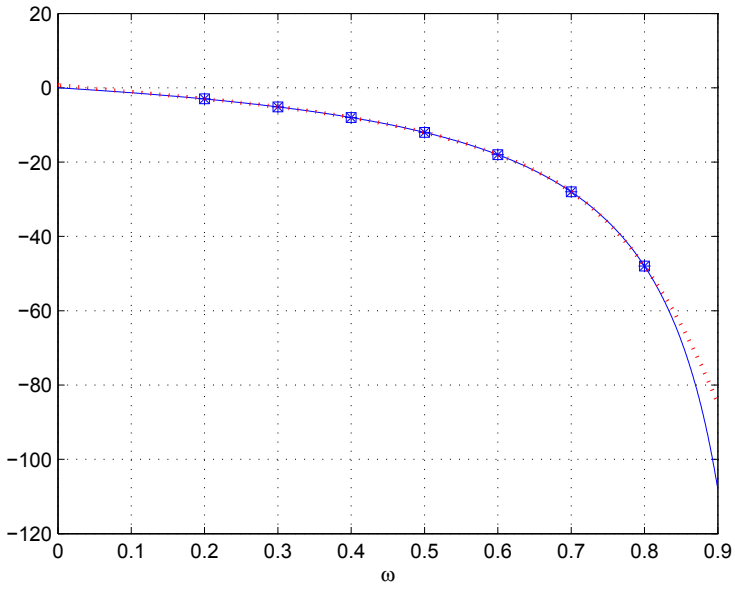

Figure 4.6: The functions $h(\pi(\omega))$ (solid line) and $\widetilde{h\circ\pi}(\omega)$ (dotted line) for $\omega \in (0, 0.9\,]$. The seven data points are represented by squares.

has a unique solution p_k such that

$$\widetilde{h\circ\pi}_{N,k}(\omega) = \kappa_k(p_k(\omega)), \tag{4.44}$$

where $\widetilde{h\circ\pi}_{N,k}(\omega)$ is obtained by (4.35).

4.3.5 Example: approximated nonlinear model of the Ćuk converter

In this section we illustrate the results of the chapter by means of two examples. Both examples are based on the averaged model of the DC-to-DC Ćuk converter [156]. We begin obtaining a scalar reduced order model which achieves moment matching at zero. In this case the exact expression of the mapping $h\circ\pi$ is known and we can compare it directly with the approximation obtained from a polynomial expansion. In the second part of the section we obtain a planar reduced order model with nonlinear state dynamics using an input generated by a nonlinear mapping l. In this case the exact expression of the mapping π is not known and the results of the chapter are used to obtain an approximation of the mapping $h \circ \pi$.

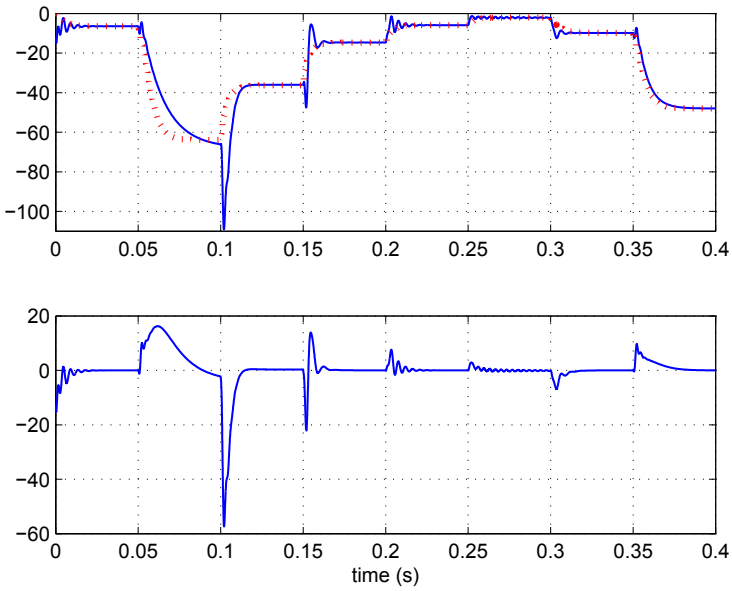

Figure 4.7: Top: time histories of the output of system (4.45) (solid line) and of the reduced order model (dotted line) for the set of input given by ω equal to 0.35, 0.85, 0.75, 0.55, 0.33, 0.15 and 0.45. Bottom: error between the two outputs.

Table 4.1: Parameters of the model of the DC-to-DC Ćuk converter [156]

$L_1 = 10\ mH$	$L_3 = 10\ mH$	$C_2 = 22\ \mu F$
$C_4 = 22\ \mu F$	$E = 12\ V$	$G = 0.0447\ S$

The averaged model of the DC-to-DC Ćuk converter is given by the equations [156]

$$\begin{aligned} L_1\dot{x}_1 &= -(1-u)x_2 + E, & C_2\dot{x}_2 &= (1-u)x_1 + ux_3, \\ L_3\dot{x}_3 &= -ux_2 - x_4, & C_4\dot{x}_4 &= x_3 - Gx_4, \\ y &= x_4, \end{aligned} \tag{4.45}$$

where $x_1(t) \in \mathbb{R}_{\geq 0}$ and $x_3(t) \in \mathbb{R}_{\leq 0}$ describe the currents, $x_2(t) \in \mathbb{R}_{\geq 0}$ and $x_4(t) \in \mathbb{R}_{\leq 0}$ the voltages, L_1, C_2, L_3, E and G positive parameters and $u(t) \in (0, 1)$ a continuous control signal which represents the slew rate of a pulse width modulation circuit used to control the switch position in the converter. In the remaining of the chapter we used the numerical values given in [156], reported in Table 4.1, to simulate system (4.45).

Moment at zero

In [3], a scalar reduced order model of system (4.45) achieving moment matching at zero has been given. It has been shown that the moment of the system is described by

$$h(\pi(\omega)) = E\frac{\omega}{\omega - 1}.$$

We simulated system (4.45) with $u = \omega(0)$, where $\omega(0)$ switched every 0.05s between the values of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7 and 0.8. We have then applied the results of the chapter selecting the horizon length w equal to 1 and fixing the value of $\widetilde{h \circ \pi}(\omega)$ to the one just before the switching time. Then a cubic interpolation has been used to fit the seven points obtained from the simulations. Fig. 4.6 shows the function $h(\pi(\omega))$ (solid line) and the approximation $\widetilde{h \circ \pi}(\omega)$ (dotted line) for values of $\omega \in (0, 0.9\,]$. The seven data points are represented by squares. Note that the approximation is close to the actual moment of the system for $\omega \in (0, 0.85]$.

The reduced order model is chosen as a linear system of order $\nu = 1$ with eigenvalue equal to -230 (to yield a reduced order model with "response time" comparable to the one of the system). The top plot in Fig. 4.7 shows the time histories of the output of system (4.45) (solid line) and of the reduced order model (dotted line) for the set of inputs given by ω equal to 0.35, 0.85, 0.75, 0.55, 0.33, 0.15 and 0.45. Note that no one of these values is one of the seven data points. The error between the two outputs is shown in the bottom graph of Fig. 4.7. The figure shows an overall good approximation of system (4.45). The error is comparable with the one shown in the simulation in [3] and it is due to the transient error $\varepsilon(t)$ caused by approximating a four dimensional nonlinear system with a scalar model.

Planar reduced order model

In the previous example the approximated moment of the system has been compared with the exact mathematical expression. However, the example is very simple since the generator is scalar and the input is constant. In this section we present the case in which the input is generated by the equations

$$\dot{\omega}_1 = -75\omega_2, \qquad \dot{\omega}_2 = 75\omega_1, \qquad u = \max\left(0.15, \frac{1}{2}\left(\omega_1\omega_2 + \omega_1^3 + \frac{1}{2}\right)\right).$$

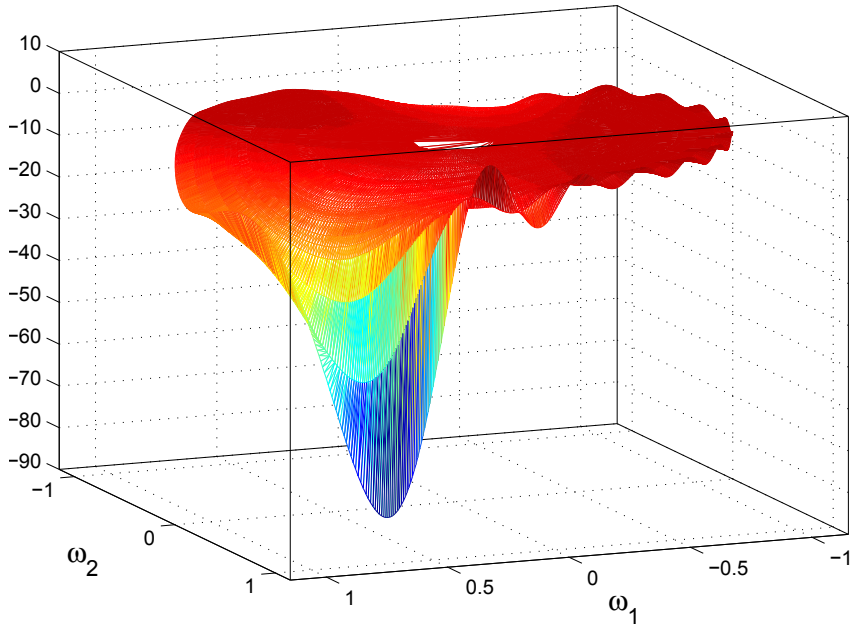

Figure 4.8: Surface obtained from the data points computed with $\omega(t)$ and $y(t)$.

Note that this produces a positive input signal with higher order harmonics.

Simulations have been carried out with a few values of $\omega(0)$ selected in the set $[0.75, 0.75] \times [0.75, 0.75]$, which generate inputs u in the set $[0.15, 0.93]$. For each simulation, once reached the steady-state, the time-history of $y(t)$ has been extracted over one period and a map between $y(t)$ and each pair $(\omega_1(t), \omega_2(t))$ has been extrapolated. Then the simulation has been repeated for other values of $\omega(0)$. Fig. 4.8 shows the extrapolated $\widetilde{h \circ \pi}(\omega)$. An interesting area is the foremost/blue region of the surface, where the input u is equal to 0.93. This is close to 1: for this value the output of the converter becomes negatively unbounded.

We use a two-dimensional polynomial fitting of order five to obtain an analytical expression of the moment for $\omega(0) \in [0.6, 0.6] \times [0.6, 0.6]$, which generate inputs u in the set[2]

[2] A polynomial fitting of order five does not properly approximate the moment in the entire region of Fig. 4.8. However, this is the highest order currently available in MATLAB™ (function *fit* with the option *poly25*). Since the coding of a better fitting function is out of the scope of this Thesis, we have restricted the interpolation to a smaller area.

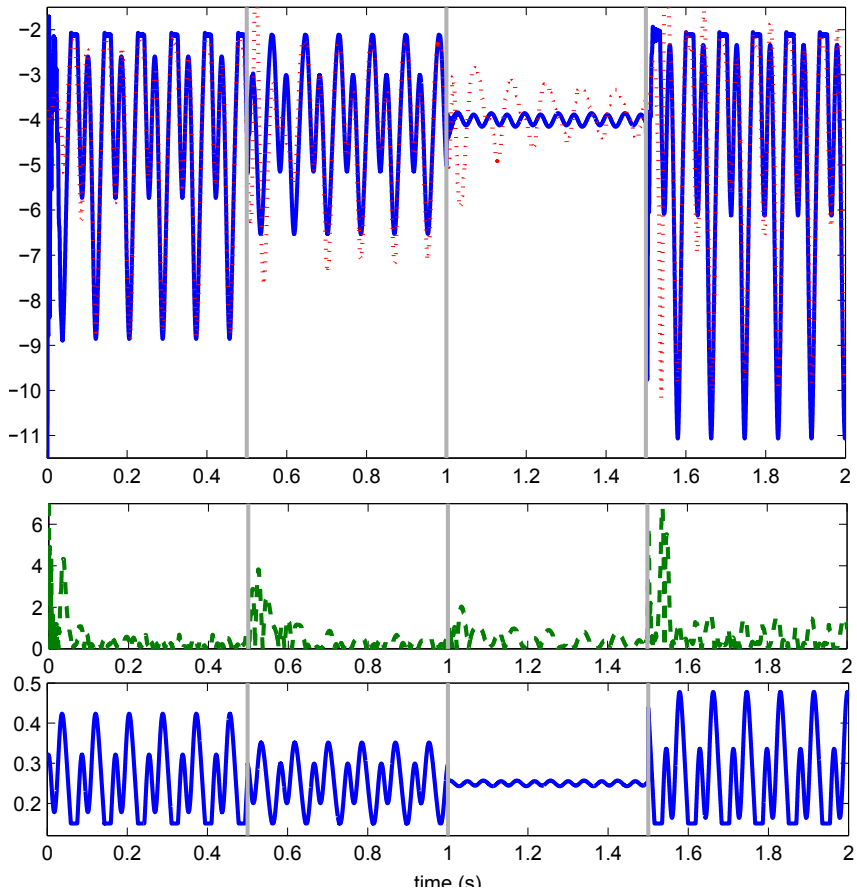

Figure 4.9: Top graph: time histories of the output of system (4.45) (solid/blue line) and of the nonlinear reduced order model (dotted/red line) for the input sequence represented in the bottom graph. The switching times are indicated by solid/gray vertical lines. Middle graph: absolute error (dashed/green line) between the two outputs.

$[0.15, 0.62]$, yielding

$$\begin{aligned}\widetilde{h \circ \pi}(\omega)_{15,\eta} = \; & -3.781 - 3.116\omega_1 + 0.641\omega_2 - 1.71\omega_1^2 - 9.964\omega_1\omega_2 - 1.243\omega_2^2 - 12.23\omega_1^2\omega_2 \\ & +5.185\omega_1\omega_2^2 - 1.764\omega_2^3 - 6.876\omega_1^2\omega_2^2 - 3.026\omega_1\omega_2^3 + 0.8862\omega_2^4 + 14.63\omega_1^2\omega_2^3 \\ & -1.709\omega_1\omega_2^4 + 1.173\omega_2^5. \end{aligned} \tag{4.46}$$

This polynomial approximation fits well the red region of the surface in Fig. 4.8, whereas in the blue region it does not decrease as fast as the actual output of the system. This suggests that the following results can be improved if another fitting function is used.

The reduced order model is chosen as in Proposition 8, with $\delta_\eta = 220 \begin{bmatrix} 1 & 1 \end{bmatrix}^\top$, namely

$$\begin{aligned}
\dot{\xi}_1 &= 75\xi_2 + 220\left(u - \max\left(0.15, \frac{1}{2}\left(\xi_1\xi_2 + \xi_1^3 + \frac{1}{2}\right)\right)\right), \\
\dot{\xi}_2 &= -75\xi_1 + 220\left(u - \max\left(0.15, \frac{1}{2}\left(\xi_1\xi_2 + \xi_1^3 + \frac{1}{2}\right)\right)\right), \\
\psi &= \widetilde{h\circ\pi}(\xi)_{15,\eta}.
\end{aligned}$$

The top graph in Fig. 4.9 shows the time histories of system (4.45) (solid/blue line) and of the reduced order model (dotted/red line) for the input sequence represented in the bottom graph. The input is obtained switching $\omega(0)$ every 0.5s (the switching times are indicated by solid/gray vertical lines). $\omega(0)$ takes, in order, the values of $(-0.45, -0.45)$, $(-0.25, -0.45)$, $(0.15, 0.05)$ and $(0.5, 0.5)$. The middle graph in Fig. 4.9 shows the absolute error (dashed/green line) between the two outputs. We note that the larger absolute error is in the third and fourth simulation. In the third, the error is due to poor transient performance and the problem could be alleviated with a selection of δ_η as a function of ξ. The poor approximation for the fourth input is caused by the fact that the input signal lives at the edge of the area approximated by (4.46), where $\widetilde{h\circ\pi}(\omega)$ is not well-fitted.

4.4 Conclusion

We have presented a theoretical framework and a collection of techniques to obtain reduced order models by moment matching from input/output data for linear (time-delay) systems and nonlinear (time-delay) systems. The approximations proposed in the chapter have been exploited to construct families of reduced order models. We have shown that these models asymptotically match the moments of the unknown system to be reduced. Conditions to enforce additional properties upon the reduced order model have been discussed. The use of the algorithm is illustrated by several examples, experimental and simulated.

Chapter 5

Model reduction of power systems with preservation of physical properties

5.1 Introduction

Since the mathematical models used to describe power systems can easily have hundreds of states, inputs and outputs, the simulation of power systems for dynamic analysis, trajectory sensitivity analysis and control design is a computationally intensive task. Hence, the model reduction problem, some times also referred to as the problem of *dynamic equivalencing* in the power and energy community [157], is central to modern research on power systems. At the basis of the use of dynamic equivalencing in power systems there is the idea of distinguishing between a *study area*, the description of which is maintained in full detail, and an *external area*, consisting of the remaining part of the network, which is reduced. The study area can be precisely analyzed and controlled while still considering the interaction of the interconnection of such area with a less faithful representation of a larger external area. Thus, when doing model reduction of power systems the main aim is not to obtain the best possible approximation, *e.g.* minimizing an error norm on the frequency response of the system, but having a dynamic behavior of the interconnection between the reduced order model and the study area as close as possible to the one of the original nonlinear power system. In other words, we are interested in the faithful

reproduction of the behavior of the system for a specific class of input signals, neglecting the behavior outside the operating conditions.

Historically, coherency-based methods have been used in model reduction of power systems, see *e.g.* [158–161]. These methods are based on the physical properties of the electrical machines connected to the network. The idea is to find coherent generators, *i.e.* machines which behave similarly when the same input is applied. Once coherent generators are identified a dynamic equivalent generator is used to replace them. In recent years, the power systems community has started to be interested in reduction techniques based on mathematical properties instead of physical ones. One of the reasons of this interest is the flexibility of having a reduction technique that is not based on the physics of the generators and, as a consequence, the possibility of reducing networks with renewable energy sources. Among these methods, balanced truncation and Krylov projectors have been successfully used in power systems reduction, see *e.g.* [157, 162, 163]. However, the real applicability of the reduced order models obtained with mathematical methods is limited by the fact that the preservation of some physical characteristics, such as some specific modes, is of paramount importance for applications in power systems. Poorly damped modes, also called electromechanical modes [164], are important in the small-signal stability analysis of a power system since they are responsible for most of the oscillating behavior. Similarly, slow modes characterize how fast the response of the system reaches the steady-state. While previous attempts to maintain these modes are essentially *ad hoc*, since with the classical Krylov methods it is very hard to preserve a certain set of eigenvalues, the problem of assigning a prescribed set of modes has been solved with the approach reported in Chapter 2, see also [3].

The first contribution of this chapter is the validation of the model reduction techniques presented in Chapter 2 and the study of the behavior of the obtained reduced order models. In particular, the methods have been used to assign arbitrary eigenvalues to the reduced order model, *i.e.* maintaining slow and poorly damped modes. In this chapter it is shown that it is not necessary to increase the order of the reduced order model to improve the quality of the approximation of the system: to achieve this goal it may be sufficient to select a different set of eigenvalues to be preserved in the reduced order model. More-

over, it is also shown that increasing the order of the reduced model gives no guarantee of improving significantly the quality of the approximation if a wrong set of eigenvalues is preserved.

The second contribution of the chapter consists in extending the techniques given in Chapters 2 and 4 to multi-input, multi-output (MIMO) systems. Power systems are not only described by a large number of states, they also have a high number of inputs and outputs. To deal with this, we discuss the so-called tangential interpolation problem which is used in the model reduction of MIMO systems. An algorithm to approximate the tangential directions is given and connections with the literature are drawn. Since classical efficient Krylov algorithms suffer the usual drawback of not preserving specific modes, we present a low computational complexity algorithm for the model reduction of MIMO systems which is able to preserve specific properties. This algorithm is a MIMO version of the algorithm presented in Chapter 4. The algorithm, exploiting the inputs and outputs of the system instead of operations on large matrices, offers a computationally efficient method to approximate MIMO reduced order models. The simulation section illustrates the performance of the algorithm, the design of the reduced order model and the dynamic behavior of the interconnection of the study area with the approximated model.

The rest of the chapter is organized as follows. In Section 5.2 we introduce the problem of the reduction of MIMO systems. In Section 5.3 we derive the algorithm for the approximation of the moments of MIMO systems. In Section 5.4.1 we describe the nonlinear model which represents the study area of the power system and the linear model which represents the external area. In Section 5.4.2 the NETS-NYPS benchmark system is described. In Section 5.4.3 an heuristic to approximate the tangential directions is given. In Section 5.4.4 the reduced order model is designed. In Section 5.4.5 the use of the algorithm is illustrated and in Section 5.4.6 the fault response of the interconnection between the study area and the reduced order model is simulated. In this section we also perform a study of the importance of preserving the correct set of modes versus increasing the order of the reduced order model.

All the results of this chapter are original contributions developed in fulfillment of my PhD course and they have been published in the conference paper [14] and in the journal

paper [15].

5.2 Model Reduction of MIMO Systems

Consider a linear MIMO, continuous-time, system described by the equations

$$\dot{x} = Ax + Bu, \qquad y = Cx, \tag{5.1}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^p$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$. Let $W(s) = C(sI - A)^{-1}B : \mathbb{C} \mapsto \mathbb{C}^{p\times m}$ be the associated transfer function and assume that (5.1) is minimal. If we characterize the moments of this MIMO system as in Definition 1, we have $mp\nu$ interpolation conditions (the mp components of W at ν points), whereas the dimensions of $C\Pi$ is only $p\nu$. While a possible solution to this problem is to inflate the order of the reduced order model, this substantially increases the dimension of the model and the reduction may even no longer make sense if $m\nu \geq n$, which is a likely situation in the analysis of power systems. A workaround to this drawback consists in "merging" some of the conditions together. However, the merging of the interpolation conditions has to be done in a specific way to preserve the information of the original individual moments. This is the idea behind the tangential interpolation approach, initially proposed in [165]. The tangential interpolation approach can be embedded in the framework we have presented with the following definition.

Definition 17. Let $s_i \in \mathbb{C} \setminus \sigma(A)$ and $l_i \in \mathbb{R}^{m\times 1}$. The *k-moment of system (5.1) at* s_i *along* l_i is the complex number

$$\eta_k(s_i) = \frac{(-1)^k}{k!}\left[\frac{d^k}{ds^k}C(sI - A)^{-1}B\right]_{s=s_i} l_i,$$

with $k \geq 0$ integer.

Hence, the moment matching conditions become [63, 120]

$$\frac{(-1)^k}{k!}\left[\frac{d^k}{ds^k}W(s)\right]_{s=s_i} l_i = \frac{(-1)^k}{k!}\left[\frac{d^k}{ds^k}\hat{W}(s)\right]_{s=s_i} l_i, \tag{5.2}$$

with $i = 1, \ldots, \nu$, where $\hat{W}(s)$ is the transfer function of the reduced order model.

Remark 31. The $p\nu$ tangential interpolation conditions (5.2) are weaker than the original $pm\nu$ matching conditions. Since this approach consists in replacing m equations with one, the resulting model is not in general as good as a model which interpolates all the m conditions. This is the price to pay to maintain the order of the reduced order model independent from the number of inputs. Moreover, this suggests that the selection of the vectors l_i has to be done with care to obtain a reliable reduced order model. ■

Remark 32. In [162] the authors present the block Krylov approach of [166]. This approach, predating the development of the tangential interpolation theory in [165], inflates the dimension of the reduced order model to satisfy the $pm\nu$ matching conditions. However, the authors have recognized this drawback and, in fact, they do not apply this method in the application part of the paper. In fact, claiming "*in order to limit the size of the Krylov subspaces, we consider that the matrix B [...] is the sum of the input matrices*" they are applying a single input algorithm on an approximated system. They have also pointed out that "*in general, such kind of heuristics for economy in the size of the base [...] does not work well*". Interestingly, a theoretical explanation for the poor general performance can be given revisiting their approximated approach in the tangential framework. It can be easily shown that adding the columns of the matrix B corresponds to using moment matching with tangential directions $l_i = [1 \ \dots \ 1]^\top$. Thus, the performance can be improved selecting different directions. ■

Exploiting Definition 17, Lemma 2 can be adapted to the MIMO case. Let $S \in \mathbb{R}^{\nu\times\nu}$ be a non-derogatory matrix and $L = [l_1\, l_2\, \dots\, l_\nu] \in \mathbb{R}^{m\times\nu}$, $l_i \in \mathbb{R}^{m\times 1}$, $i = 1, \dots, \nu$, be such that the pair (L, S) is observable. Then the moments of the system along L are in one-to-one relation with $C\Pi$, with $\Pi \in \mathbb{R}^{n\times\nu}$ the unique solution of the Sylvester equation

$$A\Pi + BL = \Pi S. \tag{5.3}$$

Finally, the family of systems [3]

$$\dot{\xi} = (S - GL)\xi + Gu, \qquad \psi = C\Pi\xi, \tag{5.4}$$

with $S - GL \in \mathbb{R}^{\nu\times\nu}$, $G \in \mathbb{R}^{\nu\times m}$ and $C\Pi \in \mathbb{R}^{p\times\nu}$ contains all the models of dimension ν interpolating the moments of system (5.1) at S along L if G is such that $\sigma(S) \cap \sigma(S-GL) =$

$\emptyset$. Hence, we say that *system (5.4) is a model of (5.1) at S along L*. System (5.4) is a *reduced order model of system (5.1) at S along L* if $\nu < n$.

Remark 33. The family of systems (5.4) represents reduced order models of system (5.1) for any matrix L, *i.e.* for any tangential direction. However, as already remarked, the quality of the approximation depends on the choice of L. We also note that the tangential directions depend upon the reduced order model and the reduced order model depends upon the tangential directions. Thus, the determination of the directions is a difficult problem. ■

Remark 34. In the interpolation framework, algorithms have been proposed to approximate iteratively the vectors l_i, see *e.g.* [44]. However, these solutions do not allow to select the interpolation points nor to preserve a prescribed set of eigenvalues. While these algorithms may work when specific physical properties are not important, in the case of power systems the preservation of slow and poorly damped modes is important for giving a physical meaning to the reduced order models [14]. ■

5.3 A low complexity MIMO algorithm for the computation of the moments

In this section we extended to MIMO systems Algorithm 1 presented in Chapter 4 for the computation of the moments of a SISO system from input/output data. Although the algorithm has been primarily devised to compute the moments when the matrices A, B, C are not available, it has also the advantage of being a computationally fast method for the approximation of the moments of a system. In fact, as already pointed out, for SISO systems the algorithm has a computational complexity of $\mathcal{O}(g\nu^{2.373})$ (g is a small scalar that will be defined later), whereas the Arnoldi or Lanczos procedures for the approximation of the moments have a computational complexity of $\mathcal{O}(\nu n^2)$ (both complexities are multiplied by p in the MIMO case).

Remark 35. We do not claim that the algorithm we are presenting here can be used to approximate the moments from actual input/output data of power systems. In fact, for a power system, which is a nonlinear system, we do not have any guarantee on the approximation given by the algorithm when measurements generated by such a nonlinear system are used. However, in addition to propose this algorithm for the fast computation

of the moments, this chapter lays also the foundations to solve the problem of input/output model reduction from real data in two ways: the input/ouput data of a power system can be filtered by means of a linear filter; the nonlinear version of the algorithm proposed in Chapter 4 can be extended to MIMO systems and used with the measured data. However, the use of these methods for the reduction of power systems is not trivial and it would deserve its own dedicated research. For the aims of this chapter, we limit ourself to use the algorithm as a mean to approximate the moments without solving equation (5.3), which is computationally expensive. ■

Theorem 18. Consider the interconnection of system (5.1) with the signal generator

$$\dot{\omega} = S\omega, \qquad u = L\omega. \tag{5.5}$$

Assume $\sigma(A) \subset \mathbb{C}_{<0}$, $\sigma(S) \subset \mathbb{C}_0$ and that the triple $(L, S, \omega(0))$ is minimal. Let $0 \leq t_0 < t_1 < \cdots < t_{k-h} < \cdots < t_k < \cdots < t_q$, with $h > 0$ and $q \geq h$ and $T_k^h = \{t_{k-h+1}, \ldots, t_{k-1}, t_k\}$ and assume that the elements of T_k^ν are such that $\operatorname{rank}\left(\begin{bmatrix} \omega(t_{k-\nu+1}) & \ldots & \omega(t_k) \end{bmatrix}\right) = \nu$ for all k. Define the time-snapshots $\widetilde{R}_k \in \mathbb{R}^{g\nu \times \nu}$ and $\widetilde{\Upsilon}_k^j \in \mathbb{R}^{g\nu}$ as

$$\widetilde{R}_k = \begin{bmatrix} \omega(t_{k-w+1}) & \ldots & \omega(t_{k-1}) & \omega(t_k) \end{bmatrix}^\top$$

and

$$\widetilde{\Upsilon}_k^j = \begin{bmatrix} y^j(t_{k-w+1}) & \ldots & y^j(t_{k-1}) & y^j(t_k) \end{bmatrix}^\top,$$

where $y^j(t_i)$ is the j-th row of $y(t_i)$. If $g = 1$, the matrix $\widetilde{R}_k$ has full rank and

$$\operatorname{vec}(\widetilde{C^j\Pi}_k) = (\widetilde{R}_k^\top \widetilde{R}_k)^{-1} \widetilde{R}_k^\top \widetilde{\Upsilon}_k^j, \tag{5.6}$$

is an approximation of $C^j\Pi$, with C^j the j-th row of C, *i.e.* there exists a sequence $\{t_k\}$ such that

$$\lim_{k\to\infty} \widetilde{C^j\Pi}_k = C^j\Pi.$$

Proof. The matrix Π defined in equation (5.3) is such that the condition

$$\dot{x}(t)|_{t_k} = \Pi S\omega(t_k) + Ae^{At_k}(x(0) - \Pi_k\omega(0)) \tag{5.7}$$

holds. Consider the first equation of system (5.1) computed at t_k, namely

$$\dot{x}(t)|_{t_k} = Ax(t_k) + BL\omega(t_k). \tag{5.8}$$

Substituting (5.7) in equation (5.8) yields

$$\Pi S\omega(t_k) + Ae^{At_k}(x(0) - \Pi_k\omega(0)) = Ax(t_k) + BL\omega(t_k)$$

and multiplying on the left-hand side by C^jA^{-1} we obtain

$$C^jA^{-1}\Pi S\omega(t_k) + C^je^{At_k}(x(0) - \Pi_k\omega(0)) = C^jx(t_k) + C^jA^{-1}BL\omega(t_k).$$

The matrix $\widetilde{C^j\Pi}_k$ defined in equation (5.6) is such that

$$C^jx(t_k) = y^j(t_k) = \widetilde{C^j\Pi}_k\omega(t_k), \tag{5.9}$$

which yields

$$C^jA^{-1}\Pi S\omega(t_k) + C^je^{At_k}(x(0) - \Pi_k\omega(0)) = \widetilde{C^j\Pi}_k\omega(t_k) + C^jA^{-1}BL\omega(t_k)$$

and

$$\left(\widetilde{C^j\Pi}_k + C^jA^{-1}(BL - \Pi S)\right)\omega(t_k) = C^je^{At_k}(x(0) - \Pi\omega(0)).$$

Using equation (5.3), yields

$$\left(\widetilde{C^j\Pi}_k - C^j\Pi\right)\omega(t_k) = C^je^{At_k}(x(0) - \Pi\omega(0)).$$

The rest of the proof follows from Lemma 15 and Theorem 13 in Chapter 4. □

Then a MIMO version of Algorithm 1 can be formulated.

Algorithm 2. Let k be a sufficiently large integer. Select $\eta_j > 0$, with $j = 1, \ldots, p$ sufficiently small. Select the integer $g \geq 1$.

1: Construct the matrices $\widetilde{R}_k$ and $\widetilde{\Upsilon}_k^j$ for all $j = 1, \ldots, p$.

2: **If** rank $\left(\widetilde{R}_k\right) = \nu$ **then** compute $\widetilde{C^j\Pi}_k$ solving equation (5.6) for all $j = 1, \ldots, p$.
Else increase g.

If $k - g\nu < 0$ **then** restart the algorithm selecting a larger initial k.

3: **If** $\left|\left|\widetilde{C^j\Pi}_k - \widetilde{C^j\Pi}_{k-1}\right|\right| > \dfrac{\eta_j}{t_k - t_{k-1}}$ for some η_j **then** $k = k+1$ **go to** 1.

4: **Stop**.

Remark 36. If $g = 1$ the matrix $\widetilde{R}_k$ is full rank. However, in practice numerical approximations in the computation of $\widetilde{R}_k$ may cause loss of rank. For this reason we formulate the algorithm with the possibility of increasing the number of measurements. As a consequence Algorithm 2 gives a least square approximation of the moments. According to our experience, good values of g can usually be selected in the set $[1, 10]$ depending on the system. In the simulations presented in this chapter $g = 4$. ■

5.4 Application to Power Systems

5.4.1 Power system model

We describe a power system composed of n_m-machines and n_b-bus with the classical model, see [167,168], which is normally used in the literature of model reduction of power systems, see *e.g.* [157, 162, 163]. The model is described by the differential equations

$$\begin{aligned}\dot{\delta}_i &= \omega_i - \omega_s,\\ \frac{2H_i}{\omega_s}\dot{\omega}_i &= T_{M_i} - D_i(\omega_i - \omega_s) - E_i^2 G_{ii} - E_i \sum_{j=1,\, j\neq i}^{n_m} \left(E_j G_{ij}\cos(\delta_i - \delta_j) + E_j B_{ij}\sin(\delta_i - \delta_j)\right),\end{aligned} \tag{5.10}$$

with $i = 1, \ldots, n_m$, where δ_i and ω_i are the rotor angle and angular velocity, respectively, of the i-th machine, ω_s is the reference angular velocity, H_i and D_i are the inertia and damping coefficients, respectively, of the i-th machine, E_i is the internal voltage of the i-th machine, $Y_{ij} = G_{ij} + \iota B_{ij}$ is the admittance between the machines i and j, G_{ii} is the self-conductance of the i-th machine and T_{M_i} is the mechanical input power of the i-th machine.

In the literature on model reduction of power systems the study area and the external area are sometimes modeled as two separate entities interconnected each other with n_p-tie-lines, see *e.g.* [157, 162]. However, this is a somewhat strong approximation. In fact, note that if the two power systems, study area and external area, are interconnected then we have a unique large power system and the power flow analysis which defines the parameters

of system (5.10) has to be updated. Using the tie-lines only to exchange the input and output of the two systems gives the considerable simplification that the number of inputs and outputs corresponds to the arbitrary number of tie-lines.

On the contrary, in this chapter the division in study area and external area is a pure exercise of labeling. In fact, the whole power system is described by equations (5.10) and the division in study and external area can be done over every region of the power system using the actual buses as interconnecting lines between the two areas. This approach has the advantage of improving the fidelity of the simulation of the power system. The drawback is that the number of inputs of the external area is the number of machines of the study area and the number of outputs of the external area is the number of machines of the external area, and *vice versa*. However, since with the considered method the dimension of the reduced order model does not depend upon the number of inputs and outputs, a large number of inputs and outputs is not an issue for the technique we are presenting.

Thus, consider an external areas composed of e_m-machines and a study areas composed of s_m-machines, with $s_m + e_m = n_m$. The study area is described by system (5.10) for $i = e_m + 1, \dots, s_m$. The input of the study area (output of the external area) is δ_j, with $j = 1, \dots, e_m$. The external area is described by the linearization of system (5.10) around an equilibrium point, namely

$$\begin{bmatrix} \Delta\dot{\delta} \\ \Delta\dot{\omega} \end{bmatrix} = \begin{bmatrix} 0 & I \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \Delta\delta \\ \Delta\omega \end{bmatrix} + \begin{bmatrix} 0 \\ B_2 \end{bmatrix} \Delta u, \qquad y = \begin{bmatrix} C_1 & 0 \end{bmatrix} \begin{bmatrix} \Delta\delta \\ \Delta\omega \end{bmatrix}, \tag{5.11}$$

where $\Delta\delta_i = \delta_i - \delta_i^0$, $\Delta\omega_i = \omega_i - \omega_i^0$, with $i = 1, \dots, e_m$, $\Delta u_j = u_j - u_j^0 = \delta_j - \delta_j^0$, with $j = e_m + 1, \dots, n_m$, $(\delta^0, \omega^0, u^0)$ is an equilibrium point and the remaining matrices are defined as $A_{21} = \omega_s \operatorname{diag}(2H_i)^{-1} K$, $A_{22} = -\omega_s \operatorname{diag}(2H_i)^{-1} \operatorname{diag}(D_i)$ and $B_2 = \omega_s \operatorname{diag}(2H_i)^{-1} Z$, with the components of K and Z defined as

$$K_{ij} = E_i E_j B_{ij} \cos(\delta_i^0 - \delta_j^0) - E_i E_j B_{ij} \sin(\delta_i^0 - \delta_j^0), \tag{5.12}$$

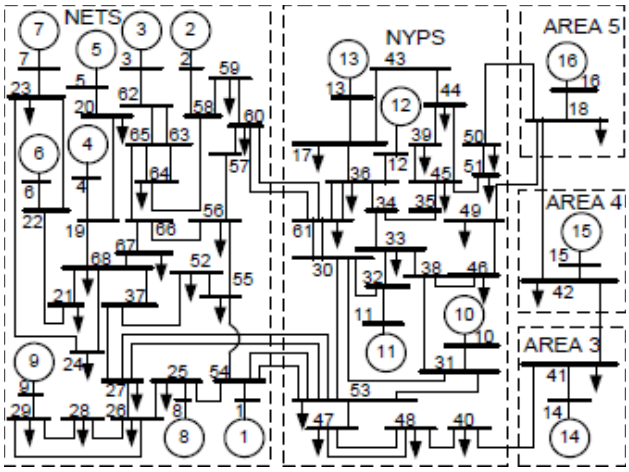

Figure 5.1: Line diagram of the 68-bus system, see [164].

for $i = 1, \ldots, e_m$, $j = 1, \ldots, e_m$ and $j \neq i$,

$$K_{ii} = \sum_{j=1,\, j\neq i}^{e_m+s_m} -E_i E_j B_{ij} \cos(\delta_i^0 - \delta_j^0) + E_i E_j B_{ij} \sin(\delta_i^0 - \delta_j^0), \tag{5.13}$$

for $i = 1, \ldots, e_m$, and

$$Z_{ik} = E_i E_k B_{ik} \cos(\delta_i^0 - \delta_k^0) - E_i E_k B_{ik} \sin(\delta_i^0 - \delta_k^0), \tag{5.14}$$

for $i = 1, \ldots, e_m$ and $k = e_m + 1, \ldots, s_m$.

5.4.2 NETS-NYPS benchmark system

The theory presented is validated on the interconnected New England test system (NETS) and New York test system (NYPS) 68-bus, 16-machine, 5-area power system, shown in Fig. 5.1, see [164]. The study area, composed of the machines 14, 15 and 16, is interconnected with the bus-lines $18-50$, $18-49$ and $41-40$ to the external area, composed of the machines from 1 to 13. The separation in study and external area follows the geographical distribution of the power system and the tie-lines are actual bus-lines of the power system. The system to be reduced has $n = 26$, $m = 3$, $p = 13$. The parameters of system (5.10) have been computed in MATLAB^{TM} as follows. With the script *Init_MultiMachine.m*, which can be downloaded from [164], the line data, the power flow results, the generator direct axis transient reactance and the inertia values H_i of the machines have been gener-

ated. Following the theory presented in [167], see also [168], we have computed the reduced bus admittances Y_{ij}, the equilibrium voltages E_i and the equilibrium angles δ_i. The mechanical input powers T_{M_i} have been computed from these quantities and equations (5.10) written at the equilibrium point. The damping coefficients D_i have been generated with the function *rand*. Finally, the quantities of the linearized system have been computed directly from (5.12)-(5.13)-(5.14).

5.4.3 Approximation of L

To determine the reduced order model we have to select the interpolation points (by constructing the signal generator) and the eigenvalues that we want to maintain. In this section we present the heuristic that has been used to construct the matrix L of the signal generator.

Algorithm 3. Let $r = 1$ and $j = 1$. Consider C^j the j-th row of C, B^r the r-th column of B and l_i^r the r-th element of l_i.

1. Consider the system $(A, \tilde{B} = B^1, C^1)$ and design a reduced order model with $l_i^1 = 1$, for all $i = 1, \ldots, \nu$, *i.e.* select the interpolation points and the desired eigenvalues of the reduced order model to achieve the desired approximation of the resulting SISO system.

2. **If** $r < m$ **then** $r = r + 1$. Using the same interpolation points and desired eigenvalues compute a reduced order model of the multi-input system $(A, \tilde{B} = [\tilde{B} \ B^r], C^1)$ as $\min_{l_i^r} ||W - \hat{W}||_{\mathcal{H}_2}$, where $||\cdot||_{\mathcal{H}_2}$ is the $\mathcal{H}_2$ norm. **Repeat.**
   Else with the obtained matrix L, the same interpolation points and desired eigenvalues compute a reduced order model of (A, B, C).

3. **Stop**.

This heuristic is justified by the observation that L is involved in combining m elements. Thus, we may expect that we can ignore the fact that the system is multi-output in the approximation of L. In addition, one can expect that the set of interpolation points and prescribed eigenvalues which have been selected to obtain a desired approximation on the truncated SISO system can be used on the MIMO system. This is justified by the observation that we still want to maintain an approximation on the truncated system as

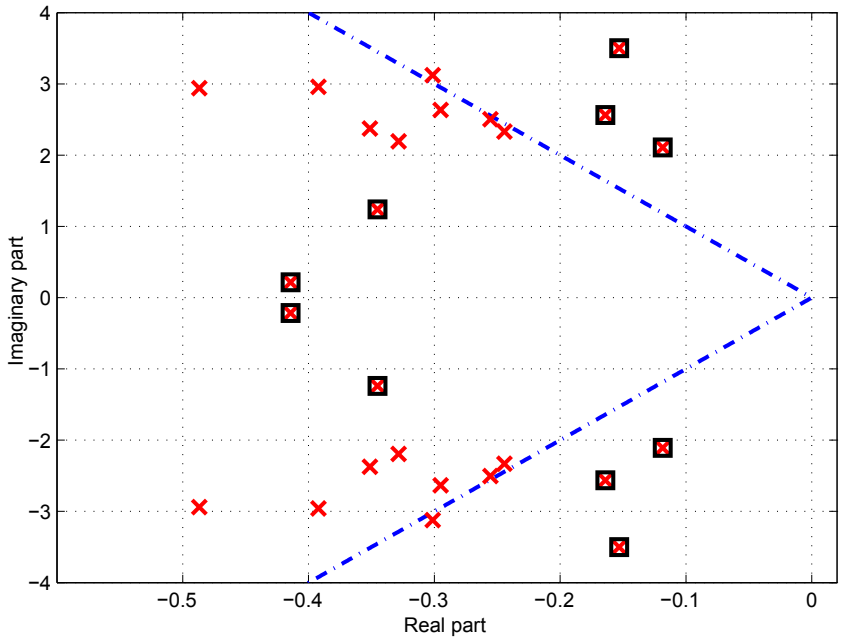

Figure 5.2: Eigenvalues of the linear system (5.11) (crosses) and of the reduced order model (5.4) (squares). The dash-dotted lines represent the 10% damping ratio.

close as possible to the one we have designed. As a consequence of these observations, the approximation of the $m\nu$ elements of L has been reduced to the problem of determining $m-1$ independent scalars l_i^r.

Remark 37. There is no guarantee that this heuristic works in general. However, our simulations have shown that the resulting L is at least locally optimal with respect the $\mathcal{H}_2$ norm. A small variation of any of the obtained elements of L causes a rapid increase of the $\mathcal{H}_2$ norm. ■

5.4.4 Design of the reduced order model

The selection of the interpolation points and of the eigenvalues of the reduced order model used in this section corresponds to the "Case 4" ($\nu = 10$) analyzed in Section 5.4.6 and we therefore refer the reader to that section. Therein with a fault behavior analysis it is shown that good time-domain performance can be obtained with a reduced order model of order between 6 and 12 if the correct slow and poorly damped modes are preserved. System (5.11) has been simulated with the input generated by the signal generator (5.5). The matrix of the generator has been selected as $S = \text{diag}(S_0, 2.12\,\bar{S}, 2.79\,\bar{S}, 3.42\,\bar{S}, 5\,\bar{S})$,

with

$$S_0 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \bar{S} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

which corresponds to choosing the interpolation points at 0 (zero and first order moment), 0.3374 Hz, 0.444 Hz, 0.5443 Hz and 0.7958 Hz (all zero order moments). The reduced order model is computed with L approximated using Algorithm 3 and $C\Pi$ approximated with Algorithm 2. We determine the six least damped and the four slowest eigenvalues of system (5.11) and we assign them to the reduced order model (5.4). In Fig. 5.2 the eigenvalues of system (5.11) are represented with crosses, whereas the eigenvalues of the reduced order model are depicted with squares. In the figure the modes in the area between the two dash-dotted lines are well-damped (more than 10% damping ratio), whereas the others are considered poorly damped [164].

5.4.5 Approximation of the moments

In this section we illustrate the performance of Algorithm 2 by showing how the approximated reduced order model improves as k in Algorithm 2 increases. The surfaces in Fig. 5.3 represent the magnitude (left graphs) and phase (right graphs) of the elements $W_{1,1}$ (top), $W_{10,2}$ (middle), $W_{11,3}$ (bottom) of the transfer matrix of the reduced order model as a function of t_k, with $3.4448 \leq t_k \leq 48.4299\,s$. For comparison, the solid/black lines indicate the magnitude and phase of the respective elements of the transfer matrix of the reduced order model for the exact moments $C\Pi$, *i.e.* computed solving equation (5.3). The figures show how the approximated magnitude and phase of the reduced order model (5.4) evolve over time and approach the respective quantities of the exact reduced order model as $t_k \to \infty$. Finally note that we have chosen to show these three particular components of the transfer matrix because the rest of the components are very similar to these. These are the components that present the most distinctive graphical features. Fig. 5.4 shows the magnitude of the elements $W_{1,1}$ (top), $W_{10,2}$ (middle) and $W_{11,3}$ (bottom) of the transfer matrix of system (5.11) (solid lines) and of the reduced order model (dotted lines). We note that the curves of the reduced order model are close to the curves of the system along all the frequencies. As expected, the approximation is not uniformly good

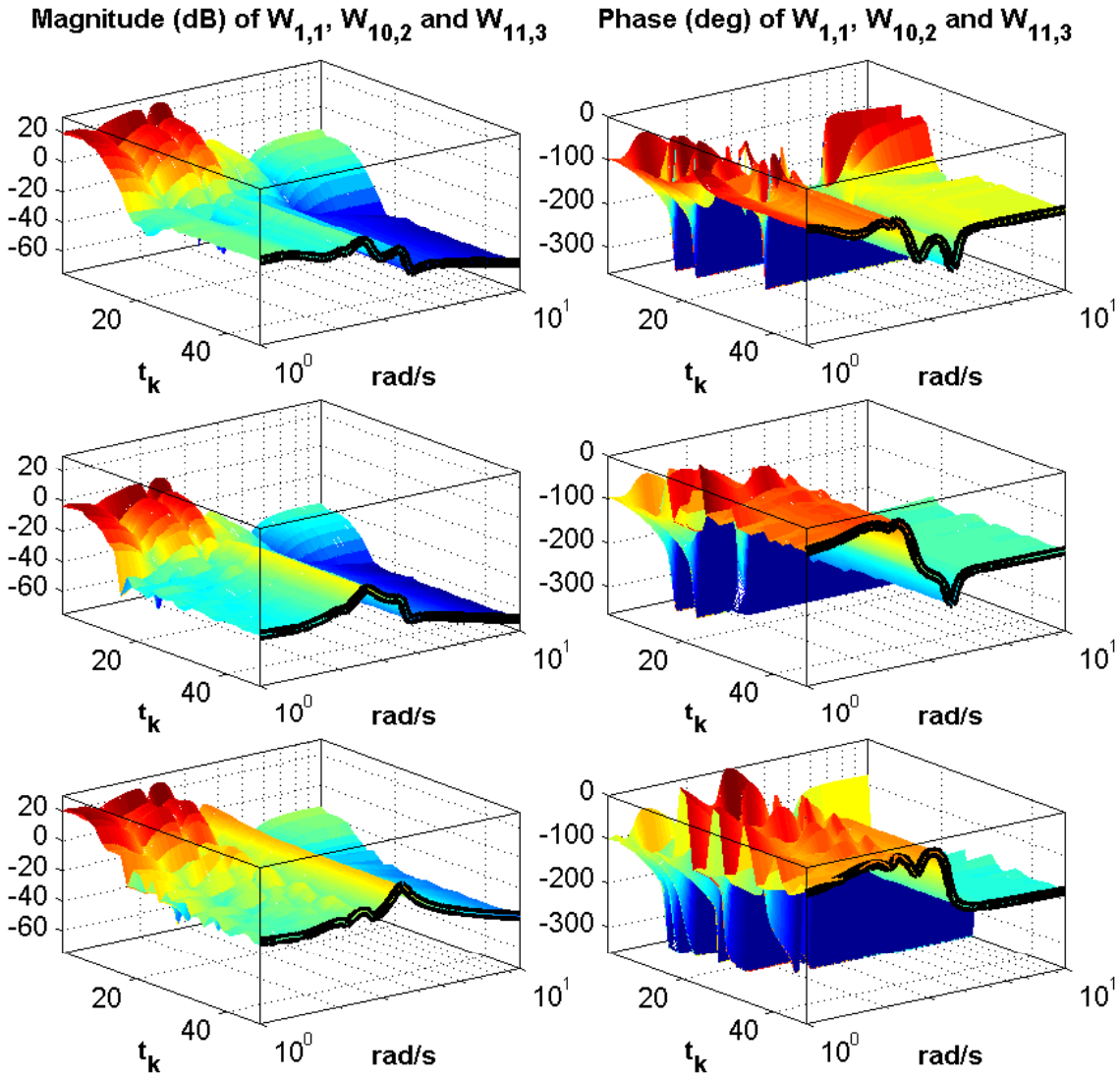

Figure 5.3: The mesh represents the magnitude (left graphs) and phase (right graphs) of the elements $W_{1,1}$ (top), $W_{10,2}$ (middle), $W_{11,3}$ (bottom) of the transfer matrix of the reduced order model as a function of t_k, with $3.4448 \leq t_k \leq 48.4299\,s$. The solid/black lines indicate the magnitude and phase of the respective elements of the transfer matrix of the reduced order model (5.4) for the exact moments $C\Pi$, *i.e.* computed solving equation (5.3).

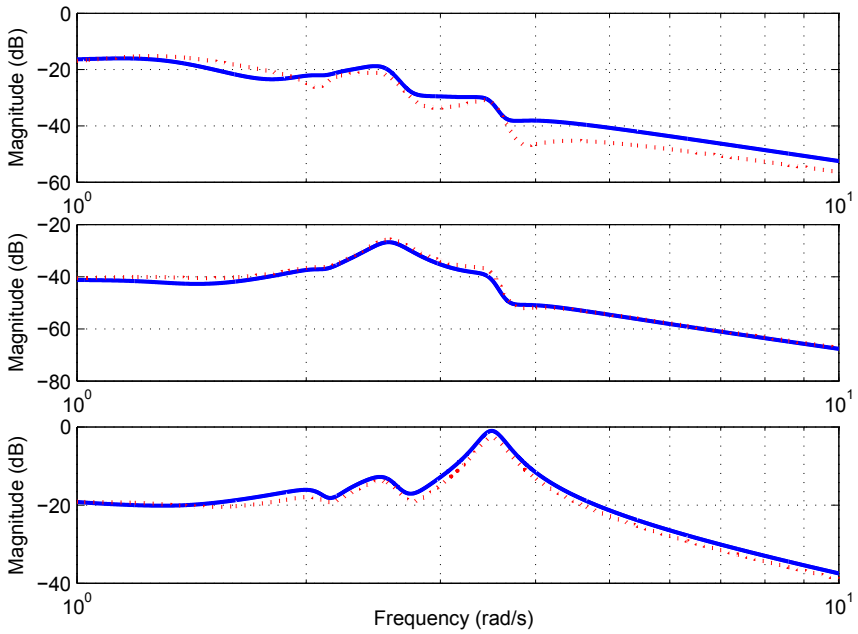

Figure 5.4: Magnitude of the elements $W_{1,1}$ (top), $W_{10,2}$ (middle) and $W_{11,3}$ (bottom) of the transfer matrix of system (5.11) (solid lines) and of the reduced order model (5.4) (dotted lines).

for all the elements of W. In fact, the curve in the top graph does not approximate the given system as well as the one of the other two graphs. This is caused by the use of the tangential interpolation, *i.e.* we are trying to capture more information maintaining the same number of parameters (the order of the reduced order model).

5.4.6 Selection of the modes to be preserved

In this section we study the problem of selecting the dimension of the reduced order model and the modes to be preserved through a fault behavior analysis of the interconnected system. We first challenge the common belief that to improve the approximation it is necessary to increase the number of interpolation points, *i.e.* the dimension of the reduced order model. Then we show that to improve the approximation it is sufficient to preserve specific modes, namely poorly damped and slow modes. Note that for each case that we consider we show only one randomly chosen Bode plot (for the sake of this chapter, the 2nd diagonal term of the transfer function) out of the total thirty-nine. This section justifies the selection of the order $\nu = 10$ and of the interpolation points and eigenvalues to be preserved that have been used in the previous sections.

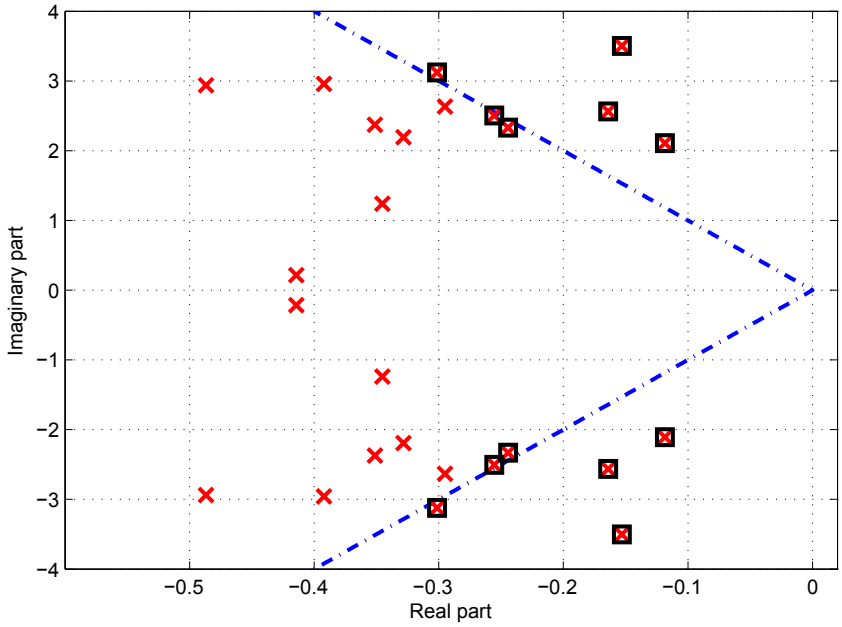

Figure 5.5: Case 1. Eigenvalues of the linear system (5.11) (crosses) and of the reduced order model (5.4) (squares). The dash-dotted lines represent the 10% damping ratio.

Case 1: we start with computing a reduced order model of dimension $\nu = 12$. The interpolation points have been chosen to be at 0 (zero and first moment), at 0.234 Hz, 0.3374 Hz, 0.444 Hz, 0.5443 Hz, 0.7958 Hz (all zero moments). Note that the solution of equation (5.3) can be computed with the function *Sylvester* of MATLAB. However, the result is very imprecise for the MIMO system considered. Then we have implemented a custom function based on the Kronecker product, see *e.g.* [5]. We determine the twelve least damped eigenvalues of system (5.11) and we assign them to the reduced order model (5.4) (see Fig. 5.5). Fig. 5.6 shows the Bode plot of the 2nd diagonal term of the transfer matrix of system (5.11) (solid lines) and of the reduced order model (dotted lines). We can see that the two graphs starts to diverge already around 0.0318 Hz and that the behavior from medium to high frequencies is very different. A dynamic simulation of the power system is performed. A self-clearing fault at bus 14 of the study area occurring at $t = 1\,s$ and cleared at $1.15\,s$ is simulated. Fig. 5.7 shows the angular velocities (large) and respective absolute errors (insert) of the study area when this is connected to the nonlinear system describing the external area (solid lines) and when it is connected to the reduced order model (dotted lines). The approximation given by the reduced order model is totally

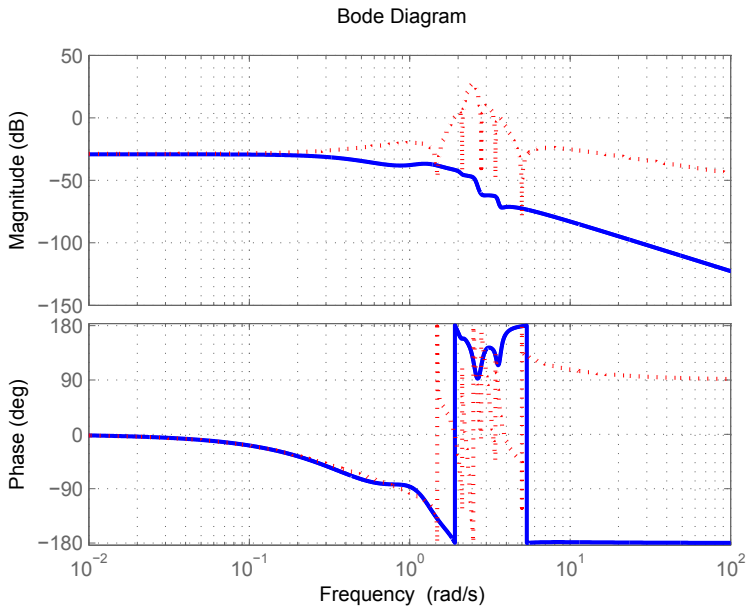

Figure 5.6: Case 1. Bode plot (2nd diagonal term of the transfer function) of the linear system (5.11) (solid lines) and of the reduced order model (5.4) (dotted lines) for the set of eigenvalues shown in Fig. 5.5.

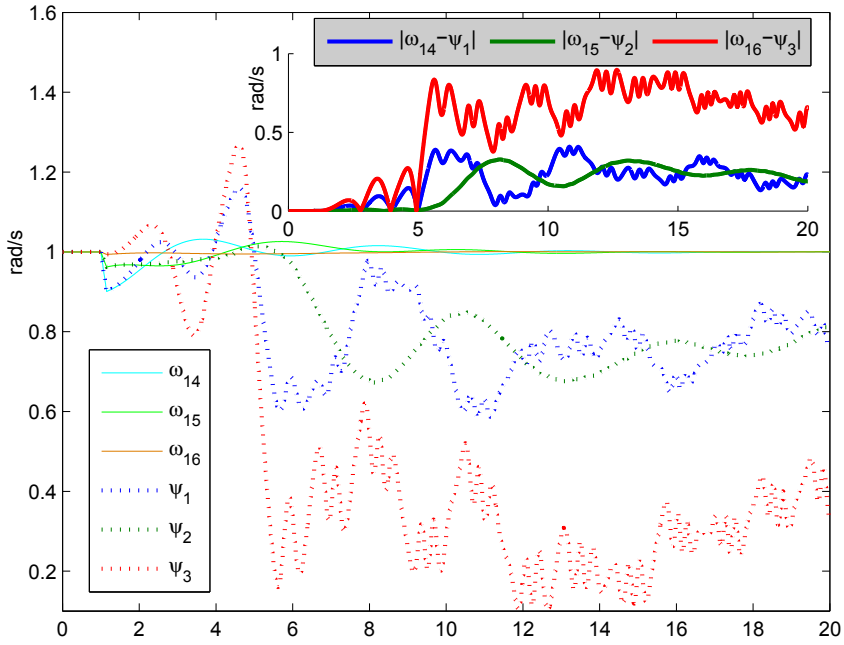

Figure 5.7: Case 1. Large: angular velocities of the study area when this is connected to the nonlinear system describing the external area (solid lines) and when it is connected to the reduced order model (5.4) (dotted lines) with the eigenvalues shown in Fig. 5.5. Insert: absolute errors between the time histories.

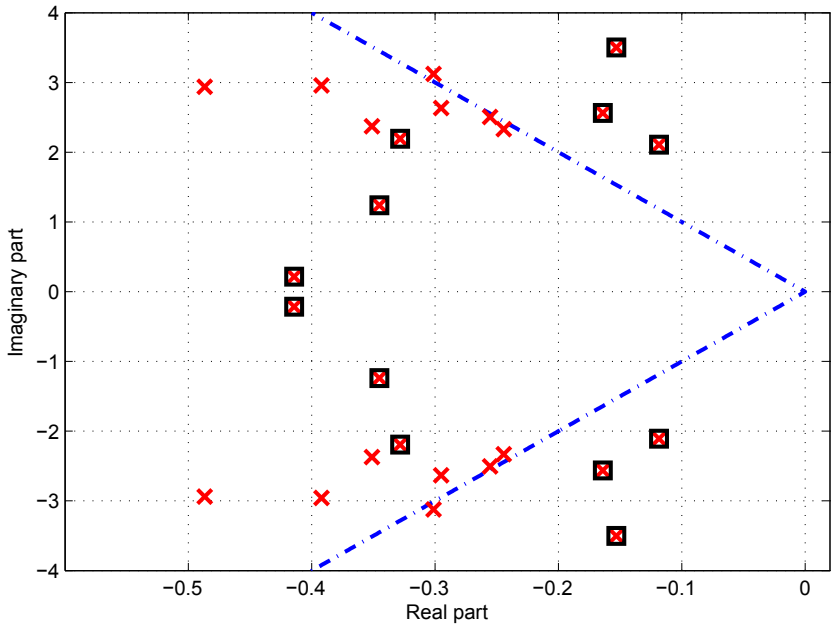

Figure 5.8: Case 2. Eigenvalues of the linear system (5.11) (crosses) and of the reduced order model (5.4) (squares). The dash-dotted line represent the 10% damping ratio.

unsatisfactory. Although Fig. 5.6 shows a good approximation of the steady state (low frequency), the interconnection of the reduced order model and the study area is unstable. In fact, the fault generates high frequency oscillations which make the trajectories of the system exit the region of attraction of the equilibrium point.

Case 2: the common approach used to improve the quality of the approximation is to increase the order of the reduced model with the hope that the new reduced order model be able to better capture the dynamics of the system to be reduced. We show that this is not necessary. We keep the same order $\nu = 12$ maintaining the same matrices S and L. This time we assign as eigenvalues the six most poorly damped modes and the six slowest modes (see Fig. 5.8). Fig. 5.9 shows the Bode plot of the 2nd diagonal term of the transfer matrix of system (5.11) (solid lines) and of the reduced order model (dotted lines). We note that the two graphs are close. The dynamic simulation graph for this case is omitted since it is similar to the next, more interesting, case. Thus we see that the eigenvalues retained in the reduced order model play a role that can be more important of the order of the reduction.

Case 3: to strengthen this last observation we now decrease the order of the reduced

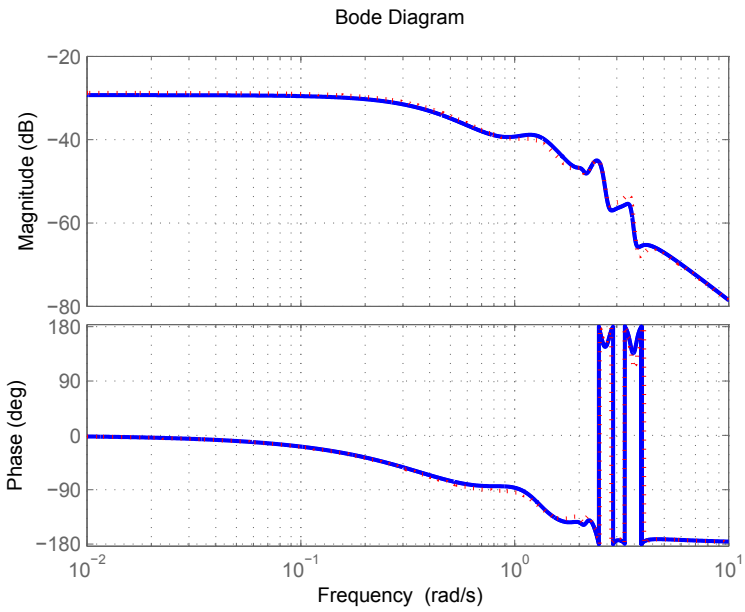

Figure 5.9: Case 2. Bode plot (2nd diagonal term of the transfer function) of the linear system (5.11) (solid lines) and of the reduced order model (5.4) (dotted lines) for the set of eigenvalues shown in Fig. 5.8.

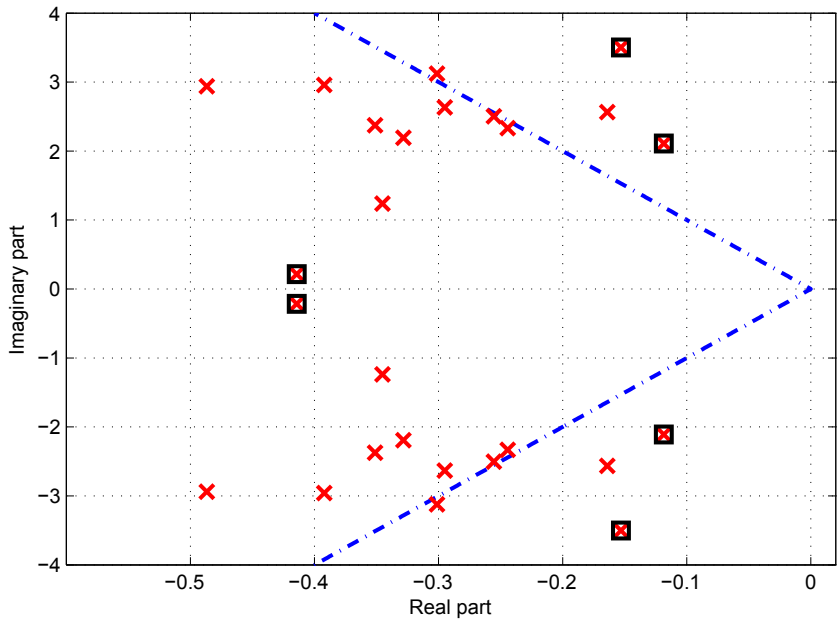

Figure 5.10: Case 3. Eigenvalues of the linear system (5.11) (crosses) and of the reduced order model (5.4) (squares). The dash-dotted lines represent the 10% damping ratio.

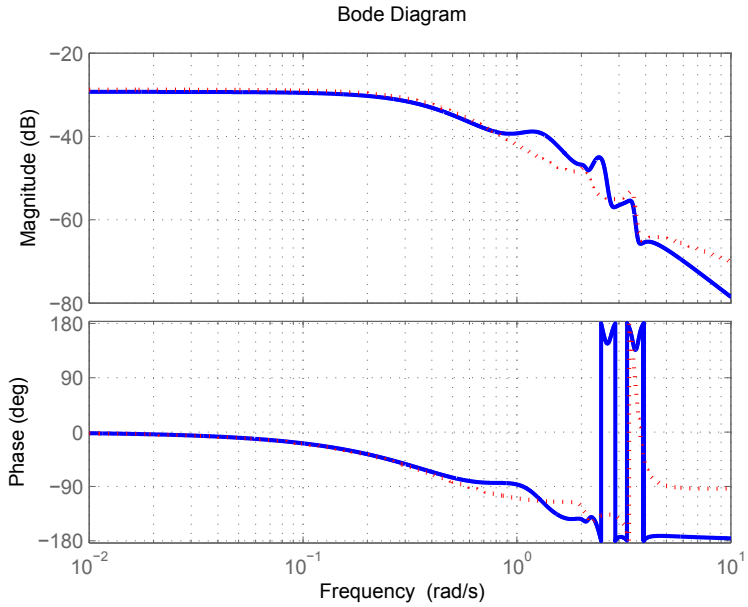

Figure 5.11: Case 3. Bode plot (2nd diagonal term of the transfer function) of the linear system (5.11) (solid lines) and of the reduced order model (5.4) (dotted lines) for the set of eigenvalues shown in Fig. 5.10.

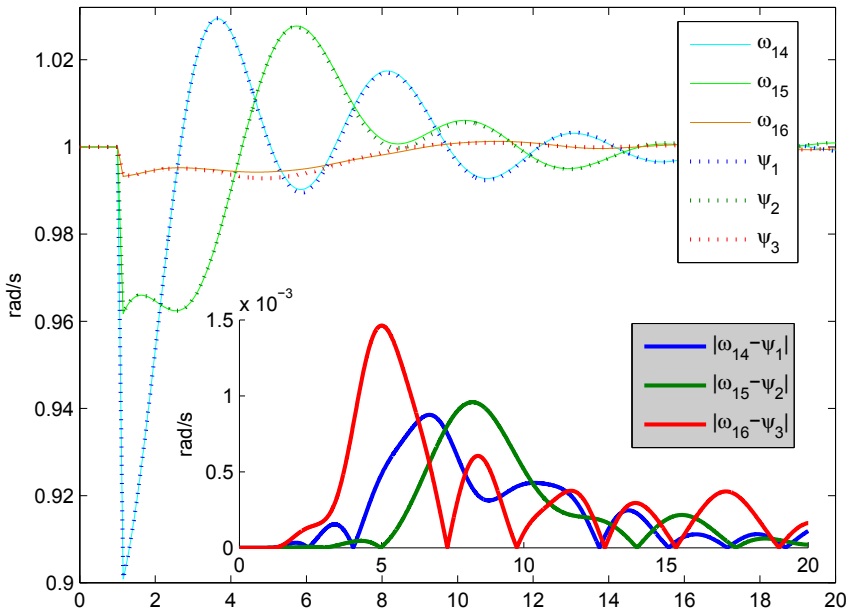

Figure 5.12: Case 3. Large: angular velocities of the study area when this is connected to the nonlinear system describing the external area (solid lines) and when it is connected to the reduced order model (5.4) (dotted lines) with the eigenvalues shown in Fig. 5.10. Insert: absolute errors between the time histories.

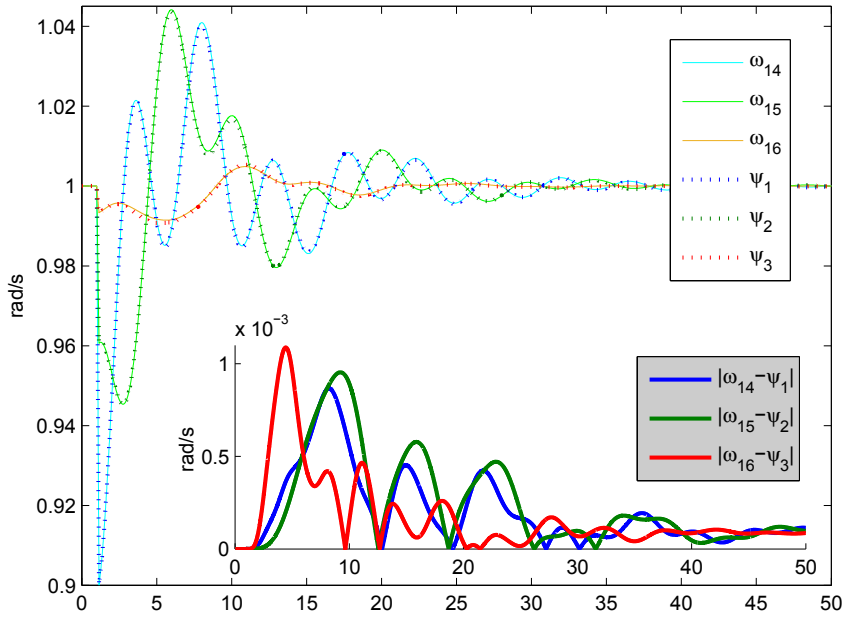

Figure 5.13: Case 4. Large: angular velocities of the study area when this is connected to the nonlinear system describing the external area (solid lines) and when it is connected to the reduced order model (5.4) (dotted lines) with the eigenvalues shown in Fig. 5.2. Insert: absolute errors between the time histories.

model to $\nu = 6$. The interpolation points have been chosen as a subset of the previous case, namely 0 (zero and first moment), 0.3374 Hz and 0.5443 Hz (all zero moments). This time we maintain the four most poorly damped modes and the two slowest modes (see Fig. 5.10). Fig. 5.11 shows the Bode plot of the 2nd diagonal term of the transfer matrix of system (5.11) (solid lines) and of the reduced order model (dotted lines). We see that in this case the reduced order model is a better approximation with respect to case 1 ($\nu = 12$ with a bad selection of eigenvalues) but obviously a worse approximation with respect to case 2 ($\nu = 12$ with a good selection of eigenvalues). However, for the dynamic behavior of interest this approximation is sufficiently good. In fact, Fig. 5.12 shows the angular velocities (large) and respective absolute errors (insert) of the study area when this is connected to the nonlinear system describing the external area (solid lines) and when it is connected to the reduced order model (dotted lines). We see that the two time histories are almost indistinguishable.

Case 4: we finally study the case $\nu = 10$ that we have used in the previous sections. The order of 10 is a compromise between the minimal and maximal dimensions that we

have analyzed. Fig. 5.13 shows the angular velocities (large) and respective absolute errors (insert) of the study area when this is connected to the nonlinear system describing the external area (solid lines) and when it is connected to the reduced order model (dotted lines). We see that also in this case the two time histories are almost indistinguishable. Thus, the analysis and control of the nonlinear system describing the study area can be performed using the interconnection of such area with the reduced order model instead of the full nonlinear description of the external area: in fact, the dynamic response to a fault is almost identical but the number of the equations is reduced.

5.5 Conclusion

In this chapter we have provided a validation of the theory developed in Chapters 2 and 4. We have presented a low complexity algorithm for the fast estimation of the moments of MIMO systems. The estimated moments have been exploited for the model reduction of large-scale interconnected power systems. The technique that we have demonstrated offers, simultaneously, a low computational complexity approximation of the moments and the possibility to easily enforce constraints on the reduced order model. This possibility has been used to preserve selected slow and poorly damped modes which are important both from a mathematical and physical point of view. The problem of the choice of the so-called tangential directions has also been studied and an heuristic for their approximation has been given. The techniques have been validated with the study of the dynamic response of the NETS-NYPS benchmark system.

Chapter 6

Model reduction by matching the steady-state response of explicit signal generators

6.1 Introduction

In this chapter we generalize the model reduction by moment matching with regards to the class of input signals considered. In Chapter 2 we have shown that the moments of a linear system are in one-to-one relation with the eigenvalues of the dynamic matrix S of the exogenous system $\dot{\omega} = S\omega$. Now, we consider input signals generated by a linear exogenous system represented in explicit form[1], *i.e.* an implicit (differential) form may not exist. This direction of investigation is motivated by a large number of applications in which standard operating conditions are associated to non-continuous or non-differential input signals. For instance, power converters are controlled by means of pulse width modulated (PWM) signals or sawtooth signals[2], see *e.g.* [171], [172], [173], and [174]. To achieve this extension, a new integral equation to characterize the moments is given. Under specific assumptions, the equivalence of this new definition and the one based on the Sylvester equation given in Chapter 2 is proved. Finally a new family of reduced order models is presented and connections with the family of models given in Chapter 2 are

[1] The terminology is taken from [169, 170]. See also Definition 18.

[2] Note that the use of the results of this chapter for the analysis of power converters is the topic of Chapter 8.

drawn.

The rest of the chapter is organized as follows. In Section 6.2 we give a formal description of the problem addressed. In Section 6.3 the definition of moment is reformulated for explicit signal generators. In Section 6.4 a family of reduced order models for linear systems is introduced. Finally Section 6.5 contains some concluding remarks.

All the results of this chapter are original contributions developed in fulfillment of my PhD course and they have been published in the conference paper [16] and in the journal paper [17].

6.2 Problem formulation

Since the moments of the system described by the equations

$$\dot{x} = Ax + Bu, \qquad y = Cx, \tag{6.1}$$

are related to the solution of the Sylvester equation

$$A\Pi + BL = \Pi S, \tag{6.2}$$

is based on the availability of a differential representation of the signal generator, namely equation

$$\dot{\omega} = S\omega, \qquad u = L\omega. \tag{6.3}$$

However, there are notable applications in which this may not be the case. For instance, as already noted in the introduction, the input of a dynamical system describing a power electronic device can often be a PWM wave which cannot be represented as the output of a system described by smooth differential equations. For this reason we introduce the following definition.

Definition 18. Let x, with $x(t) \in \mathbb{R}^n$, be the state of a dynamical system Σ. Let u, with $u(t) \in \mathbb{R}^m$, be the input of Σ. Let t_0 and $x_0 = x(t_0)$ be the initial time and the initial state, respectively. If there exists a function $\phi : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ such that

$$x(t) = \phi(t, t_0, x_0, u), \tag{6.4}$$

for all $t \geq t_0$, we call equation (6.4) the *representation in explicit form* [169], or the *explicit model*, of Σ.

Assume $\phi(t, t_0, x_0, u)$ has a continuous derivative with respect to t for every t_0, x_0 and u, and there exists a function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ continuous for each t over $\mathbb{R}^n \times \mathbb{R}^m$ such that

$$\dot{x} = f(x, u). \tag{6.5}$$

We call the differential equation (6.5) the *representation in implicit form* [170], or the *implicit model*, of Σ.

The goal of the chapter is to extend the theory developed in Chapter 2 for linear signal generators which have an implicit model to signal generators in explicit form.

To gain insight on the use of explicit models for signal generators, we look now at the linear generator in equation (6.3) making two trivial observations. The first one is that the interpolation points s_i, eigenvalues of S, are the poles of the Laplace transform of the free output response of system (6.3). The second one is that an input signal which is the sum of sinusoidal signals can be generated by equation (6.3) defining the matrix S such that the eigenvalues of S are in relation with the angular frequencies of the sinusoidal signals. We now try to reinterpret these two observations for a signal generator which does not have an implicit model. To begin with consider a square wave $\sqcap(t)$ defined as

$$\sqcap(t) = \operatorname{sign}(\sin(t)) = \begin{cases} 1, & (k-1)\pi < t < k\pi, \\ 0, & t = k\pi \text{ or } t = (k+1)\pi, \\ -1, & k\pi < t < (k+1)\pi, \end{cases}$$

with $\operatorname{sign}(0) = 0$, and $k = 1, 3, 5, \ldots, +\infty$. The Laplace transform of this function is

$$\mathcal{L}(\sqcap(t)) = \frac{1 - e^{-s\pi}}{s(1 + e^{-s\pi})},$$

and this has the poles

$$\bar{s} = 0, \qquad \bar{s}_j = (2j+1)\iota,$$

with $j = -\infty, \ldots, -1, 0, 1, \ldots, +\infty$. We see, then, that if we "insist on" the relation

between the implicit model and the Laplace transform and if we want to compute a reduced order model with the technique presented in Chapter 2, we have to interpolate at infinitely many points s_i. This is suggested also by the second observation. Since the function $\sqcap(t)$ is periodic, it admits a Fourier series, namely

$$\sqcap(t) = \frac{4}{\pi} \sum_{k=1,3,5,\ldots,+\infty}^{\infty} \frac{1}{k} \sin(kt).$$

Then, consistently with the second observation we again have to interpolate at infinitely many points. Both these observations suggest that we could describe $\sqcap(t)$ by means of the infinite dimensional system [175]

$$\dot{\omega} = \begin{bmatrix} \ddots & \ddots & & & & & & \\ \ddots & +2\iota & 0 & & & & & \\ & 0 & +\iota & 0 & & & & \\ & & 0 & 0 & 0 & & & \\ & & & 0 & -\iota & 0 & & \\ & & & & 0 & -2\iota & \ddots & \\ & & & & & \ddots & \ddots & \end{bmatrix} \omega,$$

with output $\sqcap = \hat{P}\omega$ for some "matrix" $\hat{P}$. This shows a limitation of the current approach. In fact, to interpolate at infinitely many points we need an "infinite dimensional mapping" Π and an infinite dimensional reduced order model (see [114]).

To overcome these issues we consider signal generators in explicit form. Thus, consider

$$\omega(t) = \Lambda(t, t_0)\omega_0, \qquad u = L\omega, \tag{6.6}$$

with $\Lambda(t, t_0) \in \mathbb{R}^{\nu \times \nu}$ such that $\Lambda(t_0, t_0) = I$. Note that this model provides a very general class of models which contains the implicit model (6.3), but that describes several other signal generators. For instance, any periodic signal can be described by (6.6) with the property

$$\Lambda(t, t_0) = \Lambda(t - T, t_0), \quad t \geq T + t_0, \tag{6.7}$$

where T is the period of the signal u. Alternatively (6.6) can represent a signal generator described by a time-varying system of the form

$$\dot{\omega} = S(t)\omega, \qquad u = L\omega, \tag{6.8}$$

in which case $\Lambda(t, t_0)$ is (with additional assumptions, *e.g.* the semigroup property) the transition matrix associated to (6.8) [176, Section 3].

Since the definition of moment given in this chapter is based on the existence of the steady-state response of system (6.1) driven by (6.6), we need to introduce further hypotheses on the class of input signals (6.6).

Assumption 14. The vector $\omega(t)$ defined in equation (6.6) has a strictly proper Laplace transform with non-negative poles.

Assumption 14 is a standard condition for the existence of a well-defined steady-state response of the state of system (6.1) driven by (6.6) [177–179].

Assumption 15. The matrix valued function $\Lambda(t, t_0)$ is non-singular for all $t \geq t_0$.

Assumption 15 is essential to have uniqueness of the solution $\omega(t)$ of (6.6). Note, in fact, that it is always satisfied by a generator of the form (6.3) and it is required for the uniqueness of the solution of system (6.8) (see *e.g.* [176]).

Assume now that there exists a set $\mathcal{T} \subset \mathbb{R}_{\geq 0}$ in which $\Lambda(t, t_0)$ is differentiable with respect to t and consider the time-varying system described by the equation

$$\dot{z}(t) = G(t)^\top z(t), \tag{6.9}$$

with $G(t) = -\dot{\Lambda}(t, t_0)\Lambda(t, t_0)^{-1}$. Let $\Phi(t, t_0)$ be the transition matrix of system (6.9).

Assumption 16. The function $G(t)$ is piecewise continuous with respect to t. Moreover, there exist $T \geq t_0$ and a polynomial $p(t)$ such that $||\Phi(t, t_0)^\top|| \leq p(t)$ for all $t \geq T$.

This last technical assumption guarantees that the norm of $z(t)$ in system (6.9) does not diverge to infinity exponentially [176, Section 29] and it is needed, as shown in the next section, to guarantee that the steady-state response $x_s(t)$ of system (6.1) driven by (6.6) can be written as $x_s(t) = \Pi(t)\omega(t)$ for some matrix valued function $\Pi(t)$. Moreover, the piecewise continuity of $G(t)$ guarantees that the steady-state response is unique. Note

that Assumption 16 is a generalization of the assumption (used through the Thesis) that S in (6.3) is such that $\sigma(S) \subset \mathbb{C}_{\geq 0}$. In this case $G(t) = -S$ and if $\sigma(S) \subset \mathbb{C}_{\geq 0}$, the condition $||e^{-S(t-t_0)}|| \leq p(t)$ holds trivially.

Assumption 17. The triple $(L, \Lambda, \omega(t_0))$ is minimal, *i.e.* L and $\omega(t_0)$ are such that $\sigma(\mathcal{L}(L\Lambda(t, t_0)\omega(t_0))) = \sigma(\mathcal{L}(\Lambda(t, t_0)))$.

This condition is derived from the observation that the minimality of $(L, S, \omega(0))$ is equivalent to the condition $\sigma(\mathcal{L}(Le^{St}\omega(0))) = \sigma(S)$. This condition guarantees, together with the minimality of (A, B, C), that all the modes of ω are excited and observable in the output, and captures the requirement that one wants that the dynamics of the signal generator is fully present in the steady-state response of the system.

Remark 38. Assumptions 14, 15, 16 and 17 are "mild" assumptions. For instance, they are satisfied by the general class of discontinuous periodic signals which are considered in the remaining of the chapter. ■

As already anticipated, among all the possible signals generated by an explicit model, we are particularly interested in periodic signals, which are generated by system (6.6) with the property (6.7) and its generalizations (see the next section).

6.3 Integral definition of moment

In this section we give a definition of moment in the case in which the signal generator does not have an implicit model. We begin by showing that the interconnection of system (6.1) with the signal generator (6.6) possesses a steady-state response $x_s(t)$ described by the relation $x_s(t) = \Pi(t)\omega(t)$ for some matrix valued function $\Pi(t)$. The following result holds.

Theorem 19. Consider system (6.1) and the signal generator (6.6). Assume Assumptions 15 and 16 hold, $\sigma(A) \subset \mathbb{C}_{<0}$ and $\Lambda(t)$ is almost everywhere differentiable. Let

$$\Pi(t) = \left(e^{A(t-t_0)}\Pi(t_0) + \int_{t_0}^{t} e^{A(t-\tau)} BL\Lambda(\tau, t_0) d\tau \right) \Lambda(t, t_0)^{-1}, \tag{6.10}$$

be a family of matrix valued functions parametrized in $\Pi(t_0) \in \mathbb{R}^{n \times \nu}$. Then there exists a unique $\Pi_\infty(t_0)$ such that, for any $\Pi(t_0)$, $\lim_{t \to +\infty} \Pi(t) - \Pi_\infty(t) = 0$, where $\Pi_\infty(t)$ is the solution of (6.10) with $\Pi(t_0) = \Pi_\infty(t_0)$. Moreover, if $x(t_0) = \Pi_\infty(t_0)\omega(t_0)$ then $x(t) -$

$\Pi_\infty(t)\omega(t) = 0$ for all $t \geq t_0$, and the set $\mathcal{M}_\infty = \{(x,\omega) \in \mathbb{R}^{n+\nu} \,|\, x(t) = \Pi_\infty(t)\omega(t)\}$ is attractive.

Proof. Let $\mathcal{T} \subset \mathbb{R}_{\geq 0}$ be a set in which $\Lambda(t,t_0)$ is differentiable with respect to t. Differentiating both sides of equation (6.10) over $\mathcal{T}$ yields

$$\begin{aligned}\dot{\Pi}(t)\Lambda(t,t_0) + \Pi(t)\dot{\Lambda}(t,t_0) - Ae^{A(t-t_0)}\Pi(t_0) &= BL\Lambda(t,t_0) + A\int_{t_0}^{t} e^{A(t-\tau)}BL\Lambda(\tau,t_0)d\tau \\ &= BL\Lambda(t,t_0) + A\Pi(t)\Lambda(t,t_0) - Ae^{A(t-t_0)}\Pi(t_0).\end{aligned}$$

Then, since Assumption 15 holds, we have

$$\dot{\Pi}(t) = A\Pi(t) + BL - \Pi(t)\dot{\Lambda}(t,t_0)\Lambda(t,t_0)^{-1}. \tag{6.11}$$

Let $\Pi_1(t)$ and $\Pi_2(t)$ be the solutions of equation (6.11) with initial conditions $\Pi_1(t_0)$ and $\Pi_2(t_0)$, respectively, and define the error $\hat{E}(t) = \Pi_1(t) - \Pi_2(t)$. Then

$$\begin{aligned}\dot{\hat{E}}(t) &= A\Pi_1(t) + BL - \Pi_1(t)\dot{\Lambda}(t,t_0)\Lambda(t,t_0)^{-1} - \left(A\Pi_2(t) + BL - \Pi_2(t)\dot{\Lambda}(t,t_0)\Lambda(t,t_0)^{-1}\right) \\ &= A\hat{E}(t) - \hat{E}(t)\dot{\Lambda}(t,t_0)\Lambda(t,t_0)^{-1}\end{aligned}$$

and (see [176, Section 11]) $\hat{E}(t) = e^{A(t-t_0)}\hat{E}(t_0)\Phi(t,t_0)^\top$.

Since $\Phi(t,t_0)^\top$ is bounded by a polynomial, by Assumption 16, $\hat{E}(t)$ converges to zero. This implies that there exist motions $\Pi_\infty(t)$ to which the solutions of equation (6.11) converge, *i.e.* for any $\Pi(t_0)$ there exists a $\Pi_\infty(t_0)$ such that $\lim_{t\to+\infty} \Pi(t) - \Pi_\infty(t) = 0$. Moreover, $\Pi_\infty(t)$ is unique for any $t \geq t_0$ by the piecewise continuity of $\dot{\Lambda}(t,t_0)\Lambda(t,t_0)^{-1}$, see *e.g.* [1, Theorem 3.2].

By Assumption 15, the unique solution of system (6.1) with input u defined by equation (6.6) is

$$x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}BL\Lambda(\tau,t_0)\omega(t_0)d\tau.$$

Let $x(t_0) = \Pi_\infty(t_0)\omega(t_0)$, straightforward computations show that

$$\begin{aligned}x(t) - \Pi_\infty(t)\omega(t) &= e^{A(t-t_0)}\Pi_\infty(t_0)\omega(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}BL\Lambda(\tau,t_0)\omega(t_0)d\tau \\ &\quad - \left(e^{A(t-t_0)}\Pi_\infty(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}BL\Lambda(\tau,t_0)d\tau\right)\Lambda(t,t_0)^{-1}\Lambda(t,t_0)\omega(t_0) = 0,\end{aligned}$$

for all $t \geq t_0$. The attractivity of $\Pi_\infty(t)$ and the invariance of $x(t) = \Pi_\infty(t)\omega(t)$ imply

that the set $\mathcal{M}_\infty$ is attractive. □

Corollary 2. Under the assumptions of Theorem 19 and Assumption 14, the function $\Pi_\infty(t)\omega(t)$ is *the steady-state response* $x_s(t)$ *of* $x(t)$, *i.e.* for any $x(t_0)$ and $\omega(t_0)$, $\lim_{t\to+\infty} x(t) - \Pi_\infty(t)\omega(t) = 0$.

Remark 39. The definition of the function $\Pi_\infty(t)$ can be given as in (6.10) or, alternatively, as the unique solution of

$$\dot{\Pi}(t) = A\Pi(t) + BL - \Pi(t)\dot{\Lambda}(t, t_0)\Lambda(t, t_0)^{-1}, \tag{6.12}$$

with the initial condition $\Pi(t_0) = \Pi_\infty(t_0)$. From a practical point of view to determine $\Pi_\infty(t)$ from equation (6.10) or (6.12) is necessary to know the initial condition $\Pi_\infty(t_0)$. However, since the motion $\Pi_\infty(t)$ is attractive, any solution of the two equations converges to $\Pi_\infty(t)$, *i.e.* one could select $\Pi(t_0) = 0$. ■

Remark 40. The integral equation (6.10) or the differential equation (6.12) play the role of the Sylvester equation (6.2). Unlike when we have an implicit model of the signal generator, the matrix $\Pi_\infty(t)$ is in general a function of time. In fact, as remarked in section 6.2, infinitely many interpolation points arise whenever a periodic discontinuous signal is considered. Thus, a constant Π should have infinitely many rows and columns. ■

Definition 19. Consider system (6.1) and the signal generator (6.6). Suppose Assumptions 14, 15 and 16 hold and $\sigma(A) \subset \mathbb{C}_{<0}$. The function $C\Pi_\infty(t)\omega(t)$, where $\Pi_\infty(t)$ is the solution of equation (6.10) with $\Pi(t_0) = \Pi_\infty(t_0)$, is defined as the *moment of system (6.1) at* Λ.

Corollary 3. Consider the interconnection of system (6.1) with the signal generator (6.6). Suppose Assumptions 14, 15, 16 and 17 hold and $\sigma(A) \subset \mathbb{C}_{<0}$. Then the moment of (6.1) at Λ coincides with the steady-state response of the output of the interconnected system (6.1)-(6.6).

Proof. By the hypothesis on A and Assumptions 14, 15, 16 and 17 the steady-state response of (6.1) is well-defined and the relation $x_s(t) = \Pi_\infty(t)\omega(t)$, where $\Pi_\infty(t)$ is the unique solution of (6.10) with $\Pi(t_0) = \Pi_\infty(t_0)$, holds. By Theorem 19 the set $\mathcal{M}_\infty$ is attractive and the steady-state response of the output of the interconnected system corresponds to $C\Pi_\infty(t)\omega(t)$, which by definition is the moment of the system. □

The choice of defining the moment of (6.1) as in Definition 19 is justified by the equivalence, when an implicit model of (6.6) is available, between the new and the classical definition given in Chapter 2. In the next result we prove that, under certain hypotheses, the solutions of the Sylvester equation (6.2) and of the integral equation (6.10) are the same.

Theorem 20. Consider the signal generator (6.3), suppose $\sigma(A) \subset \mathbb{C}_{<0}$ and let $\sigma(S) \subset \mathbb{C}_{\geq 0}$. Then the unique solution of the integral equation (6.10) with $\Pi(t_0) = \Pi_\infty(t_0)$ coincides with the unique solution of the Sylvester equation (6.2).

Proof. Firstly note that Assumptions 14, 15 and 16 hold for the signal generator (6.3). Let $\tilde{\Pi}$ be the unique solution of the Sylvester equation $A\tilde{\Pi} + BL = \tilde{\Pi}S$ and $\Pi_\infty(t)$ be the unique solution of the integral equation (6.10) with $\Pi(t_0) = \Pi_\infty(t_0)$. Computing the derivative of the error $E(t) = \Pi_\infty(t) - \tilde{\Pi}$ yields

$$\dot{E}(t) = A\Pi_\infty(t) - \Pi_\infty(t)S + BL - 0 = A\Pi_\infty(t) - \Pi_\infty(t)S - (A\tilde{\Pi} - \tilde{\Pi}S) = AE(t) - E(t)S,$$

and (see [176, Section 11])

$$E(t) = e^{A(t-t_0)}E(t_0)e^{-S(t-t_0)},$$

which implies that $\Pi_\infty(t) - \tilde{\Pi}$ converges to zero. Since $\tilde{\Pi}$ is constant and $\Pi_\infty(t)$ is the limiting solution of (6.10), it follows that $\Pi_\infty(t_0) = \tilde{\Pi}$, $E(t_0) = 0$ and then $E(t) = 0$ for all $t \geq t_0$, which proves the claim. □

As anticipated in the previous section, we now focus our interest on periodic signals.

Corollary 4. Consider system (6.1) and the signal generator (6.6). Assume Assumptions 14, 15 and 16 hold. If for (6.6) the property

$$\Lambda(t, t_0) = D(t)\Lambda(t - T, t_0), \quad t \geq T + t_0, \tag{6.13}$$

holds with $D(t) \in \mathbb{R}^{\nu \times \nu}$ non-singular for all $t \in \mathbb{R}_{\geq 0}$ and $T \in \mathbb{R}_{>0}$, then equation (6.10) becomes

$$\Pi_\infty(t) = e^{AT}\Pi_\infty(t - T)D(t)^{-1} + \left[\int_{t-T}^{t} e^{A(t-\tau)}BL\Lambda(\tau, t_0)d\tau\right]\Lambda(t - T, t_0)^{-1}D(t)^{-1}. \tag{6.14}$$

If $D(t) = I$ then

$$\Pi_\infty(t) = (I - e^{AT})^{-1} \left[\int_{t-T}^{t} e^{A(t-\tau)} BL\Lambda(\tau, t_0) d\tau \right] \Lambda(t, t_0)^{-1}. \tag{6.15}$$

Proof. Equation (6.14) is obtained substituting

$$x_s(t) = \Pi_\infty(t)\omega(t) = \Pi_\infty(t)D(t)\omega(t - T)$$

and $\omega(t - T) = \Lambda(t - T, t_0)\omega(t_0)$ in

$$x_s(t) = e^{AT} x_s(t - T) + \int_{t-T}^{t} e^{A(t-\tau)} BL\Lambda(\tau, t_0)\omega(t_0) d\tau.$$

If $D(t) \equiv I$, $\omega(t) = \omega(t - T)$ and the steady-state of $x(t)$ is periodic with period T. Then $\Pi_\infty(t)\omega(t) = x_s(t) = x_s(t - T) = \Pi_\infty(t - T)\omega(t - T) = \Pi_\infty(t - T)\omega(t)$ which implies that $\Pi_\infty(t) = \Pi_\infty(t - T)$ and equation (6.15) follows. □

Remark 41. Equation (6.13) is a generalized form of equation (6.6) with the property (6.7) and describes a wide class of signals, possibly non-periodic. To show this note, for instance, that (6.3) can always be written as (6.13). In fact, for any $T \in \mathbb{R}_{\geq 0}$,

$$\omega(t) = e^{ST}\omega(t - T), \qquad \Lambda(t, t_0) = e^{S(t-t_0)}.$$

Thus (6.3) can be described by (6.13) with $D = e^{ST}$ for all t. ■

For periodic signals equation (6.15) defines in a simple form the periodic matrix $\Pi_\infty(t)$. In addition, note that this definition of $\Pi_\infty(t)$ does not depend upon the initial condition. As a result equation (6.15) can be used to find the initial condition $\Pi_\infty(0)$ of equation (6.10), as shown in the following examples.

6.3.1 Example: determination of $\Pi_\infty(t_0)$

Consider the interconnection of system (6.1) and (6.6). The matrices of (6.1) have been randomly generated with the function *rss* of MATLAB$^{\text{TM}}$. For the remaining of the

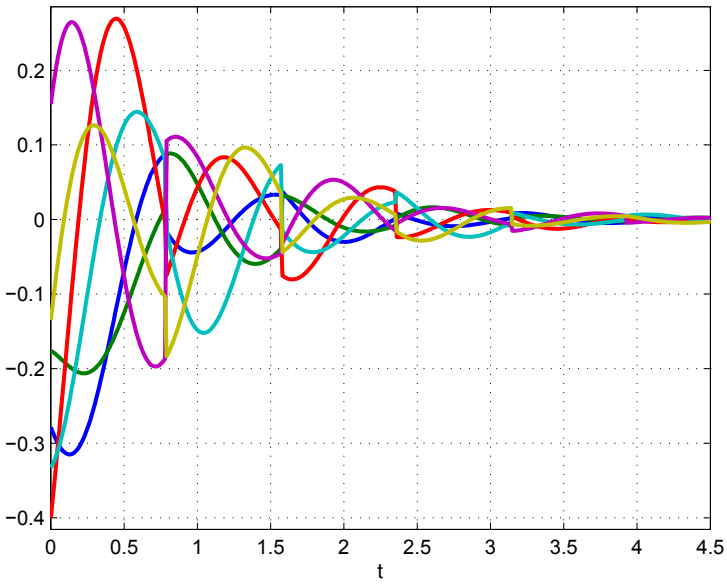

Figure 6.1: Time history of the entries of the error $\varepsilon(t) = \Pi_\infty(t) - \tilde{\Pi}(t)$, computed from equation (6.10) with initial condition $\Pi(0) = \Pi_\infty(0)$ and $\Pi(0) = \tilde{\Pi}(0) = 2\Pi_\infty(0)$, respectively.

chapter we use the selection

$$A = \begin{bmatrix} -2.439 & 2.337 & -1.776 \\ -2.933 & -1.096 & 4.221 \\ 0.09223 & -4.579 & -1.537 \end{bmatrix}, \quad \begin{aligned} B &= \begin{bmatrix} 0 & -0.7648 & -1.402 \end{bmatrix}^T, \\ C &= \begin{bmatrix} 0 & 0.4882 & 0 \end{bmatrix}, \\ L &= \begin{bmatrix} 1 & 1 \end{bmatrix}. \end{aligned} \tag{6.16}$$

Let $t_0 = 0$ and consider the matrix of square waves

$$\Lambda_\sqcap(t,0) = \begin{bmatrix} \sqcap\left(\frac{2\pi}{T}t + \frac{\pi}{2}\right) & -\sqcap\left(\frac{2\pi}{T}t\right) \\ \sqcap\left(\frac{2\pi}{T}t\right) & \sqcap\left(\frac{2\pi}{T}t + \frac{\pi}{2}\right) \end{bmatrix}. \tag{6.17}$$

From equation (6.15) computed for $t = 0$,

$$\begin{aligned} \Pi_\infty(0) = -A^{-1}(I - e^{AT})^{-1} \Big[&\left(e^{\frac{3}{4}AT} - e^{AT} + e^{\frac{1}{2}AT} - e^{\frac{1}{4}AT}\right) BL \\ &+ \left(e^{\frac{1}{2}AT} - e^{\frac{3}{4}AT} + e^{\frac{1}{4}AT} - I\right) BL\Lambda_\sqcap\left(\frac{T}{4}, 0\right) \Big], \end{aligned} \tag{6.18}$$

which for $T = \pi$ yields

$$\Pi_\infty(0) = \begin{bmatrix} 0.1760 & -0.2786 \\ 0.3327 & -0.4004 \\ 0.1349 & 0.1546 \end{bmatrix}.$$

Fig. 6.1 shows the time history of the error $\varepsilon(t) = \Pi_\infty(t) - \tilde{\Pi}(t)$, where $\Pi_\infty(t)$ is computed from equation (6.10) with initial condition $\Pi(0) = \Pi_\infty(0)$ defined in (6.18) and $\tilde{\Pi}(t)$ is computed from equation (6.10) with initial condition $\Pi(0) = \tilde{\Pi}(0) = 2\Pi_\infty(0)$. As proved in Theorem 19 the solution $\Pi_\infty(t)$ is exponentially attractive.

Remark 42. Signals similar to the ones given in (6.17) can be generated by an implicit (nonlinear) model. In fact, note that the interconnection between the signals generated by the equations

$$\dot{\omega} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \omega, \quad u = \text{sign}\left(\begin{bmatrix} 1 & 0 \end{bmatrix} \omega \right),$$

and system (6.1) can be studied with the nonlinear techniques proposed in Chapter 2. However, in this case the partial differential equation (2.23) has to be solved to compute the moment of the system and the general solution of this problem, especially when involving non-smooth functions, is difficult to find. ■

Remark 43. While from a computational point of view solving the partial differential equation (2.23) and solving the integral and the matrix inversion in $\Pi_\infty(t)$ may be equally expensive, $\Pi_\infty(t)$ is an exact solution, while the solution of the partial differential equation would probably have to be approximated. Note also that, exploiting the periodicity of the steady-state, $\Pi_\infty(t)$ has to be computed only over a period. This can be done off-line and the obtained values can then be used on-line for any time interval. ■

Remark 44. Signals similar to the ones given in (6.17) can be also generated by hybrid systems. In fact, the signal generator (6.6) can represent a class of explicit linear hybrid systems. Thus, the results of the chapter can be used to extend the model reduction method by moment matching to this class of hybrid systems. ■

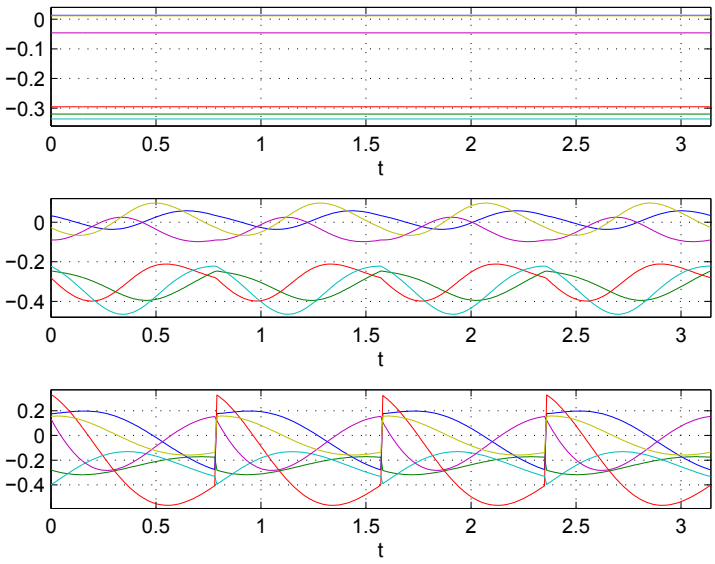

Figure 6.2: Time history of the entries of the matrices $\Pi_\sim$ (top), $\Pi_\wedge$ (middle) and $\Pi_\sqcap$ (bottom).

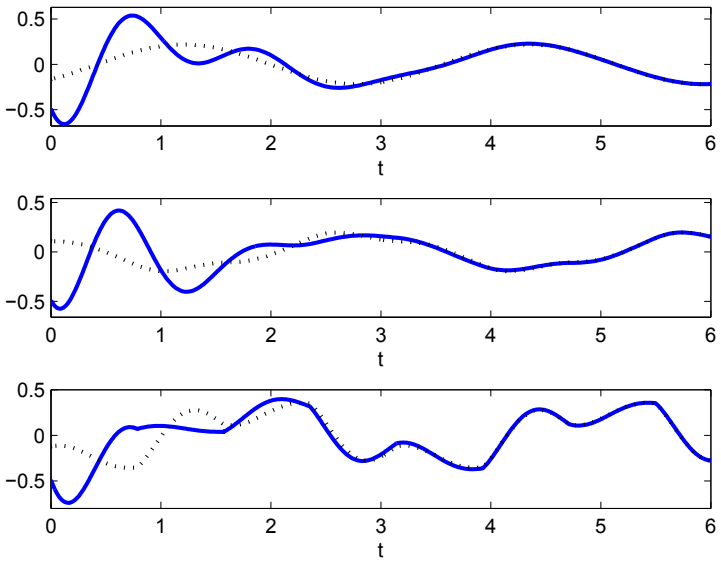

Figure 6.3: Time history of the output (solid lines) $y_\sim$ (top), $y_\wedge$ (middle) and $y_\sqcap$ (bottom) and time histories of the steady-state of the output (dotted lines) computed as $C\Pi_\sim\omega$, $C\Pi_\wedge\omega$ and $C\Pi_\sqcap\omega$.

6.3.2 Example: time-varying moments

Consider the interconnection of system (6.1) and (6.6), the selection (6.16), the matrix (6.17) and the two matrices

$$\Lambda_{\sim}(t,0) = \begin{bmatrix} \cos\left(\frac{2\pi}{T}t\right) & -\sin\left(\frac{2\pi}{T}t\right) \\ \sin\left(\frac{2\pi}{T}t\right) & \cos\left(\frac{2\pi}{T}t\right) \end{bmatrix}$$

and

$$\Lambda_{\wedge}(t,0) = \begin{bmatrix} \wedge\left(\frac{2\pi}{T}t+\frac{\pi}{2}\right) & -\wedge\left(\frac{2\pi}{T}t\right) \\ \wedge\left(\frac{2\pi}{T}t\right) & \wedge\left(\frac{2\pi}{T}t+\frac{\pi}{2}\right) \end{bmatrix},$$

where

$$\wedge(t) = \begin{cases} t \mod (\pi), & (k-1)\pi \leq t < k\pi, \\ 1-(t \mod (\pi)), & k\pi \leq t < (k+1)\pi, \end{cases}$$

with $k = 1, 3, 5, \ldots, +\infty$ and period $T = \pi$. Let $\Pi_{\sim}$, $\Pi_{\wedge}$ and $\Pi_{\sqcap}$ be the solutions of equation (6.15) and $y_{\sim}$, $y_{\wedge}$ and $y_{\sqcap}$ be the outputs of system (6.1) for $\Lambda_{\sim}$, $\Lambda_{\wedge}$ and $\Lambda_{\sqcap}$, respectively. Fig. 6.2 shows the entries of the matrices $\Pi_{\sim}$ (top), $\Pi_{\wedge}$ (middle) and $\Pi_{\sqcap}$ (bottom). We note that $\Pi_{\sim}$ is constant, whereas $\Pi_{\wedge}$ and $\Pi_{\sqcap}$ are periodic. Fig. 6.3 shows the time history of the output (solid lines) $y_{\sim}$ (top), $y_{\wedge}$ (middle) and $y_{\sqcap}$ (bottom) and of the steady-state value of the output (dotted lines) computed as $C\Pi_{\sim}\omega$, $C\Pi_{\wedge}\omega$ and $C\Pi_{\sqcap}\omega$, *i.e.* the moment of system (6.1) at Λ. We note that the outputs of the system approach the steady-state responses, as expected.

6.4 Reduced order models in explicit form

In this section, exploiting the new definition of moment, a family of systems achieving moment matching is introduced and connections with the families of models given in Chapter 2 are drawn.

Definition 20. Consider system (6.1) and the signal generator (6.6). Suppose Assumptions 14, 15, 16 and 17 hold and $\sigma(A) \subset \mathbb{C}_{<0}$. Then the system described by the equations

$$\xi(t) = \widetilde{F}(t,t_0)\xi_0 + \int_{t_0}^{t} \widetilde{G}(t-\tau)u(\tau)d\tau, \qquad \psi(t) = \widetilde{H}(t)\xi(t), \tag{6.19}$$

with $\xi(t) \in \mathbb{R}^\nu$, $\widetilde{F}(t,t_0) \in \mathbb{R}^{\nu\times\nu}$, $\widetilde{G}(t) \in \mathbb{R}^{\nu\times 1}$, is a *model of system (6.1) at (6.6)*, if there exists a unique solution $P_\infty(t)$ of the equation

$$P(t) = \left(\widetilde{F}(t,t_0)P(t_0) + \int_{t_0}^{t} \widetilde{G}(t-\tau)L\Lambda(\tau,t_0)d\tau\right)\Lambda^{-1}(t,t_0), \tag{6.20}$$

with $P(t_0) = P_\infty(t_0)$ such that for any $P(t_0)$, $\lim_{t\to+\infty} P(t) - P_\infty(t) = 0$ and

$$C\Pi_\infty(t) = \widetilde{H}(t)P_\infty(t), \tag{6.21}$$

where $\Pi_\infty(t)$ is the unique solution of (6.10), with $\Pi(t_0) = \Pi_\infty(t_0)$. System (6.19) is a *reduced order model of system (6.1) at* (Λ, L) if $\nu < n$.

Remark 45. In addition to the assumptions which guarantee the existence of the moment of system (6.1), we need two additional conditions for the existence of the reduced order model (6.19). Firstly the reduced order model has to be asymptotically stable, *i.e.* for each $\varepsilon > 0$, there exists $\delta_\varepsilon > 0$ such that if $||\xi(0)|| < \delta_\varepsilon$ then $||\xi(t)|| < \varepsilon$ and there exists a $\delta > 0$ such that if $||\xi(0)|| < \delta$ then $\lim_{t\to+\infty} \xi(t) = 0$. This guarantees that the function $P_\infty(t)$ exists, *i.e.* the moment of system (6.19) exists. Secondly we need $P_\infty(t)$ to be non-singular for all $t \in \mathbb{R}_{\geq 0}$. This guarantees the existence of $H(t)$ for all $t \in \mathbb{R}_{\geq 0}$, *i.e.* the moments are matched. ■

We now discuss how to simplify the family of models (6.19) to achieve moment matching with additional constraints.

Property 1: ideally we would like to have $P_\infty(t) = I$ (as in Chapters 2 and 3), which has proved to give a remarkable simplification in the definition of the reduced order model. We recall that the selection $P = I$ in Chapter 2 yields a family of systems, parametrized by a matrix $G \in \mathbb{R}^\nu$, which is complete, *i.e.* the family contains all systems of dimensions ν achieving moment matching.

Property 2: at the same time we would like to bring the first equation of (6.19) to a form for which we can easily enforce additional constraints. This form is described by the selection

$$\widetilde{F}(t,t_0) = e^{F(t-t_0)}, \qquad \widetilde{G}(t) = e^{Ft}G, \tag{6.22}$$

for some F and G, which makes the first equation of (6.19) a representation in explicit form of a linear time-invariant system which has an implicit model.

However, differently from the family of reduced order models given in Chapter 2, it is not possible for the reduced order model (6.19) to satisfy both properties, namely having an implicit model and be such that $P_\infty(t) = I$. Note, for instance, that if $\omega(t)$ belongs to the class of signals satisfying (6.7) and we use the selection (6.22), then the steady-state $\xi_s(t)$ is periodic and so it is $P_\infty(t)$ (see the proof of Corollary 4). Thus, this first route in which we simplify the problem fixing $P_\infty(t)$ brings the problem of determining $\widetilde{F}(t, t_0)$ and $\widetilde{G}(t)$ from equation (6.20) (which may not be easy) and, at the same time, such that additional properties, *e.g.* stability, can be enforced on the reduced order model. As a second route we can fix $\widetilde{F}(t, t_0)$ and $\widetilde{G}(t)$ with the selection given in (6.22) which, however, brings the problem of finding $P_\infty(t)$ and solving equation (6.21) with respect to $\widetilde{H}(t)$.

Both the choices are viable, however, it is easier to follow the second route, namely using the selection (6.22) and solving numerically (6.20) and (6.21). In this case the dynamics of the state ξ can be described by a linear system in implicit form for which the solution to the problem of selecting F and G such that additional properties are satisfied is given in Chapter 2.

Remark 46. Both routes lead to models which are equivalent to the reduced order models given in Chapter 2 if $\Lambda(t, t_0) = e^{S(t-t_0)}$. Following the first route easy computation shows that $\widetilde{F}(t, t_0) = e^{(S-GL)(t-t_0)}$, $\widetilde{G}(t) = e^{(S-GL)t}G$, with G free, is a solution of equation (6.20) with $P_\infty(t) = I$. Following the second route, by Theorem 20, equation (6.20) is equivalent to the Sylvester equation associated to the general family of reduced order models given in Chapter 2. ■

Remark 47. As already discussed, if $\Lambda(t, t_0) = e^{S(t-t_0)}$ and $P_\infty(t) = P = I$, then in (6.22) $F = S - GL$ with G free. The family of models (6.19) reduces to the family parametrized in G given in Chapter 2. Note also that a different selection $\bar{P} \neq P = I$ yields the same class of models through a change of coordinates in the state space representation, namely $\bar{F} = \bar{P}(S - GL)\bar{P}^{-1}$, $\bar{G} = \bar{P}G$, $\bar{H} = C\Pi\bar{P}^{-1}$, and there is still only one free parameter G. On the other hand, in the selection (6.22), we have two free parameters, namely F and G which can be totally independent of each other (this is possible since $\widetilde{H}(t)$ in (6.21) is not constant). ■

We summarize these observations in the next statement in which we give a family

of reduced order models in the case of periodic input signals.

Proposition 9. Consider system (6.1) and the signal generator (6.6) with the property (6.7). Suppose Assumptions 14, 15, 16 and 17 hold and $\sigma(A) \subset \mathbb{C}_{<0}$. Then the system

$$\dot{\xi} = F\xi + Gu, \qquad \psi = C\Pi_\infty(t)P_\infty(t)^{-1}\xi(t), \tag{6.23}$$

with $\xi(t) \in \mathbb{R}^\nu$, $F \in \mathbb{R}^{\nu\times\nu}$, $G \in \mathbb{R}^{\nu\times 1}$ and $\Pi_\infty(t)$ defined in (6.15), is a *model of system (6.1) at* (Λ, L), if $\sigma(F) \subset \mathbb{C}_{<0}$ and

$$P_\infty(t) = (I - e^{FT})^{-1}\left[\int_{t-T}^{t} e^{F(t-\tau)}GL\Lambda(\tau, t_0)d\tau\right]\Lambda(t, t_0)^{-1}, \tag{6.24}$$

is non-singular for all $t \in \mathbb{R}_{\geq 0}$.

Proof. In (6.19) select $\widetilde{F}(t, t_0)$ and $\widetilde{G}(t)$ as in (6.22). Since the signal generator (6.6) is periodic and $\sigma(F) \subset \mathbb{C}_{<0}$, equation (6.20) takes, by Corollary 4 the form given in (6.24). By assumption, $P_\infty(t)$ is non-singular and, thus, the matching condition (6.21) can be solved with respect to $\widetilde{H}(t)$, giving $\widetilde{H}(t) = C\Pi_\infty(t)P_\infty(t)^{-1}$. □

We conclude this section with a numerical example to illustrate this last result.

6.4.1 Example: a numerical example

Consider the interconnection of system (6.1) and (6.6) with Λ defined as in (6.17). The matrices of (6.1), with $n = 15$, have been selected as $A = \text{diag}(-1/n, -2/n, \ldots, -1)$, $B = [-1/n \ -2/n \ \ldots \ -1]^\top$, $C = B^\top$, with the initial state $x(t_0) = [n/2 \ n/2{-}1 \ \ldots \ -n/2{+}1]^\top$. Using the family (6.23), two reduced order models of system (6.1) have been computed. The first one is described by the selection $F = \text{diag}(-1, -0.0667)$ (which are the slowest and fastest modes of A) and $G = [0.0771 \ \ -0.0251]^\top$, while the second one by $F = \text{diag}(-0.9333, -1)$ (which are the two fastest modes of A) and $G = [0.0891 \ 0.008]^\top$. In both case G is such that $P_\infty(t)$ is non-singular for all $t \in \mathbb{R}_{\geq 0}$. The top graph of Fig. 6.4 shows the time history of the output y (solid line), of the output ψ_I of the first reduced order model (dotted line) and of the output ψ_{II} of the second reduced order model (dash-dotted line). The bottom graph shows the time history of $|y - \psi_I|$ (dotted line) and of $|y - \psi_{II}|$ (dash-dotted line). The

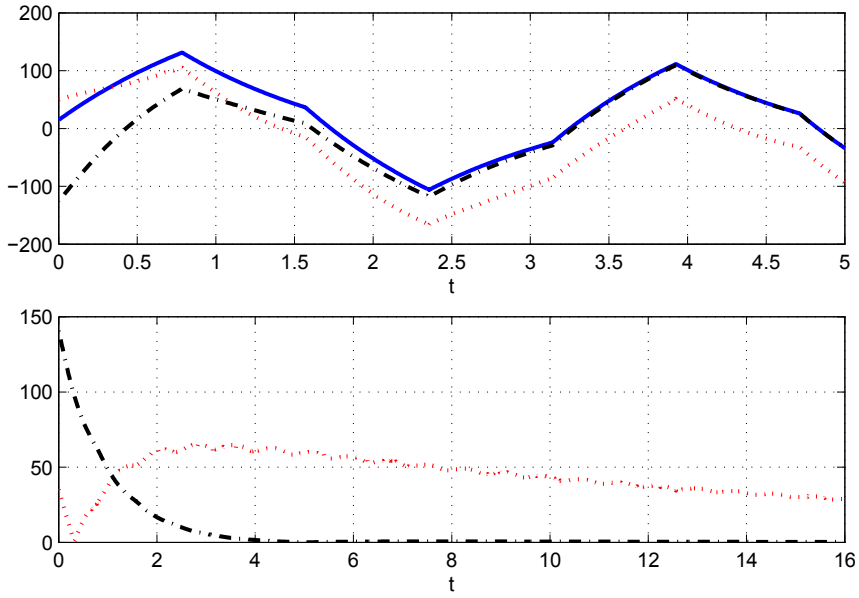

Figure 6.4: Top: time history of the output y (solid line) of system (6.1), with (6.17), of the output ψ_I (dotted line) of system (6.23) and of the output ψ_{II} (dash-dotted line) of system (6.23) . Bottom: time history of $|y - \psi_I|$ (dotted line) and of $|y - \psi_{II}|$ (dash-dotted line).

figure shows that the output response of the second model converges to the output of the system much quicker than the one of the first model.

Remark 48. The dynamics of (6.23) is the dynamics of a linear system in implicit form "filtered" with the dynamics of a linear system in explicit form. ■

6.5 Conclusion

In this chapter the limitations of the description of moment which uses the matrix S have been investigated. With the final aim of solving the model reduction problem for a class of input signals generated by a linear exogenous system which does not have an implicit (differential) form, a time-varying parametrization of the steady-state of the system has been used to extend, exploiting an integral matrix equation, the definition of moment to this class of input signals. The equivalence of the new definition and the one based on the Sylvester equation has been proved under specific conditions. Special attention has been given to periodic signals for the wide range of practical applications where these are used. Reduced order models matching the steady-state response of explicit signal generators

have been given for linear systems and several connections between classical reduced order models and the new ones have been drawn. The chapter is completed with numerical examples.

Chapter 7

Model reduction for linear singular systems

7.1 Introduction

Singular systems, also known as *generalized state-space systems*, *descriptor systems* or *differential-algebraic systems*, have been largely investigated because of their capacity of modeling a variety of physical behaviors such as power systems, electrical networks, chemical processes, biological systems and social economic systems, see [180], [181]. The important characteristic of this class of systems is that they combine dynamic (differential) equations, referred as *slow subsystem* and static (algebraic) equations, referred as *fast subsystem*. They are often the result of the interconnection of several subsystems and the overall singular system has often large dimensionality. This justifies the interest in the model reduction problem for this class of systems. A first result of model reduction for singular systems has been proposed in [182] in which the chained aggregation method has been used. As pointed out in [183] this method is computationally intensive and another approach based on balanced realization has been proposed. However, the reduced order model obtained therein is a non-singular one and as a consequence the technique is not able to maintain the impulsive characteristic of singular systems. In [184] a singular reduced order model based on the Nehari's approximation algorithm [185] has been presented. Other results on model reduction via covariance approximation [186] and using singular value decomposition [187] have also been proposed. In [188] the authors propose

a moment matching method to reduce multi-input, multi-output (MIMO) systems. The method proposed therein is based in isolating the slow and fast subsystem and in reducing only the slow part (the non-singular subsystem), leaving unreduced the fast part (the singular subsystem).

In this chapter we extend the model reduction method developed in Chapter 2 to singular systems. Combining the interpolation-based and the steady-state-based description of moment a partitioned projector Π is computed. In the projector, the contribution of the fast subsystem is separated by the contribution of the slow subsystem. The information on the fast subsystem is clearly encoded in the projector and it is not lost by the moment matching technique. Moreover, the output of reduced order models based on this projector approximates the output of the system only when consistent initial conditions are taken in account. In addition, the partitioned projector allows to define several families of reduced order models which have great flexibility in maintaining independently specific properties of the slow and fast subsystems. Thus, purely fast, purely slow and a "simple" family of reduced order models are proposed and the possibility of matching with the impulsive controllability constraint is discussed.

Remark 49. The contribution of this chapter is substantially different from [188], in both results and meaning. In [188] the fast subsystem is not reduced at all. The reason the authors give is that if the algebraic part is reduced than the frequency error between the system to be reduced and the reduced order model grows unbounded. Thus, their projector is partitioned in a classical projector (for non-singular systems) and the identity. In the method proposed here the reduction can operate independently on the algebraic part, on the differential part, or on a combination thereof. Moreover, we believe that the ability of reducing the algebraic part is relevant even though the error grows unbounded. In fact, as shown in the example in Section 7.4.5, a reduced order model which has the same output response of the system to be reduced can be determined even though for high frequencies the error grows unbounded. Moreover, in many applications the unbounded part is filtered out with a low-pass filter or, alternatively, the behavior for high frequencies is not of interest as long as some degree of divergence at high frequency is maintained (both the frequency response of the system and of the reduced order model grow unbounded). Finally note that the results in this chapter are instrumental in deriving a model reduction

theory for nonlinear singular systems. ■

The rest of the chapter is organized as follows. In Section 7.2 we recall some basic results for singular systems which are instrumental for the remaining. In Section 7.3 the interpolation-based description of moment is extended to singular systems. Then, combining it with the steady-state-based description of moment, a new formulation of the projector Π is given. In Section 7.4 several families of reduced order models are given and the possibility to retain, reduce or eliminate the fast subsystem is investigated. Moreover, a result to enforce impulsive controllability on the reduced order model is proposed. Two numerical examples are used to illustrate the results. Finally Section 7.5 contains some concluding remarks and future directions of research.

All the results of this chapter are original contributions developed in fulfillment of my PhD course and they have been published in the conference paper [18]. The problem of model reduction by moment matching for nonlinear singular systems is currently under investigation.

7.2 Preliminaries on Singular Systems

In this section some basic results for singular systems are recalled. Consider a linear, single-input, single-output, continuous-time, singular system described by the equations[1]

$$E\dot{x} = Ax + Bu, \qquad y = Cx, \tag{7.1}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $E \in \mathbb{R}^{n\times n}$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$ and $C \in \mathbb{R}^{1\times n}$. Assume that u is piecewise continuously differentiable.

Definition 21. [180] Let $\mathcal{E} \subset \mathbb{C}$ be the set of complex numbers s such that $\det(sE-A) \neq 0$. For any two matrices E and A, the pencil (E, A) is called *regular* if $\mathcal{E} \not\equiv \emptyset$.

Lemma 16. [180] The pencil (E, A) is regular if and only if there exist two nonsingular matrices Q and Γ such that

$$QE\Gamma = \mathrm{diag}(I, N), \qquad QA\Gamma = \mathrm{diag}(A_1, I),$$

[1] The results can be extended to MIMO systems as discussed in Chapter 5.

where $N \in \mathbb{R}^{n_2 \times n_2}$ is nilpotent with degree[2] h and $A_1 \in \mathbb{R}^{n_1 \times n_1}$, with $n_1 + n_2 = n$.

Assume system (7.1) is regular, then Q and Γ can be selected such that system (7.1) can be written in the so-called *first equivalent form*, namely

$$\begin{array}{c} \textit{slow} \\ \textit{subsystem} \end{array} \left\{ \begin{array}{rcl} \dot{x}_1 &=& A_1 x_1 + B_1 u, \\ y_1 &=& C_1 x_1, \end{array} \right. \qquad \begin{array}{c} \textit{fast} \\ \textit{subsystem} \end{array} \left\{ \begin{array}{rcl} N\dot{x}_2 &=& x_2 + B_2 u, \\ y_2 &=& C_2 x_2, \end{array} \right. \tag{7.2}$$

$$y = y_1 + y_2$$

with the coordinate transformation $\begin{bmatrix} x_1^\top & x_2^\top \end{bmatrix}^\top = \Gamma^{-1} x$, where $x_1 \in \mathbb{R}^{n_1 \times n_1}$, $x_2 \in \mathbb{R}^{n_2 \times n_2}$, and

$$QE\Gamma = \operatorname{diag}(I, N), \quad QB = \begin{bmatrix} B_1^\top & B_2^\top \end{bmatrix}^\top, \quad QA\Gamma = \operatorname{diag}(A_1, I), \quad C\Gamma = \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

The state response of system (7.1) is given by

$$x(t) = \Gamma \begin{bmatrix} I \\ 0 \end{bmatrix} \left(e^{A_1 t} x_1(0) + \int_0^t e^{A_1 (t-\tau)} B_1 u(\tau) d\tau \right) - \Gamma \begin{bmatrix} 0 \\ I \end{bmatrix} \sum_{i=0}^{h-1} N^i B_2 \frac{d^i}{dt^i} u(t). \tag{7.3}$$

Differently from standard linear systems, singular systems have a unique solution only if $x(0) = \Gamma \begin{bmatrix} x_1^\top(0) & x_2^\top(0) \end{bmatrix}^\top$ is a *consistent* initial condition, namely it is such that

$$x(0) = \Gamma \begin{bmatrix} I \\ 0 \end{bmatrix} x_1(0) - \Gamma \begin{bmatrix} 0 \\ I \end{bmatrix} \sum_{i=0}^{h-1} N^i B_2 \frac{d^i}{dt^i} u(t) \Big|_{t=0},$$

where $x_1(0)$ can be freely selected.

7.3 Definition of moment

In this section we establish a one-to-one relation between the moments of system (7.1), the unique solution of a generalized Sylvester equation and the steady-state response of the output of a particular interconnected system.

[2] A nilpotent matrix is a square matrix N such that $N^k = 0$ for some positive integer k. The smallest such k is called the degree of N.

Let

$$W(s) = C(sE - A)^{-1}B = C_1(sI - A_1)^{-1}B_1 + C_2(sN - I)^{-1}B_2,$$

be the transfer function associated to system (7.1) and assume that the system is minimal, *i.e.* as for non-singular systems (see [180, Theorem 2-6.3]), controllable and observable.

7.3.1 Interpolation-based description of moment

We begin extending the interpolation-based description of moment. This, in conjunction with the steady-state-based description of moment, allows to obtain a special structure for the projector to be used in the model reduction process.

Definition 22. Let $s_i \in \mathcal{E}$. The 0-*moment of system (7.1) at* s_i is the complex number

$$\eta_0(s_i) = C(s_iE - A)^{-1}B.$$

The k-*moment of system (7.1) at* s_i is the complex number

$$\eta_k(s_i) = \frac{(-1)^k}{k!}\left[\frac{d^k}{ds^k}C(sE - A)^{-1}B\right]_{s=s_i},$$

with $k \geq 1$ integer.

Note that each moment is the sum of two contributions, one depending on the slow subsystem and one depending on the fast subsystem, *i.e.*

$$\eta_k(s_i) = (-1)^k C_1(s_iI - A_1)^{-(k+1)}B_1 + (-1)^k C_2N^k(s_iN - I)^{-(k+1)}B_2, \tag{7.4}$$

and if $k \geq h$,

$$\eta_k(s_i) = (-1)^k C_1(s_iI - A_1)^{-(k+1)}B_1.$$

The next result gives a relation between the moments and the solution of a generalized Sylvester equation.

Lemma 17. Let $s_i \in \mathcal{E}$. Consider system (7.1), then

$$\begin{bmatrix} \eta_0(s_i) & \dots & \eta_k(s_i) \end{bmatrix} = C\widetilde{\Pi}\Psi_k,$$

where $\Psi_k = \mathrm{diag}(1, -1, 1, \dots, (-1)^k) \in \mathbb{R}^{(k+1)\times(k+1)}$ and $\widetilde{\Pi}$ is the unique solution of the

generalized Sylvester equation

$$A\widetilde{\Pi} + BL_k = E\widetilde{\Pi}\Sigma_k, \tag{7.5}$$

with $L_k = [1\ 0\ \dots\ 0] \in \mathbb{R}^{(k+1)}$ and

$$\Sigma_k = \begin{bmatrix} s_i & 1 & 0 & \dots & 0 \\ 0 & s_i & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & s_i & 1 \\ 0 & \dots & \dots & 0 & s_i \end{bmatrix} \in \mathbb{R}^{(k+1)\times(k+1)}.$$

Proof. Let $\widetilde{\Pi} = [\widetilde{\Pi}_0\ \widetilde{\Pi}_1\ \dots\ \widetilde{\Pi}_k]$ and note that (7.5) can be rewritten as

$$\begin{aligned} A\widetilde{\Pi}_0 \ + B &= E\widetilde{\Pi}_0 s_i, \\ A\widetilde{\Pi}_1 &= E\widetilde{\Pi}_1 s_i + E\widetilde{\Pi}_0, \\ &\vdots \\ A\widetilde{\Pi}_k &= E\widetilde{\Pi}_k s_i + E\widetilde{\Pi}_{k-1}. \end{aligned}$$

As a result

$$\begin{aligned} \widetilde{\Pi}_0 &= (sE - A)^{-1}B, \\ \widetilde{\Pi}_1 &= -(sE - A)^{-1}E(sE - A)^{-1}B = \left[\frac{d}{ds}(sE - A)^{-1}B\right]_{s_i}, \\ &\vdots \\ \widetilde{\Pi}_k &= \frac{1}{k!}\left[\frac{d^k}{ds^k}(sE - A)^{-1}B\right]_{s_i}, \end{aligned}$$

which proves the claim. □

Exploiting this lemma the following result, which is a direct extension of Lemma 2, holds.

Theorem 21. Consider system (7.1) and suppose that for a set of numbers s_i, with $i = 1, \dots, \eta$, $s_i \in \mathcal{E}$. Then there exists a one-to-one relation between the moments $\eta_0(s_1)$, $\dots$, $\eta_{k_1-1}(s_1)$, $\dots$, $\eta_0(s_\eta)$, $\dots$, $\eta_{k_\eta-1}(s_\eta)$ and the matrix $C\Pi$, where Π is the unique solution of the generalized Sylvester equation

$$A\Pi + BL = E\Pi S, \tag{7.6}$$

with $S \in \mathbb{R}^{\nu\times\nu}$ any non-derogatory matrix with characteristic polynomial

$$p(s) = \prod_{i=1}^{\eta}(s - s_i)^{k_i}, \tag{7.7}$$

where $\nu = \sum_{i=1}^{\eta} k_i$ and L such that the pair (L, S) is observable.

Proof. We begin proving that if $\sigma(S) \subset \mathcal{E}$ then equation (7.6) has a unique solution. Using the vectorization transformation and the Kronecker product properties we obtain the solution

$$\operatorname{vec}(\Pi) = -(I \otimes A - S^\top \otimes E)^{-1} \operatorname{vec}(BL),$$

which is unique if and only if

$$\det\left(I \otimes A - S^\top \otimes E\right) \neq 0. \tag{7.8}$$

Let R be an invertible matrix such that $R^{-1}SR = J_S$, where J_S is the complex Jordan form of S. Multiplying on the left by $(R \otimes I)$ and on the right by $(R^{-1} \otimes I)$ we obtain the equation

$$\begin{aligned}\det\left(I \otimes A - S^\top \otimes E\right) &= \det\left((R \otimes I)(I \otimes A - S^\top \otimes E)(R^{-1} \otimes I)\right)\\ &= \det\left(I \otimes A - J_S^\top \otimes E\right).\end{aligned}$$

Then (7.8) holds if and only if

$$\prod_{i=1}^{\eta} \det\left(A - s_i E\right) \neq 0.$$

The remaining of the theorem can be proved following the same steps in the proof of Theorem 5. □

7.3.2 Steady-state-based description of moment

In this section we extend the steady-state-based description of moment to linear singular systems. This is instrumental to derive the main technical result of the chapter, namely a partition of the projector Π in slow and fast parts.

Theorem 22. Let $S \in \mathbb{R}^{\nu\times\nu}$ be any non-derogatory matrix with characteristic polyno-

mial (7.7). Consider system (7.1) and suppose that $\sigma(S) \subset \mathcal{E}$ and $\sigma(A) \subset \mathbb{C}_{<0}$. Consider the interconnection of system (7.1) with the system

$$\dot{\omega} = S\omega, \qquad u = L\omega, \tag{7.9}$$

with L and $\omega(0)$ such that the triple $(L, S, \omega(0))$ is minimal. Then, for all consistent initial conditions, there exists a one-to-one relation between the moments $\eta_0(s_1)$, ..., $\eta_{k_1-1}(s_1)$, ..., $\eta_0(s_\eta)$, ..., $\eta_{k_\eta-1}(s_\eta)$ and the steady-state response of the output of such interconnected system.

Proof. Consider the interconnection of system (7.1) with system (7.9). By the assumptions on the pencil (E, A) and $\sigma(S)$, the interconnected system has a well-defined invariant manifold given by $\mathcal{M} = \{(x, \omega) \in \mathbb{R}^{n+\nu} : x = \Pi\omega\}$, with Π the unique solution of the generalized Sylvester equation (7.6). We prove now that $\mathcal{M}$ is attractive. Consider the equation

$$\dot{\overbrace{Ex - E\Pi\omega}} = Ax + BL\omega - E\Pi S\omega,$$

in which substituting (7.6), yields

$$\dot{\overbrace{Ex - E\Pi\omega}} = A(x - \Pi\omega).$$

Let $z = x - \Pi\omega$, then

$$E\dot{z} = Az.$$

The state response of this last equation is

$$z(t) = \Gamma \begin{bmatrix} I \\ 0 \end{bmatrix} e^{A_1 t} z(0).$$

Thus, the output response of the interconnected system is

$$y(t) = C\Pi\omega(t) + C\Gamma \begin{bmatrix} I \\ 0 \end{bmatrix} e^{A_1 t}(x_1(0) - \Pi\omega(0)).$$

The claim follows observing that

$$C\Gamma \begin{bmatrix} I \\ 0 \end{bmatrix} e^{A_1 t}(x_1(0) - \Pi\omega(0))$$

describes the transient response which vanishes exponentially. □

We are now ready to give the main result of this chapter.

Theorem 23. Let $S \in \mathbb{R}^{\nu\times\nu}$ be any non-derogatory matrix with characteristic polynomial (7.7). Consider system (7.2) and suppose that $\sigma(S) \cap \sigma(A_1) = \emptyset$. Then there exists a one-to-one relation between the moments $\eta_0(s_1), \ldots, \eta_{k_1-1}(s_1), \ldots, \eta_0(s_\eta), \ldots, \eta_{k_\eta-1}(s_\eta)$ and the matrix $C\widehat{\Pi}$, where $\widehat{\Pi} = \Gamma^{-1}\Pi = \begin{bmatrix} \Pi_1^\top & \Pi_2^\top \end{bmatrix}^\top$ is the unique solution of the equations

$$\begin{aligned} A_1\Pi_1 - \Pi_1 S &= -B_1 L, \\ \Pi_2 &= -\sum_{i=0}^{h-1} N^i B_2 L S^i. \end{aligned} \tag{7.10}$$

Proof. In steady-state, using the change of coordinates

$$\begin{bmatrix} x_{1ss} \\ x_{2ss} \end{bmatrix} = \Gamma^{-1} x_{ss} = \Gamma^{-1}\Pi\omega = \begin{bmatrix} \Pi_1 \\ \Pi_2 \end{bmatrix} \omega$$

in (7.2), yields

$$\begin{aligned} \Pi_1 S\omega &= A_1\Pi_1\omega + B_1 L\omega, \\ N\Pi_2 S\omega &= \Pi_2\omega + B_2 L\omega, \end{aligned}$$

from which ω can be eliminated because the two equations hold for any ω. We readily obtain the first of equations (7.10). Applying the vectorization transformation and the Kronecker product properties to the second equation we see that Π_2 is unique if and only if

$$\det(I - s_i N) \neq 0,$$

for any $s_i \in \sigma(S)$. This holds by definition for any $s_i \in \mathbb{C}$ since N is a nilpotent matrix. Then, Π is unique if and only if $\sigma(S) \cap \sigma(A_1) = \emptyset$. To obtain the explicit expression of

Π_2, we substitute $x_{2ss} = \Pi_2\omega$ in the last n_2 equations in (7.3) obtaining

$$\Pi_2\omega = -\sum_{i=0}^{h-1} N^i B_2 L \frac{d^i}{dt^i}\omega.$$

Observing that $\dfrac{d^i}{dt^i}\omega = S^i\omega$, yields

$$\Pi_2 = -\sum_{i=0}^{h-1} N^i B_2 L S^i,$$

since the equation holds for any ω. □

This result has multiple implications. The contribution to the moment of the fast subsystem is separated in the projector $\widehat{\Pi}$ from the contribution of the slow subsystem. The condition for uniqueness of the projector is simplified and it depends only on the eigenvalues of the slow subsystem. This is actually expected since in the first equivalent form the k-th moment of the system can be written as (7.4), in which $s_i N - I$ is full-rank for any s_i.

It is clear from the form of $\widehat{\Pi}$ that the information on the fast subsystem is encoded in the projector and it is not lost by the moment matching technique. Thus, the output of reduced order models based on this projector approximates the output of the system only when consistent initial conditions are taken in account, *i.e.* when the solution of the system to be reduced exists. Moreover, this formulation gives high flexibility in maintaining the properties of the fast and slow subsystems. Depending on the particular application the fast subsystem can be preserved, reduced or eliminated independently of the reduction of the slow subsystem. This possibility is analyzed in detail in the next section.

7.4 Reduced order models

In this section several families of reduced order models achieving moment matching are presented. The possibility of obtaining purely fast or purely slow singular systems is investigated and a "simple" family of singular systems is given. In addition, the solution to the problem of matching with impulsive controllability is given.

Definition 23. Consider system (7.1) and let $S \in \mathbb{R}^{\nu\times\nu}$ be any non-derogatory matrix with characteristic polynomial (7.7). Assume $\sigma(S)\cap\sigma(A_1) = \emptyset$ and let L be such that the pair (L, S) is observable. Then the system

$$\begin{aligned} \dot{\xi}_1 &= F_1\xi_1 + G_1u, & M\dot{\xi}_2 &= \xi_2 + G_2u, \\ \psi_1 &= H_1\xi_1, & \psi_2 &= H_2\xi_2, \\ \psi &= \psi_1 + \psi_2, \end{aligned} \tag{7.11}$$

with $\xi_j(t) \in \mathbb{R}^\nu$, for $j = 1, 2$, $\psi_j(t) \in \mathbb{R}$, for $j = 1, 2$, $F_1 \in \mathbb{R}^{\nu\times\nu}$ full-rank, $M \in \mathbb{R}^{\nu\times\nu}$ nilpotent with degree $\hbar$, $G_j \in \mathbb{R}^{\nu\times 1}$, for $j = 1, 2$, and $H_j \in \mathbb{R}^{1\times\nu}$, for $j = 1, 2$, is a *model of system (7.1) at S*, if there exists a unique solution $\begin{bmatrix} P_1^\top & P_2^\top \end{bmatrix}^\top$ of the equations

$$\begin{aligned} F_1P_1 - P_1S &= -G_1L, \\ P_2 &= -\sum_{i=0}^{\hbar-1} M^iG_2LS^i, \end{aligned} \tag{7.12}$$

such that

$$\begin{aligned} C_1\Pi_1 &= H_1P_1, \\ C_2\Pi_2 &= H_2P_2, \end{aligned} \tag{7.13}$$

where $\begin{bmatrix} \Pi_1^\top & \Pi_2^\top \end{bmatrix}^\top$ is the unique solution of (7.10). System (7.11) is a *reduced order model of system (7.1) at S* if $2\nu < n$.

We analyze now some special families of reduced order models which achieve moment matching.

7.4.1 Non-singular reduced order model

Using these results, it is easy to approximate a singular system with a non-singular system. This was the first approach attempted, see [183]. The same authors and following researchers, *e.g.* see [184], deemed the method inaccurate, since the impulsive characteristic of the singular system is lost. However, in particular settings and as a first approximation, it may be useful to approximate a singular system with a non-singular one. This can be

done easily in the present framework. In fact, the system

$$\dot{\xi} = F\xi + Gu, \qquad \psi = H\xi, \tag{7.14}$$

is a *model of system (7.1) at* S, if there exists a unique solution P of the equation

$$FP - PS = -GL, \tag{7.15}$$

such that

$$C\Pi = HP, \tag{7.16}$$

where Π is the unique solution of (7.6). Note that the reduced order model (7.14) has dimension ν, whereas the reduced order model (7.11) has dimension 2ν.

7.4.2 Fast reduced order model

Finally, a completely fast reduced order model which approximates the singular system can be obtained in the present framework. The system

$$M\dot{\xi} = \xi + Gu, \qquad \psi = H\xi, \tag{7.17}$$

is a *model of system (7.1) at* S, if

$$H = -C\Pi \left(\sum_{i=0}^{\hbar-1} M^i GLS^i \right)^{-1}, \tag{7.18}$$

where Π is the unique solution of (7.6), is well-defined. The reduced order model (7.17) has dimension ν.

7.4.3 The identity family of singular reduced order models

As already shown in Chapters 2 and 3, simple families of reduced order models are identified by the choice $P = I$. In the present setting it seems natural to select $P_1 = I$ keeping

M and G_2 free. Then, the system

$$\begin{aligned}
\dot{\xi}_1 &= (S - G_1 L)\xi_1 + G_1 u, \\
M\dot{\xi}_2 &= \xi_2 + G_2 u, \\
\psi_1 &= C_1 \Pi_1 \xi_1, \\
\psi_2 &= C_2 \sum_{i=0}^{h-1} N^i B_2 L S^i \left(\sum_{i=0}^{h-1} M^i G_2 L S^i \right)^{-1} \xi_2, \\
\psi &= \psi_1 + \psi_2,
\end{aligned} \tag{7.19}$$

where $\begin{bmatrix} \Pi_1^\top & \Pi_2^\top \end{bmatrix}^\top$ is the unique solution of (7.10), is a *model of system (7.1) at S* for any G_1, G_2 and M such that (7.12) has a unique solution.

Remark 50. The family of systems (7.19) belongs to the family given in Definition 23. The families (7.14) and (7.17) do not belong to the family given in Definition 23. ■

7.4.4 Matching with impulsive controllability

Impulse controllability is a characteristic property of singular systems and corresponds to the controllability of only the fast subsystem. A condition for impulse controllability is given in [180]. System (7.1) is impulse controllable if and only if

$$\operatorname{rank} \begin{bmatrix} E & 0 & 0 \\ A & E & B \end{bmatrix} = n + \operatorname{rank}(E). \tag{7.20}$$

If the system to be reduced is impulsive controllable, it is desirable to maintain this property also in the reduced order model.

Theorem 24. System (7.11), or (7.17), or (7.19), is impulsive controllable if M and G_2 are selected such that

$$\operatorname{rank} \begin{bmatrix} M & G_2 \end{bmatrix} = \operatorname{rank}(M). \tag{7.21}$$

The proof of the theorem is straightforward, however, to the best of the author's knowledge, the condition for impulsive controllability has not been presented in the simple form given in (7.21).

Proof. Following [180]

$$\operatorname{rank}\begin{bmatrix} E & 0 & 0 \\ A & E & B \end{bmatrix} = 2n_1 + \operatorname{rank}\begin{bmatrix} N & 0 & 0 \\ I & N & B_2 \end{bmatrix}.$$

Note now that

$$n + \operatorname{rank}(E) = n_1 + n_2 + \operatorname{rank}\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}.$$

Then condition (7.20) reduces to

$$\operatorname{rank}\begin{bmatrix} E & 0 & 0 \\ A & E & B \end{bmatrix} = 2n_1 + n_2 + \operatorname{rank}\begin{bmatrix} N & B_2 \end{bmatrix} =$$
$$= n_1 + n_2 + \operatorname{rank}\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} = 2n_1 + n_2 + \operatorname{rank}(N),$$

that written for the reduced order model, yields (7.21). □

Remark 51. The apparent simplicity of (7.21) hides the fact that the condition is a PBH test for controllability of the fast subsystem. In fact, in this case the PBH test is

$$\operatorname{rank}\begin{bmatrix} I - \lambda M & G_2 \end{bmatrix} = \operatorname{rank} M, \qquad \text{for all } \lambda \in \sigma(M).$$

Since the only eigenvalue of M is zero, the PBH test reduces to (7.21). ■

Remark 52. As noted in [188], in the reduction of non-proper $W(s)$, namely systems with a nilpotent degree larger than zero, the algebraic part should be maintained unchanged to guarantee that the error between the system to be reduced and the reduced order model stays bounded. This is obviously implied by the fact that the absolute value of the frequency response grows unbounded as the the frequency increases and the slope of this growth depends on the degree of the algebraic part of the transfer function. Thus, to have a small H_2 error norm the degree of the algebraic part should be maintained unaltered. However, in many cases we are not interested in preserving the rate at which the frequency response grows unbounded, but we are interested only in the maintaining the property that it grows unbounded. Or alternatively, we may be interested only in the output response for low frequencies. In this cases, the method in [188] is not able to reduce the fast substystem. This limit is clearly averted with this approach which leaves to the

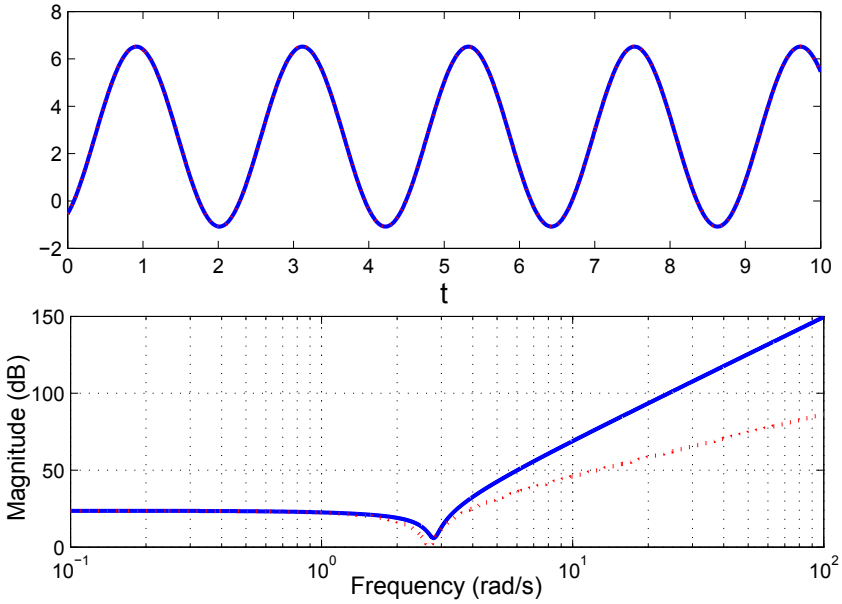

Figure 7.1: (Top) Time history of the output of system (7.22) (solid line) and of the output of the reduced order model (7.23) (dashed line). (Bottom) Magnitude of the Bode plot of system (7.22) (dashed line) and of the reduced order model (7.23) (dotted line).

designer the choice of reducing the fast subsystem, the slow subsystem, or a combination thereof. ■

7.4.5 Example: a classical fast subsystem example

This example has been proposed in [183], see also [184]. Consider the purely fast fifth order singular system described by the equations

$$N\dot{x} = x + Bu, \qquad y = Cx, \tag{7.22}$$

with

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0.1 \\ 0.2 \\ 1.8 \\ 2.5 \\ 3.0 \end{bmatrix}, \; C^\top = \begin{bmatrix} 0.1 \\ 0.3 \\ 1.2 \\ 1.8 \\ 2.8 \end{bmatrix}.$$

Let $\sigma(S) = \{0, \pm 2.85j\}$ and L and $\omega(0)$ in (7.9) be randomly generated. System (7.17) is a third order reduced order model of system (7.22), namely

$$\begin{aligned} M\dot{\xi} &= \xi + Gu, \\ \psi &= C \sum_{i=0}^{h-1} N^i BLS^i \left(\sum_{i=0}^{\hbar-1} M^i GLS^i \xi \right)^{-1} \xi, \end{aligned} \tag{7.23}$$

where G is randomly generated and

$$M = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Fig. 7.1 (top) shows the time history of the output of system (7.22) (solid line) and of the output of the reduced order model (7.23) (dashed line). The two responses overlap because a purely fast system does not have a transient response. Fig. 7.1 (bottom) shows the magnitude of the Bode plot of system (7.22) (dashed line) and of the reduced order model (7.23) (dotted line). This plot gives a more indicative information of the quality of the approximation. In fact, from the figure we see that the reduced order model is a good approximation of the system at low frequencies ($< \approx 3$ rad/s).

Remark 53. As anticipated, the method proposed in [188] cannot be used to reduce system (7.22). Conversely with the method proposed herein a reduced order model is computed and the output response of system (7.22) is perfectly matched. Note also that increasing the number of interpolation points would reduce the discrepancy shown in the bottom graph of Fig. 7.1. When the dimension of the reduced order model reaches the dimension of the system, the error becomes zero. ■

7.4.6 Example: a large-scale singular system

Consider the singular system (7.2) with $n_1 = 60$ and $n_2 = 20$. Let $A_1 = \text{diag}(-0.1, -0.2, \ldots, -6.0)$, N be a canonical nilpotent matrix of degree $h = 20$ and B_1, B_2, C_1, C_2 and L be randomly generated. The matrix S is such that $\sigma(S) = \{0, \pm 0.1j, \pm 0.5j, \pm j, \pm 1.13j, \pm 5j\}$. The reduced order model (7.19) is computed. The matrix G_1 is selected such that $\sigma(S - G_1 L) \subset \mathbb{C}_{<0}$, M is a canonical nilpotent matrix of

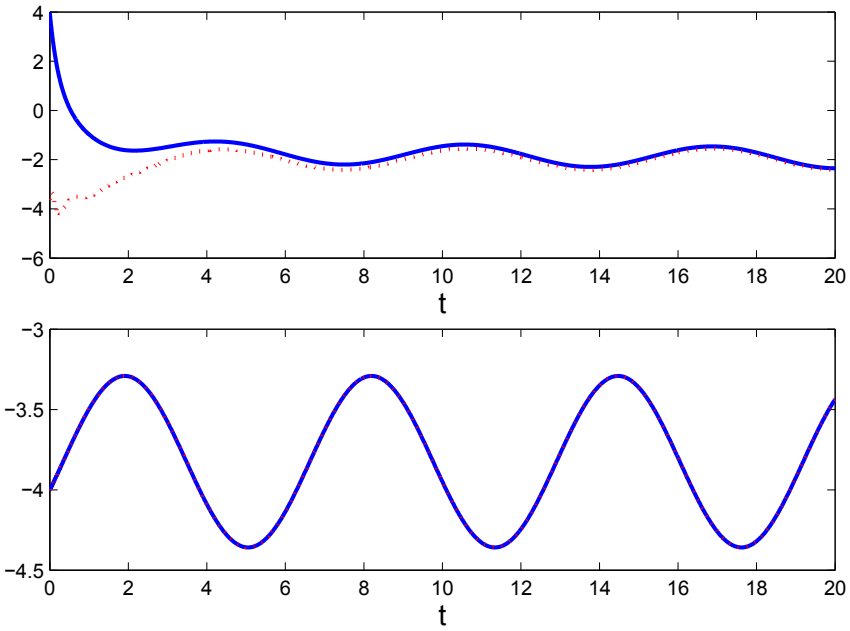

Figure 7.2: (Top) Time history of the output of the slow subsystem (solid line) and the output of the slow subsystem of the reduced order model (dotted line). (Bottom) Time history of the output of the fast subsystem (solid line) and the output of the fast subsystem of the reduced order model (dotted line).

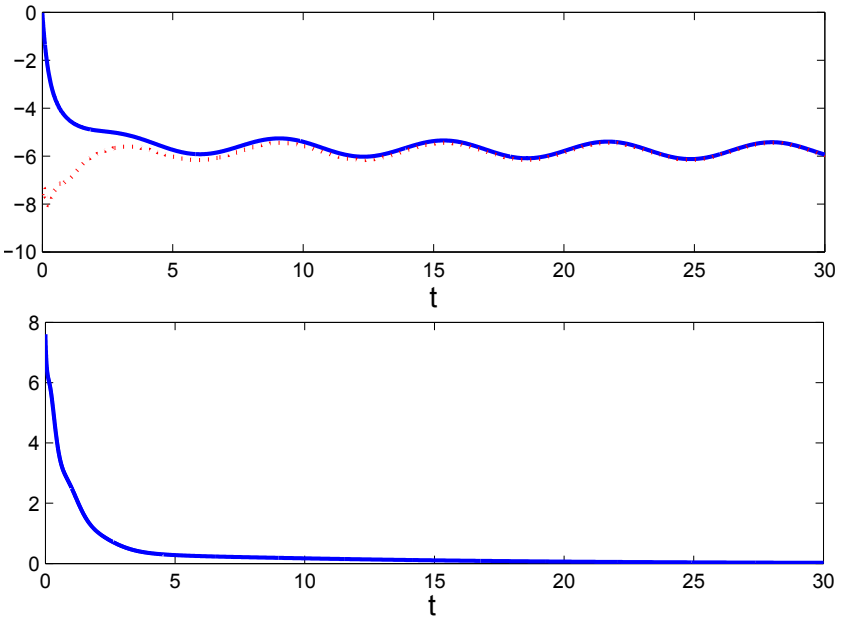

Figure 7.3: (Top) Time history of the output of the system (solid line) and of the reduced order model (dotted line). (Bottom) Absolute error between the two output responses.

degree $\hbar = 10$ and G_2 is selected as a truncation of B_2. Fig. 7.2 (top) shows the time history of the output of the slow subsystem (solid line) and the output of the slow subsystem of the reduced order model (dotted line). Fig. 7.2 (bottom) shows the time history of the output of the fast subsystem (solid line) and the output of the fast subsystem of the reduced order model (dotted line). As expected the bottom plots overlap. Fig. 7.3 (top) shows the time history of the output y of the system (solid line) and of the total output ψ of the reduced order model (dotted line). Fig. 7.3 (bottom) shows the absolute error between the two output responses. As expected the error between the two responses depends only on the response of the slow subsystem and goes to zero as the transient vanishes.

7.5 Conclusion

The model reduction technique based on moment matching has been extended to linear singular systems. Combining the interpolation-based and the steady-state-based description of moment a partitioned projector is constructed. The contribution of the slow subsystem and the contribution of the fast subsystem to the moment are separated. The information on the fast subsystem is encoded in the projector and it is not lost by the moment matching technique. Moreover, the output of reduced order models based on this projector approximates the output of the system only when consistent initial conditions are taken in account. Exploiting this partitioned projector, several families of reduced order models have been proposed. In particular, a purely fast, a purely slow and a "simple" family of reduced order models have been proposed. The possibility of maintaining the impulsive controllability property has been investigated and a few examples have been used to illustrate the results. A natural extension of the current results is represented by the model reduction of nonlinear singular systems and by the model reduction of systems with constraints.

Part II

Analysis

Chapter 8

Discontinuous phasor transform and its application to the steady-state analysis in power electronics

8.1 Introduction

Power converters are inherently discontinuous[1] systems: power electronics is described as *a branch of electrical and electronic engineering concerned with the analysis, simulation, design, manufacture, and application of switching-mode power converters* [189], [190], [191]. Since the transient response of electric circuits decays rapidly to zero, the steady-state analysis of power electronic devices is of paramount importance for their design [192], [193], [194], [195]. However, a trade-off between the level of the approximation and the difficulty of the analysis has usually to be made. Among the techniques that have been developed for this analytic task, the phasor transform represents a powerful and flexible mathematical tool which has been used for the study of the steady-state behavior of circuits powered by sinusoidal sources [196], [197]. The phasor transform greatly simplifies the dynamic analysis because it changes integro-differential equations in algebraic equations, which are computationally and analytically more easily solvable. In recent years

[1]With some abuse of terminology, the word "discontinuous" is used in this chapter for *e.g.* sources, generators, components, to mean that the generated signal is discontinuous.

several generalizations of the phasor transform have been proposed specifically for the analysis of power electronic devices [198], [199], [200], [201]. These methods are based on the assumption that the phasor is a function of time and that the input signal can be described by a complex exponential. Thus, it is assumed that the input of the circuit is a smooth signal. Periodic discontinuous signals can be analyzed using the Fourier series, obtaining a steady-state description of the response as the sum of the phasors at infinitely many frequencies. However, practically, only an approximation can be obtained and, moreover, high harmonics are usually ignored because they are numerically difficult to compute [198]. In addition, we note that these papers do not explicitly define the time-varying phasor itself. Only the inverse phasor transform is given and the phasor is found comparing two transformed quantities. This forces the designer to recognize and transform known subcircuits. For instance in [200] nine categories of subcircuits have been identified. One of those, which is not transformed, *entails thousands of different configurations.*

This chapter originates from the observation that a phasor is a very special case of moment. In fact, the phasors of an electric circuit are the components of the unique solution of the Sylvester equation (2.8) associated to the system representing the circuit. The Sylvester equation itself is proved to be the phasor transformed system describing the electric circuit. Exploiting this equivalence, the instantaneous power and the average power are defined utilizing the moments. The interest in revisiting the notion of phasor with the moment theory is twofold.

1. A direct consequence is that very efficient algorithms to compute the moments of a linear system can be used to compute the phasors of large-scale systems.

2. The most relevant reason that justifies the interest in this equivalence is that we have generalized the description of moment beyond linear systems, in particular to systems driven by discontinuous signals.

Thus, exploiting the equivalence discussed, and exploiting the results given in Chapter 6, a new phasor transform is given. The new phasor transform has the following advantages:

1. differently from other analysis methods, like the state-space averaging, the new phasor transform describes the steady-state response of an electric circuit powered by

any periodic discontinuous signal in closed-form, *i.e.* without approximations;

2. the capability of describing the steady-state response of any electrical circuit without approximations applies also to the study of the behaviour of non-ideal switches;

3. differently from [198–200] the definition of the phasor itself is given; this has the consequence that it is not necessary to recognize subcircuits (although it is still possible to proceed in this way): the phasor can be computed directly for any linear circuits of whatever complexity;

4. the formula which we provide defines the phasors of all the currents and of the integrals of all the currents in the circuit; as a consequence the steady-state v-i characteristics of all the electric quantities of the circuit can be reconstructed.

To illustrate the utilization of the results of the chapter, the steady-state behavior of a Class D CLL resonant inverter and of a wireless power transfer system with non-ideal switches (subject to the reverse recovery effect) is analyzed. We show that the new phasor transform is particularly useful when there is little damping in the circuit (scenario which "preserves" the discontinuous behavior that approximating series solutions cannot fully capture).

The rest of the chapter is organized as follows. This section continues with a precise formulation of our problem, aim and the description of the circuits which are analyzed in the chapter. In Section 8.2 the equivalence between moments and phasors is proved. The use of the results are illustrated by means of an example and the definition of power is given in terms of moments. In Section 8.3 the definition of phasor and of the inverse phasor transform are generalized to linear circuits powered by discontinuous periodic signals. The phasors of resistors, inductors and capacitors are given and the characterization of the power is described in the new phasor domain. Extensive simulations show the capability of the new phasor definition in describing the steady-state response of power electronic devices. Section 8.5 contains some concluding remarks.

All the results of this chapter are original contributions developed in fulfillment of my PhD course and they have been published in the conference paper [19] and in the journal paper [20].

8.1.1 Problem and aim

All currents and voltages in linear circuits are described by linear differential equations of the form

$$a_n \frac{d^n}{dt^n} f + a_{n-1} \frac{d^{n-1}}{dt^{n-1}} f + \cdots + a_0 f = u, \tag{8.1}$$

where $f : \mathbb{R} \to \mathbb{R}$ represents a current or voltage, $u(t) \in \mathbb{R}$ is a current or voltage input and $a_i \in \mathbb{R}$, with $i = 0, \ldots, n$. Without loss of generality we assume that $a_n \neq 0$. In the analysis of circuits is of interest to study the steady-state response of the system[2]. If the input u is a sinusoidal signal of amplitude a_u, angular frequency ω and phase ϕ, then a classical tool for the steady-state analysis of (8.1) is the phasor of $f(t)$. This is usually introduced by means of the inverse phasor transform which is defined in his simplest form as

$$f(t) = \Re \left[\overline{F} e^{j\omega t} \right], \tag{8.2}$$

where[3] $\overline{F} : \mathbb{C} \to \mathbb{C}$ is called the phasor of $f(t)$. Several generalizations of this transform have been proposed. The most general form presented in the literature is the so-called power-invariant (inverse) phasor transform [200], [199] described by

$$f(t) = \Re \left[\overline{F}(t) e^{j\omega(t)} \right], \tag{8.3}$$

where $\omega(t)$ is non-constant and the phasor $\overline{F}(t)$ is a function[4] of time and ω. From (8.3) we note that the analysis of a circuit with the phasor transform is based on the assumption that the input is a complex exponential with frequency $\omega(t)$, *i.e.* $u(t) = a_u e^{j(\omega(t)+\phi)}$.
In this chapter we revisit and extend the notion of phasor. In the first part of the chapter we show that the phasors of an electric circuit are the moments at $j\omega$ of the linear system describing the circuit. In the second part of the chapter we eliminate the assumption that the source has to be described by a complex exponential extending the notion of phasor, as defined in (8.3), to a general class of input signals including discontinuous signals, such as square waves and triangular waves.

[2]Note that this means to find the particular solution of (8.1).
[3]Phasors are frequency dependent. We omit the argument ω from the phasor $\overline{F}(\omega)$ for ease of notation.
[4]We again omit the argument ω from the phasor $\overline{F}(t, \omega)$ for ease of notation.

For the sake of clarity, we now define what we mean with phasor transform of a linear system.

Definition 24. Consider a linear, single-input, single-output, continuous-time, system described by the equations

$$\dot{x} = Ax + Bu, \qquad y = Hx, \tag{8.4}$$

with[5] $x(t) \in \mathbb{C}^n$, $u(t) \in \mathbb{C}$, $y(t) \in \mathbb{C}$, $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times 1}$ and $H \in \mathbb{C}^{1\times n}$. The *phasor transform of the linear system (8.4)* for the source $u(t) = a_u e^{j(\omega t+\phi)}$, with $a_u \in \mathbb{R}$, $\omega \in \mathbb{R}$ and $\phi \in \mathbb{R}$, is

$$\overline{X} j\omega e^{j\omega t} = A\overline{X}e^{j\omega t} + Ba_u e^{j\phi}e^{j\omega t}, \qquad \overline{Y}e^{j\omega t} = H\overline{X}e^{j\omega t}. \tag{8.5}$$

In what follows we only assume that the components in the circuit are ideal. We do not make any other simplifying assumption since we aim to describe exactly the behavior of the circuit. In Section 8.4.2 we relax the hypothesis that the components are ideal considering non-ideal switches.

Remark 54. The theory is developed applying the Kirchhoff's Voltage Law. We assume that the sources are voltage sources and the state variables and the output are currents. As a consequence we show that the moments are the phasors of the currents. An equivalent analysis based on the Kirchhoff's Current Law can be derived. ■

Remark 55. For the sake of clarity we consider the single phase case. The results can extend to the multiphase framework either following the approach in [200], or the multi-input approach presented in Chapter 5 (see also [63]). ■

8.1.2 Case studies

To illustrate the results of the chapter we study the steady-state response of resonant circuits. The interest in using this typology of circuits lies upon the fact that these are at the basis of many power converters which are normally driven by square waves. For instance, Fig. 8.1 shows a Class D CLL resonant inverter. The circuit is powered by a DC voltage V_i. The voltage across the capacitor C, which is the output of the subcircuit

[5]Usually the state $x(t)$ of a dynamical system represents real quantities. However, herein we use the complex domain because the quantities involved in the phasor analysis are complex valued.

constituted by the two transistors, is a square wave. Thus, an equivalent circuit can be derived as shown in Fig. 8.2 (see [189] for the derivation and the definition of the equivalent quantities). We note that the analysis of the Class D CLL resonant inverter (and of many other Class D inverters) boils down to the analysis of a RLC circuit driven by a square wave. The techniques presented in the second part of the chapter are developed for the study of this type of circuits. Moreover, RLC circuits arise also in other power electronic applications. For instance, Fig. 8.3 shows a wireless power transfer system [202, 203] and Fig. 8.4 shows its equivalent circuit which consists of two coupled RLC circuits.

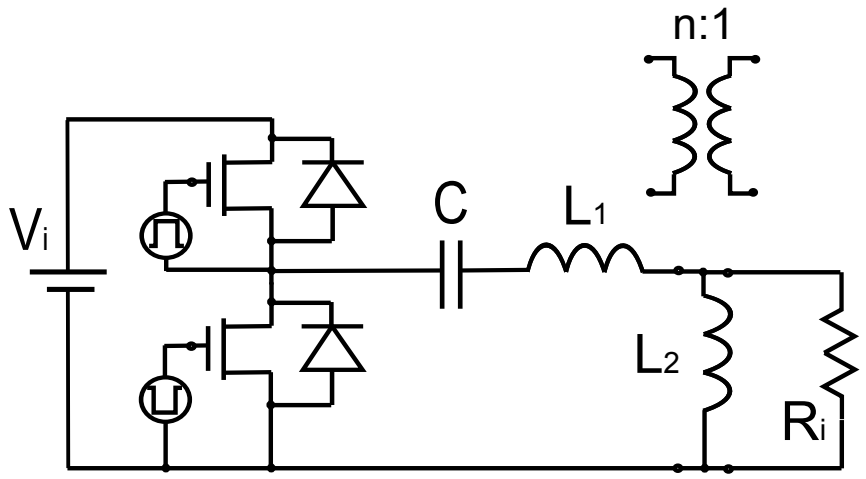

Figure 8.1: Class D CLL resonant inverter.

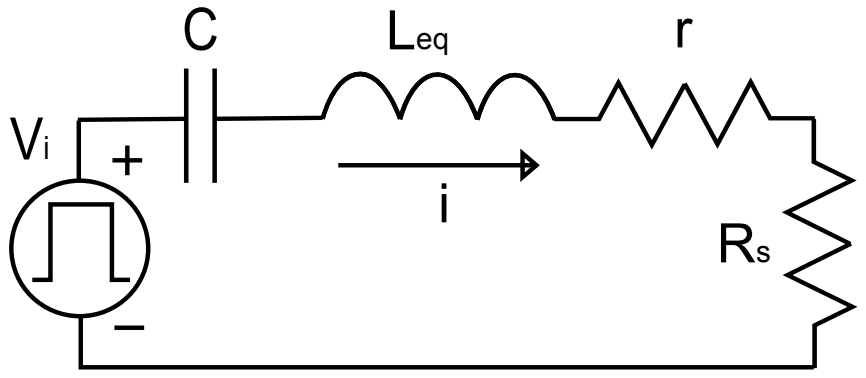

Figure 8.2: Equivalent circuit of a Class D CLL resonant inverter as derived in [189].

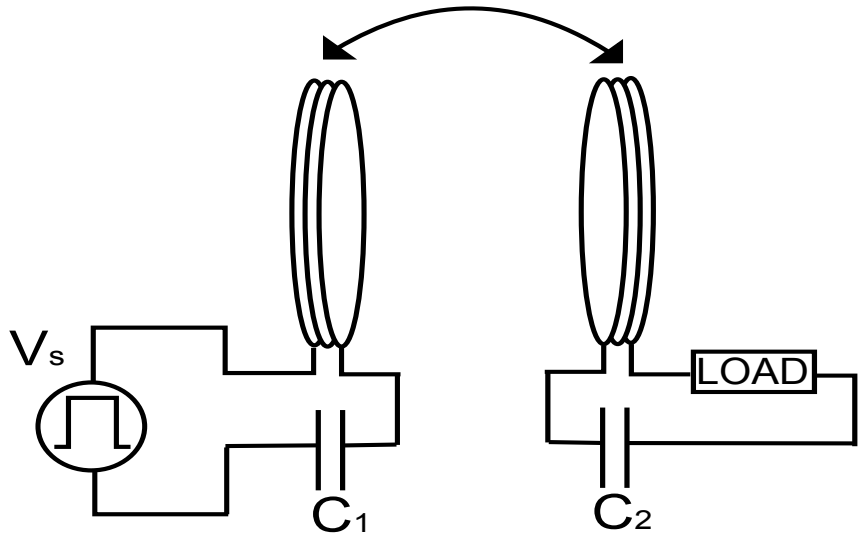

Figure 8.3: Wireless power transfer system.

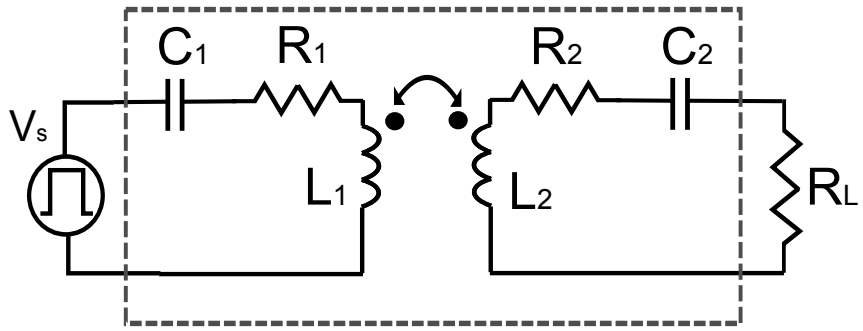

Figure 8.4: Equivalent circuit of the wireless power transfer system in Fig. 8.3.

The two examples are used as follows. In Section 8.2 we use the circuit of Fig. 8.4 because it offers more mathematical complexity and thus it is more helpful for showing some important details. In Section 8.3 we first use the circuit of Fig. 8.2 illustrating the technique for different Q factors[6]. We show that as the Q factor decreases, *i.e.* the damping increases, the difference in quality of the approximation between the Fourier series and our solution grows. Finally, we analyze the circuit of Fig. 8.4 driven by non-ideal switches. We show that the technique presented in the chapter can describe, without approximations, the reverse recovery effect which is present in the majority of commercial switches [204].

[6]The Q factor, or quality factor, is a parameter that describes the damping of an oscillator or resonator. For an ideal series RLC circuit, we use the definition Q= $\frac{1}{R}\sqrt{\frac{L}{C}}$.

8.2 Electrical equivalent of the moment theory

In this section we show that the theory of moments developed to solve the model reduction problem has an electric equivalent. In particular we prove that the phasors of an electric circuit, as defined in (8.2), are the moments of the system describing the circuit when a single complex interpolation point is selected. Then, exploiting the equivalence between moments and phasors, we define the power using the moments. Revisiting these results is instrumental for the rest of the chapter when the notion of phasor is extended to non-conventional sources: knowing the relation between moments and power is essential to give a physical meaning to the newly defined quantities. For the sake of convenience we report here the Sylvester equation (2.8), namely

$$A\Pi + B\Gamma = \Pi\Sigma, \tag{8.6}$$

and the equation of the signal generator (2.10), namely

$$\dot{\zeta} = \Sigma\zeta, \qquad u = \Gamma\zeta, \tag{8.7}$$

which will be used through the chapter. Note that for notational needs the original quantities L and S have been replaced by Γ and Σ, respectively. Finally, to streamline the presentation, we introduce the following definition.

Definition 25. The system (8.4) and the generator (8.7) are said to be in the *real convention* if the matrices A, B, H, Γ and Σ have real entries. They are said to be in the *mixed convention* if the matrices A, B and H have real entries and the matrices Γ and Σ have complex entries. They are said to be in the *complex convention* if the matrices A, B, H, Γ and Σ have complex entries.

Note that in the real convention and in the mixed convention, for all integers k with $2 \leq 2k \leq n$, the component x_{2k} of x is a current i_k, whereas the component x_{2k-1} of x is the integral $\int_{t_0}^{t} i_k(\tau)d\tau$.

8.2.1 Equivalence between moments and phasors

Herein we show that writing the phasor transform of a linear electric circuit is equivalent to writing the associated Sylvester equation. Moreover, the components of the solution of this Sylvester equation are the phasors of all the currents (and of the integrals of the currents) in the circuit.

Proposition 10. Consider the source $u(t) = a_u e^{j(\omega t+\phi)}$, with $a_u \in \mathbb{R}$, $\omega \in \mathbb{R}$ and $\phi \in \mathbb{R}$, and assume $j\omega \notin \sigma(A)$. The phasor transform of system (8.4) (see Definition 24) written in the mixed convention coincides with the Sylvester equation (8.6) with $\Sigma = j\omega$ and $\Gamma = a_u e^{j\phi}$. The components of Π, which is the unique solution of equation (8.6), are the phasors of the currents and of the integrals of the currents in the circuit.

Proof. We first compute the phasor transform of system (8.4) for the source $u = a_u e^{j(\omega t+\phi)}$, namely

$$\overline{X} j\omega e^{j\omega t} = A\overline{X} e^{j\omega t} + B a_u e^{j\phi} e^{j\omega t}, \qquad \overline{Y} e^{j\omega t} = H\overline{X} e^{j\omega t},$$

by Definition 24. Since $e^{j\omega t} \neq 0$ for all $t \in \mathbb{R}$, it can be canceled out yielding

$$\overline{X} j\omega = A\overline{X} + B a_u e^{j\phi}, \qquad \overline{Y} = H\overline{X}.$$

Thus the phasors of all the currents (and their integrals) in the circuit and the phasor of the output current are given by

$$\overline{X} = (j\omega I - A)^{-1} B a_u e^{j\phi}, \qquad \overline{Y} = H(j\omega I - A)^{-1} B a_u e^{j\phi},$$

respectively. Consider now the signal generator (8.7) with

$$\Sigma = j\omega, \quad \Gamma = a_u e^{j\phi}.$$

The associated Sylvester equation (8.6) is

$$A\Pi + B a_u e^{j\phi} = \Pi j\omega,$$

which, if $j\omega \notin \sigma(A)$, has the unique solution

$$\Pi = (j\omega I - A)^{-1} B a_u e^{j\phi},$$

which concludes the proof. □

Corollary 5. The phasor of the output response y of system (8.4) is the moment of the system at $j\omega$, namely $\overline{Y} = H\Pi$. The *inverse phasor transform* of the output current y of system (8.4) is

$$y(t) = \Re\left[H\Pi e^{\Sigma t}\right].$$

Proof. The first claim follows noting that the phasor of the output response of system (8.4) is given by

$$\overline{Y} = H\Pi = H(j\omega I - A)^{-1}Ba_u e^{\phi}.$$

To prove the second claim we note that

$$\Re\left[H\Pi e^{\Sigma t}\right] = \Re\left[\overline{Y}e^{j\omega t}\right],$$

which is the inverse phasor transform of the output response of system (8.4). □

Remark 56. Phasors are a very special case of moments of linear systems. In fact, they are the moments at the single complex interpolation point $\Sigma = j\omega$. ■

Remark 57. The higher order derivatives $\dfrac{d^n x}{dt^n}$ and the integral $\displaystyle\int_{t_0}^{t} x(\tau)d\tau$ are transformed in the phasor domain into $(j\omega)^n\overline{X}e^{j\omega}$ and $\dfrac{1}{j\omega}\overline{X}e^{j\omega}$, respectively. Similarly they are transformed in the "moment domain" into $\Sigma^n\Pi e^{j\omega}$ and $\Sigma^{-1}\Pi e^{j\omega}$, respectively. ■

Remark 58. The relation between moments and phasors established with the Sylvester equation (8.6) gives the advantage of computing the phasors using Arnoldi and Lanczos algorithms [5, Section 14.1] or the techniques presented in Chapter 4. These techniques offers very efficient methods to determine the moments, and thus the phasors, of very large-scale systems, in which the matrix inversion may be computationally expensive. ■

Remark 59. The main advantage of the equivalence that we have established is that this lends itself to the extension of the phasor transform beyond the linear framework (and the discontinuous framework of this chapter) to circuits with nonlinear elements and circuits with delays (exploiting the results in Part I of this Thesis). ■

8.2.2 Example: moments of a wireless power transfer system

We present now a worked out example which shows how to apply these results and highlights a few important aspects regarding the "convention" used in the approach.

Fig. 8.4 illustrates a wireless power transfer system [203] consisting of two coils. In this example we assume that an AC sinusoidal voltage source with an amplitude of V_s and an angular frequency of ω is applied to the transmitter coil on the input side. A load resistor R_L is connected to the receiving coil on the output side. By applying the Kirchhoff's Voltage Law to the two coils, we obtain the system of equations

$$\begin{aligned} R_1 i_1 + L_1 \frac{di_1}{dt} + \frac{1}{C_1}\int i_1 dt + M_{12}\frac{di_2}{dt} &= u(t), \\ M_{21}\frac{di_1}{dt} + R_{2L} i_2 + L_2 \frac{di_2}{dt} + \frac{1}{C_2}\int i_2 dt &= 0, \end{aligned} \tag{8.8}$$

where i_1 and i_2 are the currents flowing in the coils 1 and 2, R_1 and R_2 are the resistances, $R_{2L} = R_2 + R_L$, L_1 and L_2 are the self-inductances, C_1 and C_2 are the capacitances and $M_{12} = M_{21}$ are the mutual inductances between the two coils. We are interested in determining the amplitude and phase of the steady-state current in the receiving coil, *i.e.* the phasor $\overline{I}_2$.

We start by solving the problem with the phasor transform approach. Transforming the differential equations (8.8) we obtain the complex algebraic system

$$Z_1\overline{I}_1 + j\omega M_{12}\overline{I}_2 = V_s, \qquad j\omega M_{21}\overline{I}_1 + Z_2\overline{I}_2 = 0,$$

where $Z_1 = R_1 + j\omega L_1 - j\dfrac{1}{\omega C_1}$ and $Z_2 = R_{2L} + j\omega L_2 - j\dfrac{1}{\omega C_2}$. Solving with respect to $\overline{I}_2$ yields

$$\overline{I}_2 = \frac{-j\omega M_{21}}{Z_1 Z_2 + \omega^2 M_{21} M_{12}} V_s.$$

Now we compute the moment of the differential system (8.8) at

$$\dot{\zeta} = \Sigma\zeta, \quad u = \Gamma\zeta. \tag{8.9}$$

We begin using the "mixed convention" that, as highlighted later, is the most useful among the three representations. Thus, consider $\Gamma = V_s$ and $\Sigma = j\omega$ and the state

$$x_1(t) = \int_{t_0}^{t} i_1(\tau)\, d\tau, \quad x_2(t) = i_1(t), \quad x_3(t) = \int_{t_0}^{t} i_2(\tau)\, d\tau, \quad x_4(t) = i_2(t).$$

System (8.8) can be represented by the first order system of differential equations (8.4) with

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{L_2}{C_1\hat{L}} & -\frac{R_1 L_2}{\hat{L}} & \frac{M_{12}}{C_2\hat{L}} & \frac{M_{12}R_{2L}}{\hat{L}} \\ 0 & 0 & 0 & 1 \\ \frac{M_{21}}{C_1\hat{L}} & \frac{R_1 M_{21}}{\hat{L}} & -\frac{L_1}{C_2\hat{L}} & -\frac{L_1 R_{2L}}{\hat{L}} \end{bmatrix}, \quad \begin{aligned} B &= \begin{bmatrix} 0 & \frac{L_2}{\hat{L}} & 0 & -\frac{M_{21}}{\hat{L}} \end{bmatrix}^\top, \\ H &= \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}, \end{aligned} \tag{8.10}$$

where $\hat{L} = L_1 L_2 - M_{12} M_{21} \neq 0$. The solution of the Sylvester equation (8.6) is given by

$$\Pi = (j\omega - A)^{-1} B V_s,$$

and the moment of the system at (Γ, Σ) is given by

$$H\Pi = \epsilon_4 \Pi = \frac{j\omega^3 M_{21}}{D\hat{L}} V_s,$$

with

$$D = -\frac{\omega^2}{\hat{L}} \left(Z_1 Z_2 + \omega^2 M_{12} M_{21} \right)$$

the determinant of the matrix $(j\omega I - A)$. In a similar way we can prove that

$$\begin{bmatrix} \Pi_1 & \Pi_2 & \Pi_3 & \Pi_4 \end{bmatrix}^\top = \begin{bmatrix} \frac{1}{j\omega}\overline{I}_1 & \overline{I}_1 & \frac{1}{j\omega}\overline{I}_2 & \overline{I}_2 \end{bmatrix}^\top.$$

It is interesting to explore which relation between phasors and moments holds when the "real convention" is used. Consider system (8.4) with the matrices given in (8.10) and the matrices of the signal generator (8.9) given by

$$\Gamma = \begin{bmatrix} V_s & 0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix}.$$

The input of the system is $u = V_s \cos(\omega t)$ instead of $u = V_s \cos(\omega t) + j V_s \sin(\omega t)$. It comes with no surprises that the phasors are related to the solution of the Sylvester equation by

the relations

$$\begin{bmatrix} \Pi_{11} + j\Pi_{12} \\ \Pi_{21} + j\Pi_{22} \\ \Pi_{31} + j\Pi_{32} \\ \Pi_{41} + j\Pi_{42} \end{bmatrix} = \begin{bmatrix} \frac{1}{j\omega}\overline{I}_1 \\ \overline{I}_1 \\ \frac{1}{j\omega}\overline{I}_2 \\ \overline{I}_2 \end{bmatrix}.$$

Finally, we investigate the use of the "complex convention". Consider the coordinates $x_1(t) = i_1(t)$, $x_2(t) = i_2(t)$ and the signal generator (8.9) with $\Gamma = V_s$ and $\Sigma = j\omega$. System (8.8) can be represented by a system of integro-differential equations given by

$$\begin{aligned} \dot{x}_1 &= \frac{1}{\hat{L}}\left(-R_1L_2x_1 - \frac{L_2}{C_1}\int x_1dt + M_{12}R_{2L}x_2 + \frac{M_{12}}{C_2}\int x_2dt + L_2u\right), \\ \dot{x}_2 &= \frac{1}{\hat{L}}\left(-R_{2L}L_1x_2 - \frac{L_1}{C_2}\int x_2dt + M_{21}R_1x_1 + \frac{M_{21}}{C_1}\int x_1dt - M_{21}u\right). \end{aligned}$$

Exploiting Remark 57, we can write the Sylvester equation and find its solution, namely

$$\begin{bmatrix} \Pi_1 \\ \\ \Pi_2 \end{bmatrix} = \begin{bmatrix} j\omega\hat{L} + R_1L_2 + \frac{L_2}{j\omega C_1} & -M_{12}R_{2L} - \frac{M_{12}}{j\omega C_2} \\ -M_{21}R_1 - \frac{M_{21}}{j\omega C_1} & j\omega\hat{L} + L_1R_{2L} + \frac{L_1}{j\omega C_2} \end{bmatrix}^{-1} \begin{bmatrix} L_2V_s \\ \\ -M_{21}V_s \end{bmatrix}$$

showing easily that

$$\begin{bmatrix} \Pi_1 & \Pi_2 \end{bmatrix}^\top = \begin{bmatrix} \overline{I}_1 & \overline{I}_2 \end{bmatrix}^\top.$$

Remark 60. When the system is linear and the source generates a sinusoidal signal, the real convention and the mixed convention are redundant since each current appears twice in the matrix Π (as $\overline{I}_k$ and $\frac{1}{j\omega}\overline{I}_k$). However, when the input is not a complex exponential, the steady-state of the current i_k and of its time integral are not anymore linked by a simple scaling factor. ■

8.2.3 Definition of power from the moments

The phasor analysis is useful to determine the instantaneous and average power absorbed by a load Z at steady-state. Exploiting the relation between phasors and moments we can define these two quantities with respects to the moments. The instantaneous power is defined as

$$p(t) = v(t)i(t) = \Re\left[\overline{V}e^{j\omega t}\right]\Re\left[\overline{I}e^{j\omega t}\right]. \tag{8.11}$$

Exploiting the properties of the real part operator we write the instantaneous power as

$$p(t) = \frac{1}{2}\Re\left[\overline{V}^*\overline{I}e^{j\omega t}e^{-j\omega t}\right] + \frac{1}{2}\Re\left[\overline{VI}e^{j2\omega t}\right]. \tag{8.12}$$

Using Euler's formula it can be proved that the average power

$$P = <p(t)> = \frac{1}{2}\Re\left[\overline{V}^*\overline{I}\right] \tag{8.13}$$

is equal to the first term of (8.12), and that the second term of (8.12) has zero average. Since we have proved that $\overline{I} = H\Pi$, the instantaneous power is described by

$$p(t) = \frac{1}{2}\Re\left[\frac{(H\Pi)^*}{Z}H\Pi\right] + \frac{1}{2}\Re\left[\frac{H\Pi}{Z}H\Pi e^{j2\omega t}\right]. \tag{8.14}$$

with Z the complex impedance.

Remark 61. The equivalence between the average power P and $\frac{1}{2}\Re\left[\overline{V}^*\overline{I}\right]$ is not a definition. The relation is a consequence of the properties of the complex exponential, as highlighted in (8.12), and we may expect that this relation does not hold if the input is not a complex exponential. ■

8.3 Generalizing the phasor to sources in explicit form

Now that we have linked the theory of phasors with the moment theory, we are able to extend the phasor analysis to more general classes of sources. In this section instead of considering sinusoidal sources, we study any periodic source which has the following explicit representation

$$\zeta(t) = \Lambda(t)\zeta(0), \quad u = \Gamma\zeta. \tag{8.15}$$

with $\Lambda(t)$ such that $\Lambda(t) = \Lambda(t-T)$ for $t \geq T$. As remarked in Chapter 6, this class of signal generators includes possibly discontinuous signals, such as square waves and triangular waves, which are of great interest in circuit analysis. In this chapter we develop an electric equivalent of the results presented in Chapter 6, yielding an extension of the phasor transform to this general class of signals. We begin by giving the definition of the discontinuous phasor. We show that multiplying the phasor with the source signal

and taking the real part yields the steady-state current. Then we show the relation between voltage and current for resistors, inductors and capacitors. Finally, we show that the phasor can be used to compute the steady-state instantaneous and average power. In summary, we extend all the features and properties of the phasor analysis to discontinuous sources.

Remark 62. Since (8.15) is a periodic signal, a classical phasor analysis can be carried out exploiting the Fourier series of the signal. However, the approach that we propose achieves a closed-form expression of the phasors and of the steady-state quantities. No approximations are introduced and non-ideal components can be considered. A comparison between the new results and the approximation given by the Fourier series is illustrated in the examples. ■

Remark 63. The state-space averaging method is commonly used as a modeling technique for power electronic converters [191]. However, while the averaging method provides an approximation, the method we propose is exact. In detail, the state-space averaging technique provides a reliable approximation of the steady-state behavior of the circuit when the natural frequencies of the converter (the poles of its transfer function) are much smaller than the switching frequency. The averaging method fails to give a good approximation when the natural frequencies of the system are close or above one half of the switching frequency [205]. On the contrary, the discontinuous phasor transform does not have this restriction since it is not an approximation but an exact description of the steady-state behavior of any linear circuit powered by any discontinuous source. ■

Remark 64. We focus on periodic signals only because these are the most common sources in power electronics. Note that the analysis of this section can be extended to any, possibly non-periodic, signal generators described by an equation of the form $\zeta(t) = \Lambda(t)\zeta(0)$, as discussed in Chapter 6. ■

Remark 65. In the remaining of the section we use the mixed convention. A similar analysis for the real convention can be derived. The complex convention presents more difficulties because Corollary 6 does not have an equivalent in this convention. ■

8.3.1 Definition of the discontinuous phasor

We define now the phasor in the case in which the signal generator is described by (8.15).

Definition 26. Consider system (8.4) and the signal generator (8.15). Assume Assumptions 14, 15, 16 and 17 hold, $\sigma(A) \subset \mathbb{C}_{<0}$ and $\Lambda(t)$ is almost everywhere differentiable. The components of the function

$$\Pi_\infty(t) = (I - e^{AT})^{-1}\left[\int_{t-T}^{t} e^{A(t-\tau)} B\Gamma\Lambda(\tau)d\tau\right]\Lambda(t)^{-1}; \tag{8.16}$$

are the *discontinuous phasors* of all the currents and of all the integrals of the currents in system (8.4) for the source $\Lambda(t)$. The *discontinuous inverse phasor transform* of the steady-state output current $i(t)$ of system (8.4) is

$$i(t) = \Re\left[\overline{I}(t)\Lambda(t)\right], \tag{8.17}$$

with $\overline{I}(t) = H\Pi_\infty(t)$.

Remark 66. We remind that it is always possible to describe square waves, triangular waves and any periodic signal which is almost everywhere differentiable in the mixed convention to satisfy Assumptions 14, 15, 16 and 17. ■

Remark 67. Like in the sinusoidal case, the instantaneous currents are recovered multiplying the phasor with the source and taking the real part. ■

Remark 68. Differently from the sinusoidal case, the phasor $\overline{I}(t)$ is a time-dependent periodic function. Note that if $\Lambda(t)$ is sinusoidal, equation (8.16) defines the usual constant phasor and Π_∞ solves the Sylvester equation (8.6) (see Theorem 20). ■

Remark 69. The inverse phasor transform (8.3) introduced in [198–200] is a particular case of the more general phasor transform we have introduced. In fact, (8.3) is recovered when $\Lambda(t) = e^{j\omega(t)}$. Note however, that the new result have a remarkable advantage. In [198–200] the inverse phasor transform is introduced but the phasor itself (*i.e.* the direct phasor transform) is not defined. Thus to apply the results of those papers one must recognize simple subcircuits and find the v-i characteristics in the phasor domain of the subcircuit. Although this is possible also with the new transform (8.17) (and it is done in the following), it is not necessary. In fact, since we have the definition of the phasor (8.16) from the system matrices A and B, we can obtain the phasors of all the currents and voltages in the circuit in closed-form without the need of decomposing the circuit in pre-classified subcircuits. ■

Now that we have defined the discontinuous phasor and the discontinuous inverse phasor transform we extend the properties of the phasor circuit analysis.

8.3.2 Inductance, capacitance and resistance

Following [198–200] we describe the v-i characteristics of some common subcircuits which constitute power electronic devices. As already remarked, this is not strictly necessary if we want to compute simply the phasors of the overall circuit. However, describing simple circuits in the phasor domain improves the understanding of this new tool. Moreover, to be useful for applications we need to be able to compute the voltage across an inductor, capacitor and resistor given the phasor of the current which flows through these components. This is of paramount importance to be able to define the power and, more in general, to make this mathematical extension an accurate description of the physical quantities in the circuit. The expressions that relate voltage and current in an inductor, capacitor and resistor are, respectively,

$$v = L\frac{di}{dt}, \qquad v = \frac{1}{C}\int_0^t i\,d\tau, \qquad v = Ri. \tag{8.18}$$

Utilizing the phasor transform (8.2), it can be proved that the relations

$$\overline{V} = j\omega L\overline{I}, \qquad \overline{V} = \frac{1}{j\omega C}\overline{I}, \qquad \overline{V} = R\overline{I}, \tag{8.19}$$

hold. With the phasor transform (8.3) these become

$$\overline{V}(t) = L\dot{\overline{I}}(t) + j\omega(t)L\overline{I}(t), \qquad \dot{\overline{V}}(t) + j\omega(t)\overline{V}(t) = \frac{1}{C}\overline{I}(t), \qquad \overline{V}(t) = R\overline{I}(t), \tag{8.20}$$

When the source is described by the generator (8.15), these relations may not hold anymore. Consider, for instance, a square wave, which is described by the sum of infinitely many frequencies ω_k. As noted in [16], this observation suggests that we could describe a square wave by means of an infinite dimensional system. As a consequence, the matrix Π would have infinitely many rows and columns and the phasor would be the sum of infinitely many frequencies. It is exactly for this inability to deal with this type of signals without approximations that Definition 26 has been introduced. In fact, exploiting the

discontinuous phasor transform we obtain the following exact relations.

Theorem 25. Consider the first equation in (8.18). The relation

$$\overline{V}(t) = L\dot{\overline{I}}(t) + L\frac{\dot{\Lambda}(t)}{\Lambda(t)}\overline{I}(t), \tag{8.21}$$

holds.

Proof. Consider the first equation in (8.18). This is a scalar system with $\overline{I}(t) = \Pi_\infty(t) \in \mathbb{C}$, $A = 0$, $B = \dfrac{1}{L}$, $\Gamma = \overline{V}(t)$. The derivative of the current is

$$\begin{aligned}\frac{di}{dt} = \frac{d}{dt}\Re\left[\overline{I}(t)\Lambda(t)\right] = \Re\left[\frac{d}{dt}\left[\overline{I}(t)\Lambda(t)\right]\right] = \Re\left[\dot{\overline{I}}(t)\Lambda(t) + \overline{I}(t)\dot{\Lambda}(t)\right] \\ = \Re\left[\left(\dot{\overline{I}}(t) + \overline{I}(t)\dot{\Lambda}(t)\Lambda(t)^{-1}\right)\Lambda(t)\right]\end{aligned}$$

from which we recognize, by comparison with (8.17), that $\dot{\overline{I}}(t) + \overline{I}(t)\dot{\Lambda}(t)\Lambda(t)^{-1}$ is the phasor. □

Remark 70. If $\Lambda(t) = e^{j\omega t}$, then $\dot{\overline{I}}(t) = 0$, $\dot{\Lambda}(t)\Lambda(t)^{-1} = j\omega$ and (8.21) becomes the first relation in (8.19). If we use the phasor transform (8.3) in which the phasor $I(t)$ is not constant, (8.21) reduces to the first relation in (8.20). ■

Theorem 26. Consider the second equation in (8.18). The relation

$$\dot{\overline{V}}(t) + \frac{\dot{\Lambda}(t)}{\Lambda(t)}\overline{V}(t) = \frac{1}{C}\overline{I}(t), \tag{8.22}$$

holds.

Proof. It is similar to the proof of Theorem 25. □

Remark 71. In the mixed convention, the components with odd indeces of Π_∞, computed from (8.16), are those functions that multiplied by Λ give the steady-state of the integrals of the currents. The following result holds. ■

Corollary 6. In the mixed convention the components with odd indeces of Π_∞, computed with (8.16), are the phasors of the integrals of the currents in the circuit. Thus for the current i_k which flows in the capacitance C_k, the relation

$$\frac{\epsilon_{2k-1}^\top \Pi_\infty(t)}{C_k} = \overline{V}_k(t), \tag{8.23}$$

holds, where $\overline{V}_k$ is the phasor of the voltage across the capacitor C_k.

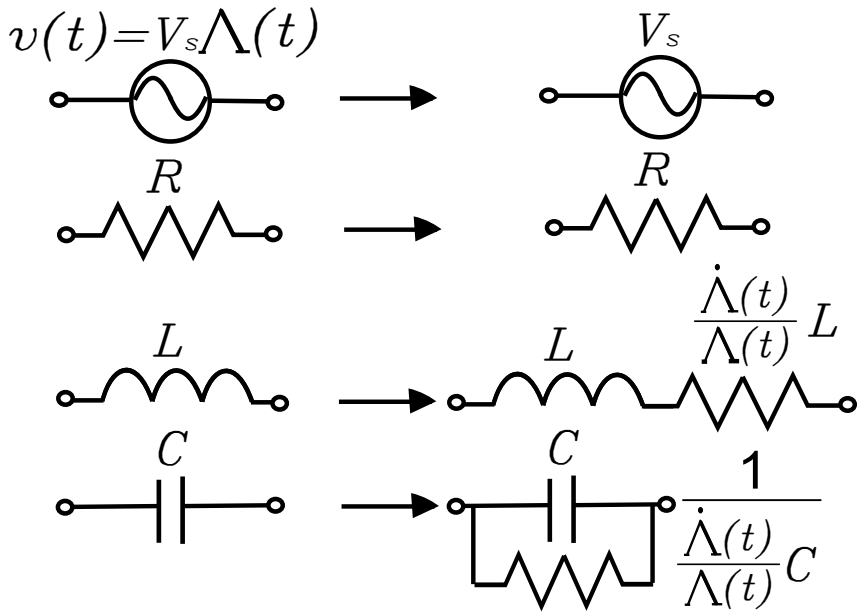

Figure 8.5: Phasor models for basic circuit elements.

This last result provides a way to compute the phasor of the voltage across a capacitor. In fact, this value comes directly from solving (8.16).

Remark 72. In [198–200] a discussion regarding the selection of the initial conditions of the equations (8.20) is missing. Being the phasor defined only as the solution of the differential equations (8.20) it is not clear how the correct solution is selected among the many identified by the differential equations. Instead, as a result of Definition 26 the initial conditions can be computed as $\Pi(0) = \Pi_\infty(0)$, using (8.16) for $t = 0$. In fact, this is the most efficient way to compute the phasors, namely solving the differential equations in the phasor domain *e.g.* (8.21) and (8.22), and using the computational expensive (8.16) only for $t = 0$. ■

Theorem 27. Consider the third equation in (8.18). The relation

$$\overline{V} = R\overline{I}(t), \tag{8.24}$$

holds.

Proof. The statement holds trivially noting that the multiplication by a real constant and the real part operator commutes. □

Fig. 8.5 shows the phasor equivalent of these circuit elements. Having defined

the differential operator, the integral operator and the multiplicative operator, the v-i characteristics of other common circuits in the phasor domain can be easily obtained. For instance, transformers and gyrators, in which voltages and currents are related by a multiplication factor, show phasor equations similar to (8.24).

Remark 73. In [198, 200], transformers with variable transformation ratio $s(t)$ are considered for their ability of replacing switches [206]. They are not needed in the present framework since the new phasors we have introduced naturally embed switches. Thus, these time-varying transformers are not analyzed further in the chapter. Note anyway that since their v-i characteristics are given by the multiplication of two time-varying functions, *i.e.* $v_t(t) = s(t)i_t(t)$, they present equations in the phasor domain similar to the ones developed in the following section for the power $p(t) = v(t)i(t)$, replacing $p(t)$ with $v_t(t)$ and $v(t)$ with $s(t)$. ■

8.3.3 Instantaneous, average and reactive power

Using the phasor transform (8.17) the instantaneous power is defined as

$$p(t) = v(t)i(t) = \Re\left[\overline{V}(t)\Lambda(t)\right]\Re\left[\overline{I}(t)\Lambda(t)\right], \tag{8.25}$$

which, exploiting the properties of the real part operator, yields

$$p(t) = \frac{1}{2}\Re\left[\overline{V}(t)^*\overline{I}(t)\Lambda(t)\Lambda(t)^*\right] + \frac{1}{2}\Re\left[\overline{V}(t)\overline{I}(t)\Lambda(t)^2\right]. \tag{8.26}$$

As in the sinusoidal case the instantaneous power is separated in two terms: the average of the first term is equal to the average power, whereas the average of the second term is zero. However, differently from the sinusoidal case, the first term is not constant, in general, and thus it is not equal to the average power. Hence, the average power and the reactive power are defined as follows.

Definition 27. In the phasor domain identified by the phasor transform (8.17), the average power P and the reactive power Q are defined as

$$P = <p(t)> = \frac{1}{2}\left\langle\Re\left[\overline{V}(t)^*\overline{I}(t)\Lambda(t)\Lambda(t)^*\right]\right\rangle, \qquad Q = \frac{1}{2}\left\langle\Im\left[\overline{V}(t)^*\overline{I}(t)\Lambda(t)\Lambda(t)^*\right]\right\rangle. \tag{8.27}$$

Equations (8.27) are consistent with the usual definition of average power and reactive

power in the complex exponential case. For the non-exponential case, equations (8.27) generalize the respective relations achievable with the phasor transforms (8.2) and (8.3). In fact, both of these always show a time-invariant transform of the average power [200]. Moreover, the sum of the average powers (or reactive powers) generated by the Fourier series expansion of the periodic signal Λ converges to the value given in (8.27). One can say more when specific signals are considered. For instance, if the input signal is a square wave, the following result holds.

Proposition 11. Assume that the signal generator (8.15) produces a square wave which takes values in the discrete set $\{-1, 1\}$. The average power and reactive power are described by

$$P = \frac{1}{4}\Re\left[\overline{V}(t)^*\overline{I}(t)\right], \qquad Q = \frac{1}{4}\Im\left[\overline{V}(t)^*\overline{I}(t)\right]. \tag{8.28}$$

In fact, note that the term $\Lambda(t)\Lambda(t)^*$ in (8.27) is constant and equal to 1 if the driving signal is a complex exponential, it is constant and equal to $\frac{1}{2}$ if the driving signal is a square wave and it is a parabola if the driving signal is a triangular wave.

8.4 Analysis of inverters and wireless power transfer systems

We now apply the newly defined discontinuous phasor transform to the steady-state description of the currents flowing in a Class D resonant inverter controlled by square waves and a wireless power transfer system with non-ideal switches (affected by the reverse recovery effect).

8.4.1 Example: analysis of a resonant inverter

The equation

$$Ri + L\frac{di}{dt} + \frac{1}{C}\int i_1 dt = u(t), \tag{8.29}$$

describes the circuit shown in Fig. 8.2. Consider now that we are interested in studying the steady-state behavior of this circuit for a particular switching function. The switching function is determined by the designer on the basis of the specific application. Thus, we assume that the switching function is a square wave with given angular frequency ω, *i.e.* $\sqcap(\omega t + \pi/2)$. This signal does not satisfy Assumption 15. However, this issue can be easily

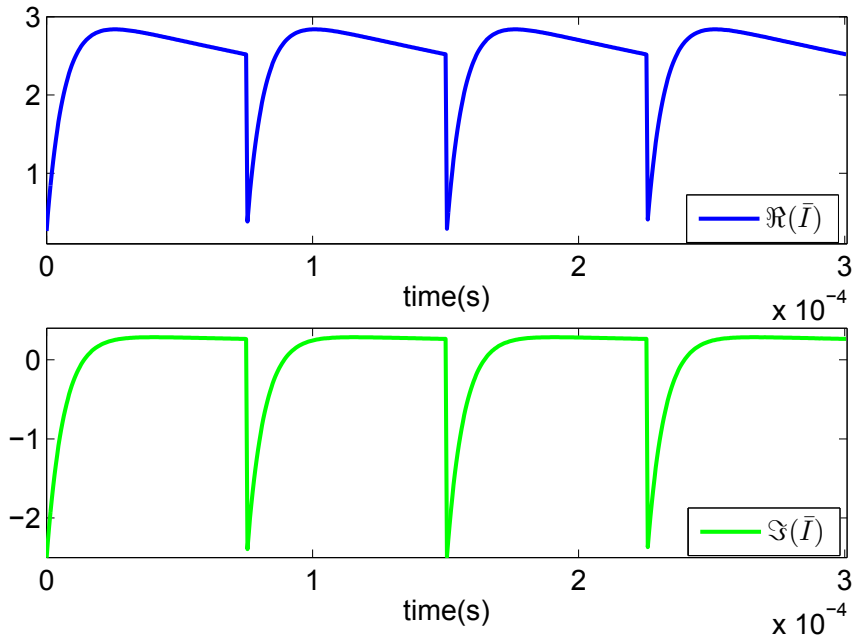

Figure 8.6: Real part (top graph) and imaginary part (bottom graph) of the phasor $\bar{I}(t)$.

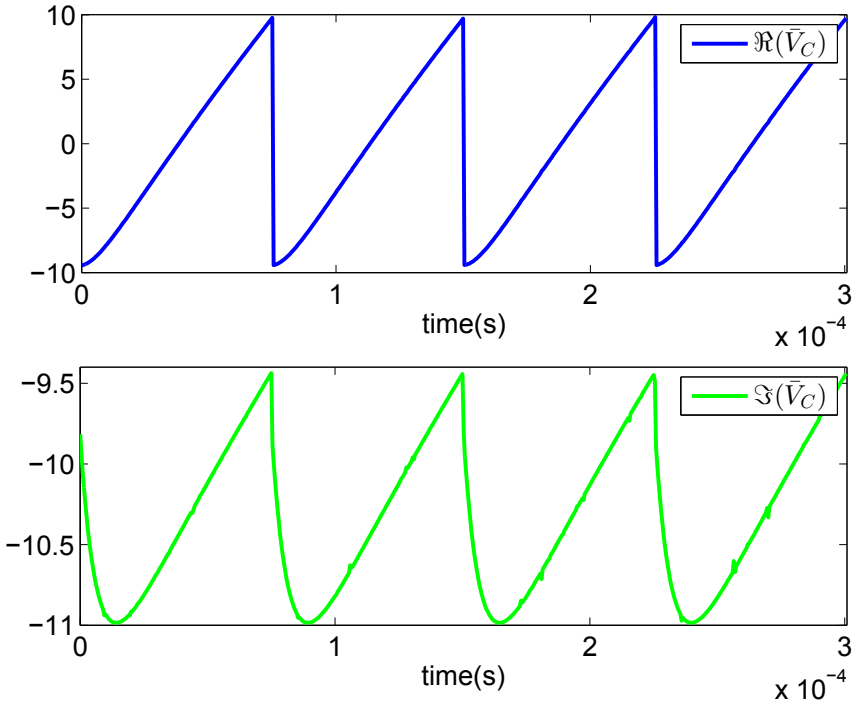

Figure 8.7: Real part (top graph) and imaginary part (bottom graph) of the phasor $\bar{V}_C(t)$.

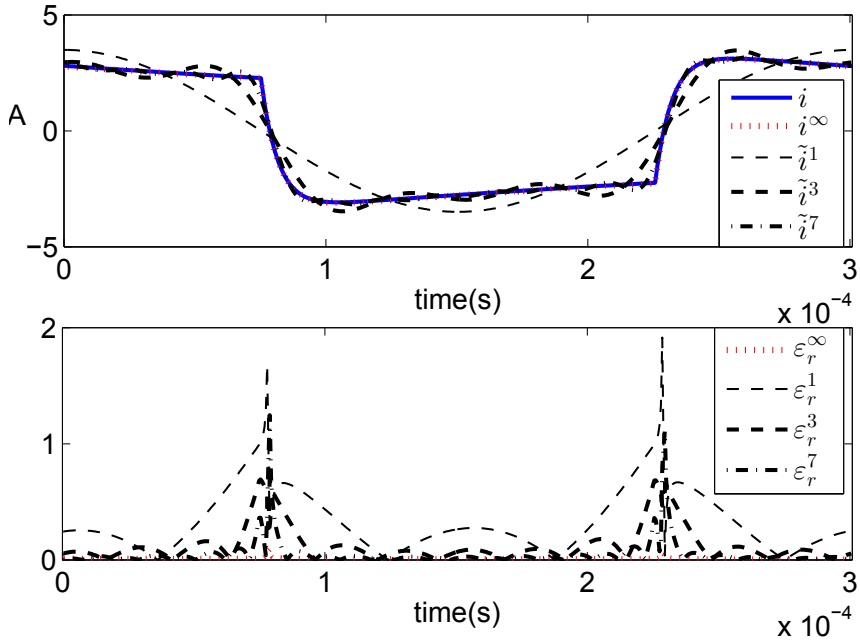

Figure 8.8: Top graph: time histories of the steady-state current $i(t)$ (solid/blue line), of the steady-state current $i^\infty(t)$ (dotted/red line) and of the approximated steady-state current computed with the Fourier series with one harmonic $\tilde{i}^1(t)$ (dashed/black line), three harmonics $\tilde{i}^3(t)$ (bold dashed/black line), and seven harmonics $\tilde{i}^7(t)$ (dash-dotted/black line). Bottom graph: resulting relative errors (with same color coding).

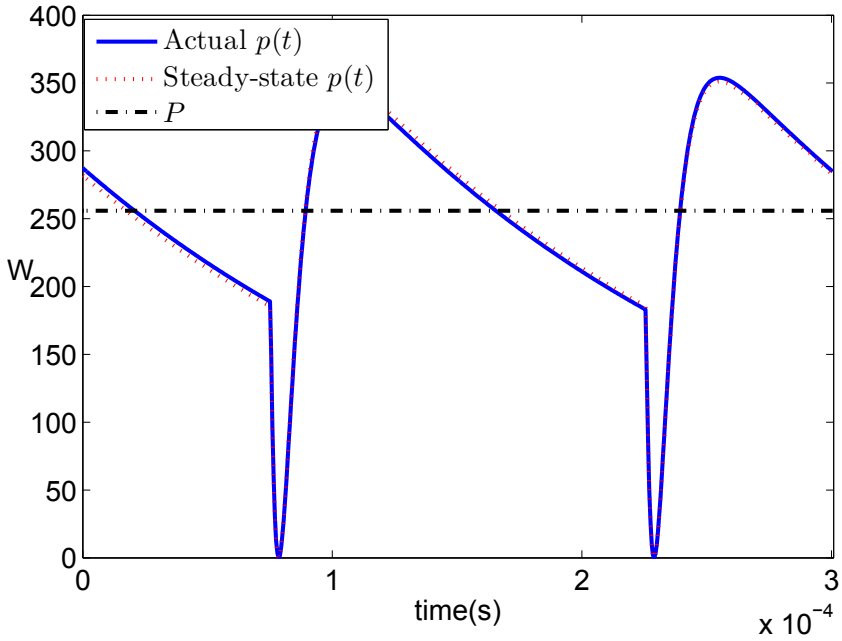

Figure 8.9: Instantaneous power absorbed by R (solid/blue line) computed from the current $i(t)$. Instantaneous power (dotted/red line) and average power (dash-dotted/black line) computed from the phasor $\overline{I}(t)$.

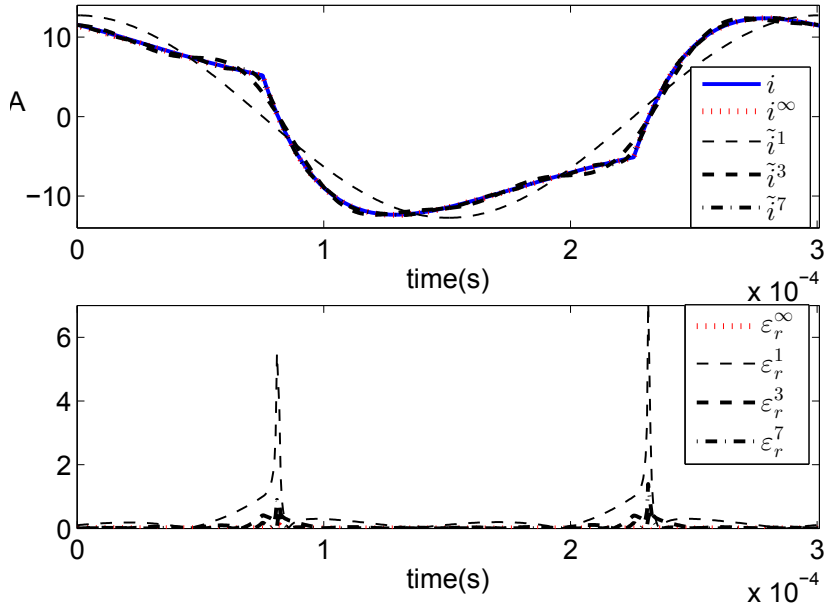

Figure 8.10: Top graph: time histories of the steady-state current $i(t)$ (solid/blue line), of the steady-state current $i^{\infty}(t)$ (dotted/red line) and of the approximated steady-state current computed with the Fourier series with one harmonic $\tilde{i}^{1}(t)$ (dashed/black line), three harmonics $\tilde{i}^{3}(t)$ (bold dashed/black line), and seven harmonics $\tilde{i}^{7}(t)$ (dash-dotted/black line). Bottom graph: resulting relative errors (with same color coding).

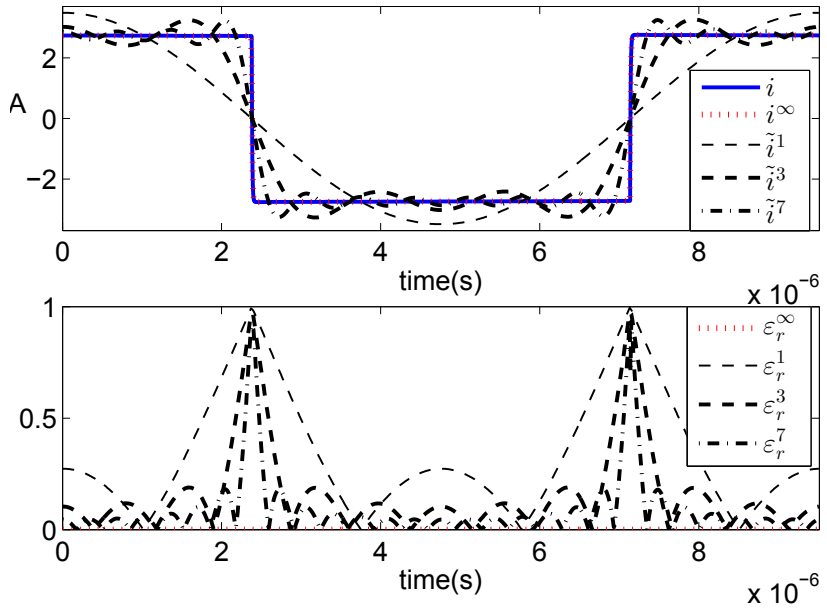

Figure 8.11: Top graph: time histories of the steady-state current $i(t)$ (solid/blue line), of the steady-state current $i^{\infty}(t)$ (dotted/red line) and of the approximated steady-state current computed with the Fourier series with one harmonic $\tilde{i}^{1}(t)$ (dashed/green line), three harmonics $\tilde{i}^{3}(t)$ (bold dashed/green line), and seven harmonics $\tilde{i}^{7}(t)$ (dash-dotted/black line). Bottom graph: resulting relative errors (with same color coding).

solved considering the extended signal $\sqcap(\omega t+\pi/2)+j\sqcap(\omega t)$. This complex signal is never equal to zero and Assumption 15 is satisfied. Note that this is in line with the smooth case when the source is described by $e^{j\omega t}=\sin(\omega t+\pi/2)+j\sin(\omega t)$.
Thus, we write equation (8.29) in the form (8.4) using the mixed convention and we consider the signal generator described by the equations

$$u(t)=V_s\zeta(t),\qquad \zeta(t)=\sqcap\left(\omega t+\frac{\pi}{2}\right)+j\sqcap(\omega t)\,, \tag{8.30}$$

with $V_s\in\mathbb{R}$ and $\sqcap(t)$ implemented as the function *square* of MATLAB^{TM}. Note that in the real convention this signal is realized by the system

$$u(t)=V_s\begin{bmatrix}1 & 0\end{bmatrix}\zeta(t),\qquad \zeta(t)=\begin{bmatrix}\sqcap\left(\omega t+\frac{\pi}{2}\right) & -\sqcap(\omega t)\\ \sqcap(\omega t) & \sqcap\left(\omega t+\frac{\pi}{2}\right)\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}. \tag{8.31}$$

In fact, in the smooth case the matrix Λ would reduce to the usual rotation matrix and $u(t)=\cos(\omega t)$. However, in the following we use the mixed convention which, as already pointed out, is more compact and it results in more efficient computation.

Analysis of a resonant inverter with Q=0.1313

The parameters for the simulation have been selected as $L=229.3\,\mu H$, $C=10\,\mu F$, $R=36.47\,\Omega$, $V_s=100\,V$ and $\omega=\dfrac{1}{\sqrt{LC}}$. Note that the inverter has a low quality factor of Q= 0.1313. The phasors are computed using equation (8.16). The formula has been implemented in MATLAB^{TM} with the function *integral* with the option *ArrayValued*. Fig. 8.6 shows the real component (top graph) and the imaginary component (bottom graph) of the phasor $\overline{I}(t)$ computed as $\epsilon_1^\top\Pi_\infty$. Note that the phasor is time-dependent, periodic (with period one forth of the period of the input source) and discontinuous. The other electrical quantities of the circuit in the phasor domain can be computed as well. For instance Fig. 8.7 shows the phasor $\overline{V}_C(t)=\epsilon_2^\top\Pi_\infty/C$, computed as in equation (8.23), of the voltage across the capacitor C. Fig. 8.8 (top graph) shows the time histories of the steady-state current $i(t)$ (solid/blue line), the steady-state current i^∞ computed from the phasor $\overline{I}(t)$ as $i^\infty=\Re\left[\overline{I}(t)\Lambda(t)\right]$ (dotted/red line) and the approximated steady-state

current computed with the Fourier series with one harmonic $\tilde{i}^1(t)$ (dashed/black line), three harmonics $\tilde{i}^3(t)$ (bold dashed/black line), and seven harmonics $\tilde{i}^7(t)$ (dash-dotted/black line). Note that the transient of $i(t)$ has been eliminated simulating 50 periods of the switching input and considering only the last period. Fig. 8.8 (bottom graph) shows the corresponding relative errors (with the same color coding), namely $\varepsilon_r^\infty = |i - i^\infty|/|i|$ and $\varepsilon_r^k = |i - \tilde{i}^k|/|i|$, with $k = 1, 3, 7$. The steady-state of the current $i(t)$ is exactly described by i^∞ (the error is identically zero). The steady-state is approximated by the Fourier series increasingly better as the number of terms grows. However, note that at the points of discontinuity a finite number of Fourier terms cannot approximate perfectly the steady-state, as it can be noted by the peaks of ε_r^7 (dash-dotted/black line). Fig. 8.9 shows the absorbed power by the load R. The solid line shows the actual steady-state instantaneous power computed from the current $i(t)$. The dotted line and dash-dotted line show the steady-state instantaneous power $p(t)$ computed as (8.26) and the steady-state average power P computed as (8.28), respectively. The figure confirms that the quantities that we have defined maintain their physical meaning since the power relations keep holding in the phasor domain.

Analysis of a resonant inverter with Q=0.4789 and Q=0.0042

We expect that as the "severity" of the discontinuity increases in the system, the discrepancy between the Fourier series and the discontinuous phasor response grows. We can check this changing a few parameters: for example setting $R = 10\,\Omega$ (Q= 0.4789) in the first case and $L = 229.3\,nH$ and $R = 36.47\,\Omega$ (Q= 0.0042) in the second case. Fig. 8.10 shows the quantities in Fig. 8.8 for the intermediate quality factor $Q = 0.4789$. Note that since the curves are smoother the current computed with the Fourier series approaches the current computed with the new phasor with fewer harmonics. On the other hand, Fig. 8.11 shows the quantities in Fig. 8.8 for the lower quality factor Q= 0.0042. In this case the error of the Fourier series increases. These figures show that the new discontinuous phasor is particularly useful to describe underdamped circuits that present "severe" discontinuities.

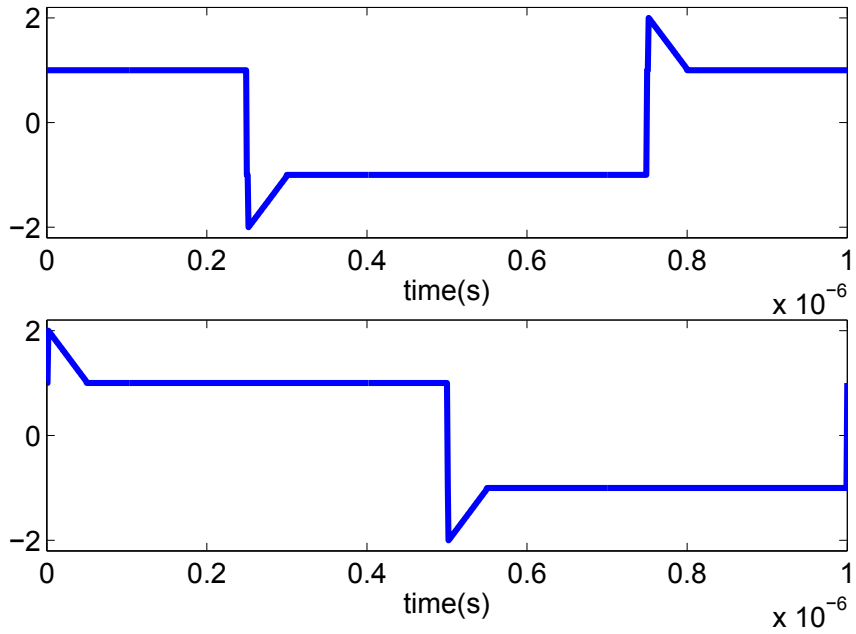

Figure 8.12: Real part (top graph) and imaginary part (bottom graph) of the signal representing a non-ideal switch with reverse recovery effect.

8.4.2 Example: analysis of a wireless power transfer system with non-ideal switches

Consider now system (8.8) represented in the mixed convention by system (8.4) with the matrices as in (8.10) and assume that the circuit contains non-ideal switches. In particular we study the effect of the reverse recovery time of the switching diode BAS19 [207]. The reverse recovery effect consists in an overshoot of the switching signal at the switching time. Most commercially available switching diodes have a reverse recovery time ranging from few nanoseconds to one microsecond [204]. The diode BAS19, which is considered a fast diode, has a reverse recovery time of 50 nanoseconds that becomes relevant at MHz switching frequencies at which some wireless power transfer systems operate [208]. Fig. 8.12 shows the behavior of the diode BAS19 at the frequency $f = 1\,MHz$. This has been modeled with equation (8.30) in which we have added a sawtooth wave at the switching times. The parameters for the simulation have been selected as $L_1 = L_2 = 8.203\,mH$, $M_{12} = M_{21} = 1.545\,mH$, $C_1 = 1.029\,nF$, $C_2 = 1.024\,nF$, $R1 = R_{2L} = 33.576\,k\Omega$, $V_s = 230\,V$ and $\omega = 2\pi f$. Fig. 8.13 shows the time histories of the 50-th period of the current i_2, the steady-state $i_2^{\infty} = \Re\left[\overline{I}_2(t)\Lambda(t)\right]$, and the approximation $\widetilde{i}_2^{15}(t)$

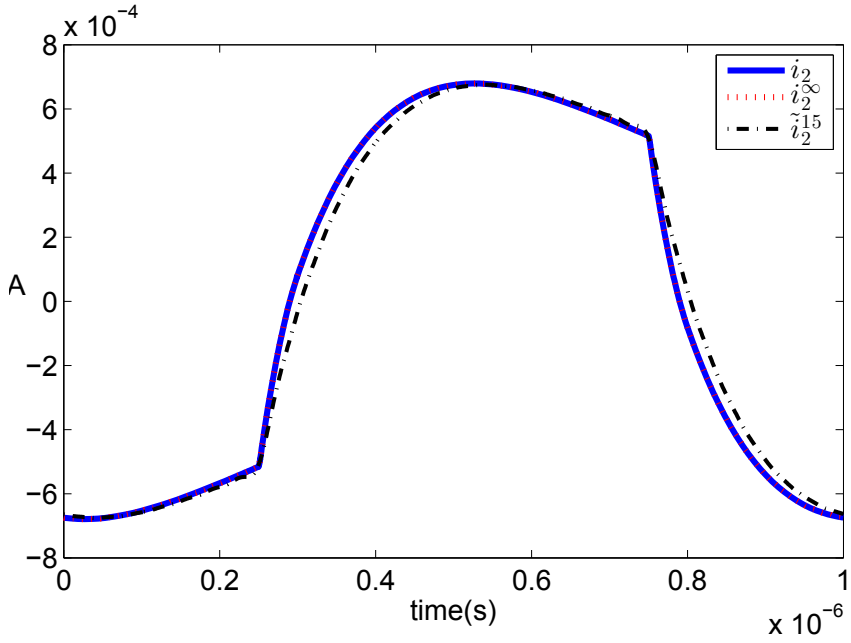

Figure 8.13: Time histories of the steady-state current $i_2(t)$ (solid/blue line), of the steady-state current $i_2^\infty(t)$ (dotted/red line) and of the approximated steady-state current $\tilde{i}_2^{15}(t)$ computed with the Fourier series with fifteen harmonics (dash-dotted/black line).

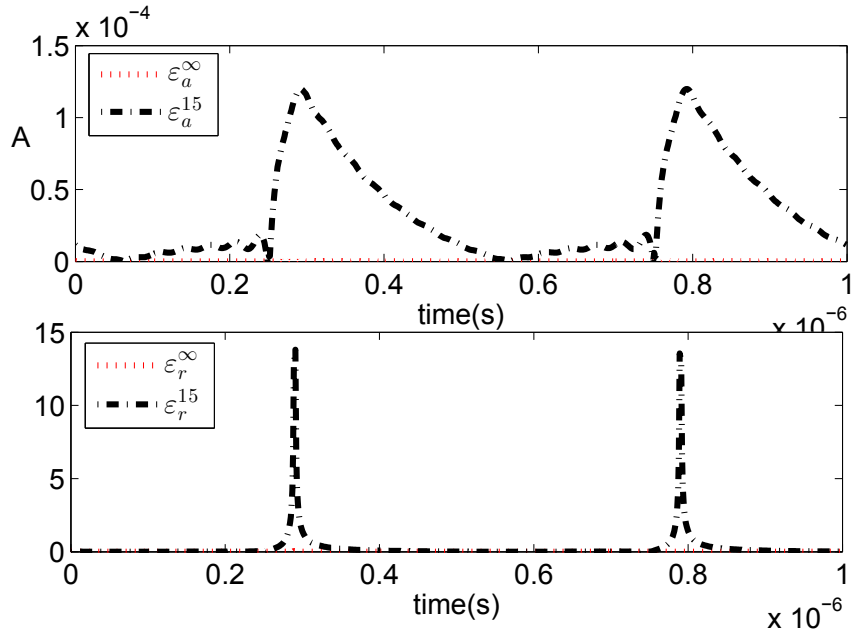

Figure 8.14: Time histories of the absolute errors between the steady-state current $i_2(t)$ and i_2^∞ (dotted/red line) and $\tilde{i}_2^{15}(t)$ (dash-dotted/black line), respectively.

computed with the Fourier series of an ideal switch with fifteen harmonics. Fig. 8.14 shows the corresponding absolute errors (top) and relative error (bottom). Note that the new discontinuous phasor can deal with the discontinuous non-ideal source without approximations. On the other hand, the Fourier series cannot fully describe this source.

8.5 Conclusion

We have shown that the phasors of an electric circuit are the moments on the imaginary axis of the linear system describing the circuit. Exploiting this relation, we have analyzed circuits powered by discontinuous sources. A new "discontinuous phasor transform", which allows to analyze in closed-form the steady-state behavior of discontinuous power electronic devices, has been defined and the v-i characteristics for inductors, capacitors and resistors have been described in terms of this new phasor transform. The new quantities maintain their physical meaning: the instantaneous power, the average power and the reactive power in the phasor domain have been defined. We have illustrated the use of this mathematical tool studying the steady-state response of power inverters and of wireless power transfer systems with non-ideal switches. The equivalence that we have established between moments and phasors has the potential of allowing further extensions of the phasor transform beyond the linear framework and the discontinuous framework to circuits with nonlinear elements and circuits with delays.

Chapter 9

Invariance-like theorems and "lim inf" convergence properties

9.1 Introduction

The qualitative study of asymptotic properties of trajectories of nonlinear systems is a key problem in systems and control theory, see *e.g.* [1, 209–212]. Among these asymptotic properties, the most important is attractiveness, which is often established by means of Lyapunov functions. Although this formulation is convenient from a practical point of view, it is in general hard to find a function that fulfills the sufficient (and in some cases necessary) conditions of the Lyapunov theorems. It is somewhat easier to find a weak Lyapunov function, *i.e.* a positive definite function with negative semi-definite time derivative along the trajectories of the systems. In this last case, for time-invariant systems, the Krasovskii-LaSalle invariance principle allows to establish attractiveness, under additional assumptions on the $\Omega-$limit sets of the solutions (see *e.g.* [213–216]).

Another tool that is used to replace the negative definiteness condition of the Lyapunov theorem and, in addition, can be used for time-varying systems is Matrosov Theorem (see *e.g.* [217–219]). This theorem allows proving attractiveness of equilibrium points, provided that a linear combination of positive semi-definite functions is positive definite and their time derivatives along the trajectories of the system have a triangular structure. However, to apply Matrosov Theorem it is necessary to assume stability of the equilibrium point.

In [220], the authors posed the question of what can be established if this stability as-

sumption and the positive definiteness assumption are removed from Matrosov Theorem. The answer to this question is that it is still possible to establish some convergence result, although, with reference to the positive definiteness assumption, not as strong as one could think borrowing from the Krasovskii-LaSalle invariance principle. We call this convergence "lim inf" convergence, in the sense that we cannot establish asymptotic stability of the equilibrium point, but we can show that there is an oscillating behavior with some "nice" asymptotic properties.

Since (the classical) Matrosov Theorem relies upon the study of a triangular system of two differential inequalities, it can be extended in different directions. A straightforward extension consists in the so-called nested Matrosov Theorem [219], in which several inequalities are considered. Another extension, which changes radically the reach of the theorem, consists in removing the triangular structure of the system of differential inequalities. In this case the Matrosov inequalities, which can also be interpreted within the framework of vector Lyapunov functions (see [221] for instance), lead (assuming additional hypotheses) to the Lyapunov formulation of the Small-gain Theorem [222].

The Small-gain Theorem is an important tool to assess the asymptotic properties of the trajectories of a system resulting from the interconnection of two or more subsystems. The Small-gain Theorem has been developed in different formulations depending on which property is used to describe the input-output behavior of each of the subsystems. For linear systems the L_p Small-gain Theorem has been successfully used in input-output formulations of the problem (see *e.g.* [223, Chapter 6] and [224]). For nonlinear systems versions based on L_p-gains (see [225]), but using Lyapunov functions, have been presented in [226–228]. In this chapter, the Lyapunov formulation given in [222] and derived from the property of input-to-state stability (ISS) (see [229, 230]) is used. Note that within this framework other formulations in which interconnections between possibly non-ISS subsystems are considered (see *e.g.* [231–233]), have been proposed. While the Small-gain Theorem is usually formulated for two interconnected subsystems it is often interesting, for practical applications, to study its large-scale version. A large-scale version of the theorem for linear systems can be found in [223], whereas a nonlinear formulation has recently been presented in [234, 235] (see also [221]).

The outcome of this chapter is a class of theorems inspired by the Krasovskii-LaSalle invariance principle that can establish "lim inf" convergence results, thus can "describe" the oscillatory behavior of the solutions of dynamical systems. These theorems lead to "lim inf" Matrosov and Small-gain Theorems which are based on a "lim inf" Barbalat's Lemma. In addition, technical assumptions to have "lim" convergence[1] are given, and the "lim inf" / "lim" relation and the role of some of the assumptions are discussed.

The rest of the chapter is organized as follows. This section continues with the formulation of the problem and with a discussion on the connections with Matrosov Theorem and the Small-gain Theorem. In Section 9.2 some properties of the so-called M-matrices are recalled. We also give a small-gain-like condition and we extend the concept of irreducibility to non-constant matrices. Section 9.3 presents a series of technical lemmas which forms the core theoretical part of the chapter. The irreducible case is analyzed and connections with Barbalat's Lemma are drawn before studying the general reducible case. In Section 9.4 the use of "linear gains" as opposed to "nonlinear gains" is justified and supported by a series of counter-examples. In Section 9.5 the theorems are applied to the study of the asymptotic behavior of solutions of dynamical systems. Sections 9.5.1 and 9.5.2 contain examples illustrating the theoretical results and Section 9.6 gives some concluding remarks.

Most results of this chapter are original contributions developed in fulfillment of my PhD course and they have been published in the conference paper [21] and in the journal paper [22]. Early versions of some of the results presented in this chapter (Lemma 23, Lemma 25 and Section 9.4) are due to [220, 236]

Notation. A continuous function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to belong to class K_∞, if it is strictly increasing, $\alpha(0) = 0$ and $\alpha(s) \to +\infty$ as $s \to +\infty$. Id denotes the identity map, *i.e.* $\mathrm{Id}(s) = s$. $[\boldsymbol{v}]_i$ denotes the i-th component of the vector $\boldsymbol{v}$ and the notation $\boldsymbol{u} \leq \boldsymbol{v}$ has to be understood component-wise.

[1] "lim" convergence is an abuse of language that we use to indicate the usual concept of convergence and that allows to clarify the distinction between the two "types of convergence".

9.1.1 Problem formulation

Motivated by the attempt to add a new tool to "comparison theory" for studying the behavior of the solutions of dynamical systems, we consider the following problem.

Problem 1. Let $i \in \{1, \dots, p\}$ and $j \in \{1, \dots, p\}$. Let $a_i : \mathbb{R}_{\geq 0} \to [-\overline{a}, \overline{a}]$ be absolutely continuous functions and $b_i : \mathbb{R}_{\geq 0} \to [0, \overline{b}]$ be continuous functions. Consider continuous, positive definite functions $\alpha_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and continuous functions $\beta_{ij} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, with $i \neq j$, satisfying $\beta_{ij}(0) = 0$, such that the differential inequalities

$$\begin{aligned} \dot{a}_1 &\leq -\alpha_1(b_1) + \beta_{12}(b_2) + \cdots + \beta_{1p}(b_p), \\ \dot{a}_2 &\leq -\alpha_2(b_2) + \beta_{21}(b_1) + \cdots + \beta_{2p}(b_p), \\ &\vdots \\ \dot{a}_p &\leq -\alpha_p(b_p) + \beta_{p1}(b_1) + \cdots + \beta_{p(p-1)}(b_{p-1}), \end{aligned} \tag{9.1}$$

hold for almost all t in $\mathbb{R}_{\geq 0}$.

The *"lim inf" / "lim" convergence problem* consists in determining the asymptotic properties of the functions b_i, more precisely, in determining conditions on the functions α_i and β_{ij} such that

$$\liminf_{t \to \infty} \sum_{i=1}^{p} b_i(t) = 0, \tag{9.2}$$

or

$$\lim_{t \to \infty} \sum_{i=1}^{p} b_i(t) = 0. \tag{9.3}$$

The key feature of the inequalities (9.1) is that the arguments b_i of the functions α_i and β_{ij} are not related *a-priori* with the functions a_k in the left-hand side. To illustrate this statement we recall the (simplest) formulation of the Matrosov Theorem and of the Small-gain Theorem.

Consider a nonlinear system described by the equation

$$\dot{x} = f(x), \tag{9.4}$$

where $x \in \mathbb{R}^n$ is the state of the system and the function $f : \mathbb{R}^n \to \mathbb{R}^n$ is locally Lipschitz. Assume there exists an equilibrium point which, without loss of generality, we choose as the origin of the coordinate system, *i.e.* $f(0) = 0$.

Theorem 28 (Matrosov Theorem [217], [218], [219], [220])**.** Consider system (9.4). Let $i \in \{1, 2\}$. Assume there exist

1. a differentiable, positive definite and radially unbounded function $V_0 : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ such that $\dot{V}_0 \leq 0$ along all the solutions of system (9.4);
2. two differentiable functions $V_i : \mathbb{R}^n \to \mathbb{R}$ and two continuous, positive semi-definite functions $h_i : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ such that the function $h_1 + h_2$ is positive definite;
3. a continuous function $\beta_{21} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, such that $\beta_{21}(0) = 0$;

satisfying, along all the solutions of system (9.4), the inequalities

$$\begin{aligned} \dot{V}_1 &\leq -h_1, \\ \dot{V}_2 &\leq -h_2 + \beta_{21}(h_1). \end{aligned} \tag{9.5}$$

Then the equilibrium $x = 0$ of system (9.4) is globally asymptotically stable.

Theorem 29 (Small-gain Theorem[2] [222, 236])**.** Consider system (9.4). Let $i \in \{1, 2\}$. Assume there exist

1. two continuously differentiable functions $V_i : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ such that the function $V_1 + V_2$ is positive definite and radially unbounded;
2. two class K_∞ functions $\alpha_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and two continuous functions β_{12}, $\beta_{21} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, such that $\beta_{12}(0) = \beta_{21}(0) = 0$;

satisfying, along all the solutions of system (9.4), the inequalities

$$\begin{aligned} \dot{V}_1 &\leq -\alpha_1(V_1) + \beta_{12}(V_2), \\ \dot{V}_2 &\leq -\alpha_2(V_2) + \beta_{21}(V_1). \end{aligned} \tag{9.6}$$

If the small-gain condition

$$\beta_{21} \circ \alpha_1^{-1} \circ \beta_{12} \circ \alpha_2^{-1} < \mathrm{Id}, \tag{9.7}$$

holds, then the equilibrium $x = 0$ of system (9.4) is globally asymptotically stable.

[2]The Small-gain Theorem is usually applied in the study of the stability properties of the equilibrium point of an interconnected system. In this chapter, following the Lyapunov formulation given in [222], we abuse the terminology saying that the Small-gain Theorem holds for the inequalities (9.6), ignoring if these arise as the result of a composition of systems.

Note that in Problem 1 and Theorem 28 and 29 differential inequalities with similar structure are studied; in Theorem 28 and 29 the inequalities must hold when the functions are evaluated along any solution. Instead in Problem 1, we restrict our attention to those particular solutions which are bounded.

We are also interested in generalizing Theorem 28, removing the stability assumption and not requiring that a linear combination of positive-semidefinite functions be positive definite (in the spirit of LaSalle invariance principle), and Theorem 29, allowing the arguments b_i of the functions α_i and β_{ij} to be not related *a-priori* with the functions a_k in the left-hand side (in the spirit of Matrosov Theorem). Note that as anticipated in [236] and illustrated in detail here, the result that we prove may not hold when the nonlinear functions α_i and β_{ij} satisfy the nonlinear small-gain condition (9.7): a more restrictive linear small-gain-like condition may be required.

9.2 Preliminary results on the test matrix

In this section we define the notion of "test matrix" associated to the inequalities (9.1) and we recall or prove properties which are instrumental to establish the results of the following sections.

Definition 28. A principal minor of order j of an $n \times n$ matrix A is the determinant of the $j \times j$ sub-matrix obtained from A by deleting $n - j$ columns and $n - j$ rows with the same indices.
A leading principal minor of order j of a matrix A is the determinant of its upper-left j by j sub-matrix and is indicated by the notation $\mathcal{M}_j(A)$.

Definition 29. [237] A Z-matrix is a matrix with non-positive off-diagonal elements.

Definition 30. [237, Condition E_{17}, Theorem 6.2.3] A Z-matrix having all its leading principal minors strictly positive is called a non-singular M-matrix.

Definition 31. [237] A matrix is reducible if, after some permutation of the rows and the columns, it can be written in a lower block triangular form. Otherwise it is said to be irreducible.

Lemma 18. [237, Theorem 6.2.7] The inverse of a non-singular M-matrix A has non-negative entries. Moreover, if A is irreducible, the inverse has strictly positive entries.

In the following we call *test matrix* the matrix Q with the (i,j) element equal to $-\beta_{ij}(b_j)$, if $i \neq j$, or to $\alpha_j(b_j)$, if $i = j$, namely[3]

$$Q = \begin{bmatrix} \alpha_1(b_1) & -\beta_{12}(b_2) & \dots & -\beta_{1p}(b_p) \\ -\beta_{21}(b_1) & \alpha_2(b_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ -\beta_{(p-1)1}(b_1) & \dots & \alpha_{p-1}(b_{p-1}) & -\beta_{(p-1)p}(b_p) \\ -\beta_{p1}(b_1) & \dots & -\beta_{p(p-1)}(b_{p-1}) & \alpha_p(b_p) \end{bmatrix}.$$

Note that Q is a Z-matrix.

When, for all $k = 1, \dots, p$, $l = 1, \dots, p$, with $k \neq l$, there exist non-negative real numbers γ_{kl} satisfying

$$\sup_{s \in (0, \bar{b}]} \frac{\beta_{kl}(s)}{\alpha_l(s)} \leq \gamma_{kl}, \tag{9.8}$$

we associate to the test matrix Q a matrix Γ defined as the matrix with off-diagonal elements equal to $-\sup_{s \in (0, \bar{b}]} \dfrac{\beta_{kl}(s)}{\alpha_j(s)}$ and diagonal elements equal to one. Again Γ is a Z-matrix.

Lemma 19. Assume the following.

1. The test matrix Q satisfies the following *linear small-gain-like condition:*
   there exists a strictly positive real number ε such that, for all
   $j = 1, \dots, p$ and all $(b_1, \dots, b_p)$ in $[0, \bar{b}\,]^p$, we have
   $$\mathcal{M}_j(Q(b_1, \dots, b_p)) \geq \varepsilon \prod_{k=1}^{j} \alpha_k(b_k). \tag{9.9}$$
2. Each function $s \mapsto \dfrac{\beta_{kl}(s)}{\alpha_l(s)}$ is bounded.

Then the matrix Γ satisfies

$$\mathcal{M}_j(\Gamma) \geq \varepsilon, \qquad \forall\, j = 1, \dots, p. \tag{9.10}$$

Proof. Condition (9.9) is equivalent to

$$\mathcal{M}_j \left(Q \operatorname{diag} \left(\frac{1}{\alpha_1}, \dots, \frac{1}{\alpha_p} \right) \right) \geq \varepsilon, \quad \forall\, j \text{ and } \forall\, b_i \in (0, \bar{b}\,].$$

[3] Omitting the arguments of Q.

By definition of supremum, there exist p sequences $\{b_{ln}\}$ such that

$$\frac{\beta_{kl}(b_{ln})}{\alpha_l(b_{ln})} \leq \sup_{s\in(0,\bar{b}]} \frac{\beta_{kl}(s)}{\alpha_l(s)} \leq \frac{\beta_{kl}(b_{ln})}{\alpha_l(b_{ln})} + \frac{1}{n}.$$

Since a minor is a polynomial in the entries of the matrix and the b_{ln} are bounded, this yields

$$\mathcal{M}_j(\Gamma) \geq \varepsilon + p\left(\frac{1}{n}\right), \qquad \forall\, j = 1, \ldots, p,$$

where $p\left(\frac{1}{n}\right)$ is a polynomial in $\frac{1}{n}$ that goes to zero as $n \to \infty$, *i.e.* $p(0) = 0$, hence the claim. □

Another way to make sure that (9.10) holds when the second assumption of Lemma 19 is satisfied is by defining a matrix $\overline{\Gamma}$ with off-diagonal elements equal to some $-\gamma_{kl}$ satisfying (9.8) and diagonal elements equal to one and check if we have

$$\mathcal{M}_j(\overline{\Gamma}) \geq \overline{\varepsilon} > 0, \qquad \forall\, j = 1, \ldots, p.$$

Indeed in this case, we have

$$\mathcal{M}_j(\Gamma) \geq \mathcal{M}_j(\overline{\Gamma})$$

This follows from the fact that Lemma 18 implies that $\mathcal{M}_j(\overline{\Gamma})$ is a non-increasing function of γ_{kl}.

We show now that, when the small-gain-like condition (9.9) is satisfied, the irreducibility of Q implies the boundedness of the functions $s \mapsto \dfrac{\beta_{ij}(s)}{\alpha_j(s)}$ on $(0, \bar{b}]$.

Lemma 20. Assume the test matrix Q satisfies the linear small-gain-like condition (9.9). If, for some index j, there exists a vector $(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^* \ldots, b_p^*)$ in $(0, \bar{b}\,]^{p-1}$ such that, for all $b_j \in (0, \bar{b}\,]$, the matrix $Q(b_1^*, \ldots, b_{j-1}^*, b_j, b_{j+1}^* \ldots, b_p^*)$ is irreducible then, for all $i \neq j$, the functions $s \mapsto \dfrac{\beta_{ij}(s)}{\alpha_j(s)}$ are bounded on $(0, \bar{b}\,]$.

Proof. By definition, the small-gain-like condition (9.9) implies that $Q(b_1, \ldots, b_p)$ is an M-matrix for all $(b_1, \ldots, b_p)$ in $[0, \bar{b}\,]^p$. It implies also that $\det(Q(b_1, \ldots, b_p))$ is strictly positive for all $(b_1, \ldots, b_p)$ in $(0, \bar{b}\,]^p$.

Let j and $(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^* \ldots, b_p^*)$ in $(0, \bar{b}\,]^{p-1}$ be such that the matrix $Q(b_1^*, \ldots, b_{j-1}^*, s, b_{j+1}^* \ldots, b_p^*)$ is irreducible for all s in $(0, \bar{b}\,]$. Omitting the argu-

ment $(b_1^*, \ldots, b_{j-1}^*, s, b_{j+1}^* \ldots, b_p^*)$ when it is not necessary, let q_k, for all k, be the k-th entry of the j-th row of $\mathrm{adj}(Q) = \det(Q)Q^{-1}$. From Lemma 18 q_k is strictly positive. Also q_k does not depend on s. Indeed the row j of Q^T depends only on s, and q_k is the determinant of the sub-matrix formed by deleting the j-th row and the k-th column of Q^T. Finally, the j-th diagonal element of the matrix identity

$$\det(Q)I = \mathrm{adj}(Q)Q,$$

yields

$$0 < \det(Q) = q_j(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^*, \ldots, b_p^*)\alpha_j(s) - \sum_{\substack{k=1 \\ k\neq j}}^{p} q_k(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^*, \ldots, b_p^*)\beta_{kj}(s),$$

for all $s \in (0, \bar{b}]$. Since for any $i \neq j$, q_i is strictly positive, this implies

$$\beta_{ij}(s) + \sum_{\substack{k=1 \\ k\neq j,i}}^{p} \frac{q_k(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^*, \ldots, b_p^*)}{q_i(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^* \ldots, b_p^*)}\beta_{kj}(s) < \frac{q_j(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^* \ldots, b_p^*)}{q_i(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^* \ldots, b_p^*)}\alpha_j(s),$$

for all $s \in (0, \bar{b}]$, for all $i \neq j$, and therefore

$$\frac{\beta_{ij}(s)}{\alpha_j(s)} < \frac{q_j(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^* \ldots, b_p^*)}{q_i(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^* \ldots, b_p^*)} \qquad \forall s \in (0, \bar{b}], \quad \forall i \neq j.$$

□

In view of this result we define what we mean by the fact that Q as a function of $(b_1, \ldots, b_p)$ is irreducible.

Definition 32. A test matrix is said to be irreducible as a function if, for each index j, there exists a vector $(b_1^*, \ldots, b_{j-1}^*, b_{j+1}^* \ldots, b_p^*)$ in $(0, \bar{b}\,]^{p-1}$ such that, for all $b_j \in (0, \bar{b}\,]$, the matrix $Q(b_1^*, \ldots, b_{j-1}^*, b_j, b_{j+1}^* \ldots, b_p^*)$ is irreducible.

The outcome of Lemma 20 is that if the inequalities in (9.1) cannot be re-written in triangular form by means of a permutation of rows and columns or more precisely if the associated test matrix is irreducible as a function, then the linear small-gain-like condition implies the existence of the matrix Γ with no additional hypotheses. In other words, when Q is irreducible as a function and (9.9) holds there is no need to assume that the functions

$s \mapsto \dfrac{\beta_{ij}(s)}{\alpha_j(s)}$ are bounded.

9.3 Main technical results

In this section we present lemmas which constitute the core theoretical part of the chapter. They will be used to establish the results of the following sections dealing with the study of the behavior of solutions of ordinary differential equations which are known to exist on $[0, +\infty)$, and taking values in a compact set, as detailed in Problem 1. For this reason we assume, without loss of generality, that all functions are bounded.

We begin with the irreducible case in the first subsection, we study the triangular reducible case in the second and we conclude with the triangular block reducible case in the last.

9.3.1 Irreducible case

Lemma 21. Let $i \in \{1, \dots, p\}$ and $j \in \{1, \dots, p\}$. Let $a_i : \mathbb{R}_{\geq 0} \to [-\overline{a}, \overline{a}]$ be absolutely continuous functions and $b_i : \mathbb{R}_{\geq 0} \to [0, \overline{b}]$ be continuous functions. Consider continuous, positive definite functions $\alpha_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and continuous functions $\beta_{ij} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, with $i \neq j$, satisfying $\beta_{ij}(0) = 0$, such that the following hold.

1. The differential inequalities

$$\begin{aligned} \dot{a}_1 &\leq -\alpha_1(b_1) + \beta_{12}(b_2) + \cdots + \beta_{1p}(b_p), \\ \dot{a}_2 &\leq -\alpha_2(b_2) + \beta_{21}(b_1) + \cdots + \beta_{2p}(b_p), \\ &\vdots \\ \dot{a}_p &\leq -\alpha_p(b_p) + \beta_{p1}(b_1) + \cdots + \beta_{p(p-1)}(b_{p-1}), \end{aligned} \tag{9.11}$$

   hold for almost all t in $\mathbb{R}_{\geq 0}$.

2. The test matrix Q associated to (9.11) is irreducible as a function and satisfies the linear small-gain-like condition (9.9).

Then we have

$$\liminf_{t \to \infty} \sum_{i=1}^{p} b_i(t) = 0. \tag{9.12}$$

If the functions b_i are uniformly continuous then we have

$$\lim_{t \to \infty} \sum_{i=1}^{p} b_i(t) = 0. \tag{9.13}$$

To prove Lemma 21, we observe in the next statement that Cesàro's summability of an integral [238], *i.e.* convergence of the mean, implies the "lim inf" convergence.

Lemma 22. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a continuous function. If

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \sigma(s)ds = 0,$$

then

$$\liminf_{t\to\infty} |\sigma(t)| = 0.$$

Remark 74. As it will be clear from the proof, if σ has constant sign then it is sufficient that σ be piecewise continuous. ■

Proof. We prove Lemma 22 by contradiction. Assume the existence of positive real numbers ε and T such that, for all $t \geq T$, $|\sigma(t)| \geq \varepsilon$. Because of continuity, σ has constant sign for all $t \geq T$. We study the case $\sigma(t) > 0$. For $\sigma(t) < 0$ similar arguments can be invoked. Integrating the previous inequality on the interval $[T, t]$ yields

$$\int_T^t \sigma(s)ds \geq \varepsilon(t - T).$$

Dividing by t, *i.e.*

$$\frac{1}{t} \int_T^t \sigma(s)ds \geq \frac{\varepsilon(t - T)}{t},$$

re-writing the integral as

$$\frac{1}{t} \int_0^t \sigma(s)ds \geq \frac{\varepsilon(t - T)}{t} + \frac{1}{t} \int_0^T \sigma(s)ds,$$

and picking t such that

$$\left| \int_0^T \sigma(s)ds \right| \leq \frac{\varepsilon}{2} t,$$

yield

$$\frac{1}{t} \int_0^t \sigma(s)ds \geq \frac{\varepsilon}{2} - \frac{\varepsilon T}{t}.$$

Since the limit for t going to $+\infty$ of the right hand side is strictly positive, we have a contradiction. □

Proof of Lemma 21. By Lemmas 19 and 20 we know the γ_{ij} defined in (9.8) exist and

the inequality (9.10) holds. Hence the i-th line in (9.11) gives

$$\dot{a}_i \leq -\alpha_i(b_i) + \sum_{\substack{j=1 \\ j\neq i}}^{p} \gamma_{ij}\alpha_j(b_j). \tag{9.14}$$

To rewrite this inequality in more compact notation let

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_p \end{bmatrix}, \qquad \boldsymbol{\alpha}(\boldsymbol{b}) = \begin{bmatrix} \alpha_1(b_1) \\ \vdots \\ \alpha_p(b_p) \end{bmatrix}.$$

Then (9.14) reads

$$\dot{a}_i \leq \begin{bmatrix} \gamma_{i1} & \dots & \gamma_{i(i-1)} & -1 & \gamma_{i(i+1)} & \dots & \gamma_{ip} \end{bmatrix} \boldsymbol{\alpha}(\boldsymbol{b}). \tag{9.15}$$

With the definition of the matrix Γ, this reduces further to

$$[\dot{\boldsymbol{a}}]_i \leq [-\Gamma\boldsymbol{\alpha}]_i. \tag{9.16}$$

Since, by (9.10), Γ has all leading principal minors with strictly positive determinant, by Lemma 18, Γ^{-1} has all positive entries, hence the relation

$$[\Gamma^{-1}\dot{\boldsymbol{a}}]_i \leq [-\boldsymbol{\alpha}]_i, \tag{9.17}$$

holds. In fact each of the inequalities in (9.17) is obtained as a weighted sum, with positive weights, of the inequalities in (9.16). Integrating from 0 to t each of these relations yields

$$\int_0^t [\boldsymbol{\alpha}(\boldsymbol{b}(s))]_i ds \leq -\int_0^t [\Gamma^{-1}\dot{\boldsymbol{a}}(s)]_i ds \leq [\Gamma^{-1}(\boldsymbol{a}(t) - \boldsymbol{a}(0))]_i.$$

Since the functions a_i are bounded, there exists a positive real number $\overline{\alpha}$ such that, for all i,

$$\int_0^t [\boldsymbol{\alpha}(\boldsymbol{b}(s))]_i ds \leq \overline{\alpha}.$$

By adding all the above inequalities we have that

$$\int_0^t \sum_{i=1}^{p} \alpha_i(b_i(s))ds < p\overline{\alpha} < +\infty, \tag{9.18}$$

hence, by Lemma 22, we conclude

$$\liminf_{t\to\infty} \sum_{i=1}^{p} \alpha_i(b_i(t)) = 0.$$

Since the functions α_i are positive definite, this implies (9.12). When the functions b_i are also uniformly continuous, the functions $t \mapsto \alpha_i(b_i(t))$ are uniformly continuous. So in this case, by Barbalat's Lemma, (9.18) gives

$$\lim_{t\to\infty} \sum_{i=1}^{p} \alpha_i(b_i(t)) = 0,$$

and therefore (9.13) follows. □

9.3.2 Triangular reducible case

Lemma 23. Let $p \geq 3$, $i \in \{1, \ldots, p\}$ and $j \in \{2, \ldots, p\}$. Let $a_i : \mathbb{R}_{\geq 0} \to [-\overline{a}, \overline{a}]$ be absolutely continuous functions and $b_i : \mathbb{R}_{\geq 0} \to [0, \overline{b}]$ be continuous functions. Consider continuous, positive definite functions $\alpha_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and continuous functions $\beta_{ij} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, with $j < i$, satisfying $\beta_{ij}(0) = 0$, such that the differential inequalities

$$\begin{aligned}
\dot{a}_1 &\leq -\alpha_1(b_1), \\
\dot{a}_2 &\leq -\alpha_2(b_2) + \beta_{21}(b_1), \\
&\vdots \\
\dot{a}_p &\leq -\alpha_p(b_p) + \beta_{p1}(b_1) + \cdots \beta_{p(p-1)}(b_{p-1}),
\end{aligned}$$

hold for almost all t in $\mathbb{R}_{\geq 0}$.

Then

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \sum_{i=1}^{p} b_i(s)ds = 0, \tag{9.19}$$

and therefore

$$\liminf_{t\to\infty} \sum_{i=1}^{p} b_i(t) = 0. \tag{9.20}$$

Remark 75. As opposed to the irreducible case given in Lemma 21, in the triangular reducible case boundedness of the functions $s \mapsto \dfrac{\beta_{ij}(s)}{\alpha_j(s)}$ does not play any role. ■

To prove Lemma 23 we use the following sufficient condition to have Cesàro's summability.

Lemma 24. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a locally integrable function. If, for all $\varepsilon > 0$, there exits a positive number μ such that

$$\left| \int_0^t \sigma(s)ds \right| \leq \varepsilon t + \mu, \tag{9.21}$$

for all $t \geq 0$, then

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \sigma(s)ds = 0.$$

Proof. Assume, by contradiction, that there exists $\delta > 0$ and a sequence $t_n \to \infty$ such that

$$\left| \int_0^{t_n} \sigma(s)ds \right| \geq \delta t_n.$$

Select $\varepsilon = \dfrac{\delta}{2}$ then, by (9.21), there exists μ such that

$$\left| \int_0^t \sigma(s)ds \right| \leq \frac{\delta}{2} t + \mu,$$

for all $t \geq 0$. For $t = t_n$, this gives

$$\delta t_n \leq \left| \int_0^{t_n} \sigma(s)ds \right| \leq \frac{\delta}{2} t_n + \mu.$$

Since there exists n such that $t_n > \dfrac{2\mu}{\delta}$, we have a contradiction. □

Remark 76. Lemma 22 and 24 provide a weaker version of Barbalat's Lemma (see *e.g.* [1]). In fact, the classical Barbalat's Lemma can be recovered when the function σ is uniformly continuous and (9.21) holds for $\epsilon = 0$. ■

Another notion that we need to introduce concerns a function φ associated with a pair of functions (α, β).

Let $\bar{b}$ be a non-negative real number. To a continuous positive definite function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and a continuous function $\beta : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, satisfying $\beta(0) = 0$, we associate the

function $\varphi : [1, +\infty) \to \mathbb{R}$ defined as

$$\varphi(\rho) = \max_{b \in [0, \bar{b}]} \left(\beta(b) - (\rho - 1)\alpha(b)\right). \tag{9.22}$$

Lemma 25. The function φ takes non-negative values and is non-increasing, Lipschitz and such that $\lim_{\rho \to +\infty} \varphi(\rho) = 0$.

Proof. Since α and β are continuous and $[0, \bar{b}\,]$ is compact, for each $\rho \in [1, +\infty)$ there exists (at least one) $b(\rho)$ in $[0, \bar{b}\,]$ such that

$$\varphi(\rho) = \beta(b(\rho)) - (\rho - 1)\alpha(b(\rho)).$$

As a result, for any $\rho' \geq \rho'' \geq 1$,

$$\begin{aligned} \varphi(\rho'') &= \beta(b(\rho'')) - (\rho'' - 1)\alpha(b(\rho'')) \geq \beta(b(\rho')) - (\rho'' - 1)\alpha(b(\rho')) \\ &\geq \beta(b(\rho')) - (\rho' - 1)\alpha(b(\rho')) = \varphi(\rho') \geq \beta(b(\rho'')) - (\rho' - 1)\alpha(b(\rho'')). \end{aligned}$$

This yields

$$0 \leq \varphi(\rho'') - \varphi(\rho') \leq (\rho' - \rho'')\,\overline{\alpha},$$

where

$$\overline{\alpha} = \max_{b \in [0, \bar{b}\,]} \alpha(b),$$

i.e. the function φ is Lipschitz and non-increasing.

Note now that, since α is continuous and positive definite, for any sequence $\{\rho_n\}$, such that $\lim_{n \to +\infty} \rho_n = +\infty$, there exists $N > 0$ and a sequence $\{b_n\} \subset [0, \bar{b}]$, satisfying $\lim_{n \to +\infty} b_n = 0$, and $\alpha(b_n) = \dfrac{1}{n\rho_n}$, for all $n \geq N$. In addition, since $b(\rho_n) \in [0, \bar{b}]$,

$$\beta(b_n) + \alpha(b_n) - \frac{1}{n} \leq \varphi(\rho_n) = \beta(b(\rho_n)) - (\rho_n - 1)\alpha(b(\rho_n)),$$

and therefore

$$0 \leq \rho_n\, \alpha(b(\rho_n)) + \beta(b_n) + \alpha(b_n) \leq \beta(b(\rho_n)) + \alpha(b(\rho_n)) + \frac{1}{n}.$$

This implies that $\lim_{n \to \infty} \alpha(b(\rho_n)) = 0$ and, since α is continuous and positive definite, that

$\lim_{n\to\infty} b(\rho_n) = 0$. Finally, since β is zero at zero and continuous,

$$\lim_{\rho\to+\infty} \varphi(\rho) = 0,$$

which also proves that φ takes non-negative values. □

Note that Lemma 25 holds also for a linear combination of functions $\beta(b)$. In this case we use the notation

$$\varphi_j(\rho) = \max_{b_j\in[0,\bar{b}]} \left(\beta_j(b_j) - (\rho-1)\alpha_j(b_j)\right), \tag{9.23}$$

with $\beta_j(b_j) = \sum_{\substack{i=1\\ i\neq j}}^{p} k_i\beta_{ij}(b_j)$, where the weights k_i are non-negative.

We are now ready to prove Lemma 23.

Proof of Lemma 23. The claim is proved by contradiction. To simplify the discussion consider the case $p = 3$, which contains all ingredients necessary for the general proof. Let $\varphi_2 : [1, +\infty) \to \mathbb{R}$ be defined as

$$\varphi_2(\rho) = \max_{b_2\in[0,\bar{b}]} \left(\beta_{32}(b_2) - (\rho-1)\alpha_2(b_2)\right).$$

Let also ε be an arbitrarily chosen strictly positive real number. Since by Lemma 25, φ_2 is non-increasing and $\lim_{\rho\to+\infty} \varphi_2(\rho) = 0$, we can select $\psi_2(\varepsilon)$ in $[1, +\infty)$ such that

$$\varphi_2(\psi_2(\varepsilon)) \leq \frac{\varepsilon}{2}.$$

Let $\varphi_1 : [1, +\infty) \to \mathbb{R}$ be defined as

$$\varphi_1(\rho) = \max_{b_2\in[0,\bar{b}]} \left(\psi_2(\varepsilon)\beta_{21}(b_1) + \beta_{31}(b_1) - (\rho-1)\alpha_1(b_1)\right).$$

Similarly, for all $\varepsilon > 0$ we can select $\psi_1(\varepsilon)$ in $[1, +\infty)$ such that

$$\varphi_1(\psi_1(\varepsilon)) \leq \frac{\varepsilon}{2}.$$

Note now that

$$\begin{aligned}\dot{\overbrace{\psi_1(\varepsilon)a_1+\psi_2(\varepsilon)a_2+a_3}} &\leq -\psi_1(\varepsilon)\alpha_1(b_1)+\psi_2(\varepsilon)\beta_{21}(b_1)+\beta_{31}(b_1)-\psi_2(\varepsilon)\alpha_2(b_2)+\beta_{32}(b_2)-\alpha_3(b_3)\\ &\leq (\varphi_1(\psi_1(\varepsilon))+\varphi_2(\psi_2(\varepsilon)))-(\alpha_1(b_1)+\alpha_2(b_2)+\alpha_3(b_3))\,.\end{aligned}$$

As a result

$$\begin{aligned}\int_0^t \sum_{i=1}^{3}\alpha_i(b_i(s))ds \leq\ & (\varphi_1(\psi_1(\varepsilon))+\varphi_2(\psi_2(\varepsilon)))t+(\psi_1(\varepsilon)a_1(0)+\psi_2(\varepsilon)a_2(0)+a_3(0))\\ & -(\psi_1(\varepsilon)a_1(t)+\psi_2(\varepsilon)a_2(t)+a_3(t))\\ \leq\ & (\varphi_1(\psi_1(\varepsilon))+\varphi_2(\psi_2(\varepsilon)))t+2(\psi_1(\varepsilon)+\psi_2(\varepsilon)+1)\overline{a}\\ \leq\ & \varepsilon\, t\ +\ 2(\psi_1(\varepsilon)+\psi_2(\varepsilon)+1)\overline{a}.\end{aligned}$$

Since ε is arbitrary, the claim follows by Lemmas 22 and 24.
In the case $p > 3$ the claim can be proved along the same lines defining $p-1$ functions φ_j. □

Remark 77. If in Lemma 23 we assume that the functions $\dfrac{\beta_{ij}}{\alpha_j}$ are bounded and the functions b_i are uniformly continuous, then we have the "lim" convergence result

$$\lim_{t\to\infty}\sum_{i=1}^{p} b_i(t) = 0.$$

In fact in the previous proof we can pick

$$\varepsilon = 0, \quad \psi_2 = 1+\gamma_{21}, \quad \psi_1 = 1+\gamma_{31}+\gamma_{32}\psi_2$$

and follow the same arguments as in the proof of Lemma 21. ■

9.3.3 Triangular block reducible case

We are now ready to study the triangular block reducible case that can be regarded as a generalization of the previous results. To this end, let $s_l = r_l - r_{l-1}$, with $s_0 = r_0 = 0$, be the dimension of the column vectors

$$\boldsymbol{a}_l = \begin{bmatrix} a_{(r_{l-1}+1)} & a_{(r_{l-1}+2)} & \dots & a_{r_l}\end{bmatrix}^T, \qquad \boldsymbol{b}_l = \begin{bmatrix} b_{(r_{l-1}+1)} & b_{(r_{l-1}+2)} & \dots & b_{r_l}\end{bmatrix}^T,$$

and define

$$\boldsymbol{\alpha}_l(\boldsymbol{b}_l) = \begin{bmatrix} \alpha_{(r_{l-1}+1)}(b_{(r_{l-1}+1)}) & \alpha_{(r_{l-1}+2)}(b_{(r_{l-1}+2)}) & \dots & \alpha_{r_l}(b_{r_l}) \end{bmatrix}^T,$$

$$\boldsymbol{\delta}_l(\boldsymbol{b}_l) = \sum_{j=(r_{l-1}+1)}^{r_l} \begin{bmatrix} \delta_{(r_{l-1}+1)j}(b_j) \\ \delta_{(r_{l-1}+2)j}(b_j) \\ \vdots \\ \delta_{(r_l)j}(b_j) \end{bmatrix}, \text{ with } \delta_{kj} = \begin{cases} -\alpha_j \text{ if } k = j \\ \\ \beta_{kj} \text{ if } k \neq j \end{cases}$$

$$\boldsymbol{\mu}_{lm}(\boldsymbol{b}_m) = \sum_{j=(r_{m-1}+1)}^{r_m} \boldsymbol{\mu}_{lm}(b_j) = \sum_{j=(r_{m-1}+1)}^{r_m} \begin{bmatrix} \beta_{(r_{l-1}+1)j}(b_j) \\ \beta_{(r_{l-1}+2)j}(b_j) \\ \vdots \\ \beta_{(r_l)j}(b_j) \end{bmatrix}.$$

Proposition 12. Let $i \in \{1, \dots, p\}$ and $j \in \{1, \dots, p\}$. Let $a_i : \mathbb{R}_{\geq 0} \to [-\overline{a}, \overline{a}]$ be absolutely continuous functions and $b_i : \mathbb{R}_{\geq 0} \to [0, \overline{b}]$ be continuous functions. Consider continuous, positive definite functions $\alpha_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and continuous functions $\beta_{ij} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, with $i \neq j$, satisfying $\beta_{ij}(0) = 0$. Let $\boldsymbol{a}_l$, $\boldsymbol{b}_l$, $\boldsymbol{\delta}_l$ and $\boldsymbol{\mu}_{lm}$ be vectors of dimension s_l, with components obtained from the a_i's, b_i's, α_i's and β_{ij}'s, such that the following hold.

1. The differential inequalities

$$\begin{aligned} \dot{\boldsymbol{a}}_1 &\leq \boldsymbol{\delta}_1(\boldsymbol{b}_1), \\ \dot{\boldsymbol{a}}_2 &\leq \boldsymbol{\mu}_{21}(\boldsymbol{b}_1) + \boldsymbol{\delta}_2(\boldsymbol{b}_2), \\ &\vdots \\ \dot{\boldsymbol{a}}_q &\leq \boldsymbol{\mu}_{q1}(\boldsymbol{b}_1) + \boldsymbol{\mu}_{q2}(\boldsymbol{b}_2) + \dots + \boldsymbol{\delta}_q(\boldsymbol{b}_q), \end{aligned} \tag{9.24}$$

   with $r_q = p$, hold for almost all t in $\mathbb{R}_{\geq 0}$.

2. The matrix Q_l for each diagonal element $\boldsymbol{\delta}_l$ is irreducible as a function and satisfies the linear small-gain-like condition (9.9).

Then

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t \sum_{i=1}^{p} b_i(s) ds = 0, \tag{9.25}$$

and therefore

$$\liminf_{t\to\infty} \sum_{i=1}^{p} b_i(t) = 0. \tag{9.26}$$

Remark 78. As expected, (9.26) holds with no additional restrictions on the off-diagonal elements $\boldsymbol{\mu}_{lm}$. However, as discussed in Remark 77, if all functions $\dfrac{\beta_{ij}}{\alpha_j}$ in the off-diagonal element $\boldsymbol{\mu}_{lm}$ are bounded and the b_i are uniformly continuous then (9.26) can be replaced by

$$\lim_{t\to\infty} \sum_{i=1}^{p} b_i(t) = 0. \tag{9.27}$$

■

Proof. We consider the case with $q = 3$ blocks, namely

$$\begin{aligned}
&\dot{\boldsymbol{a}}_1 \leq \boldsymbol{\delta}_1(\boldsymbol{b}_1),\\
&\dot{\boldsymbol{a}}_2 \leq \boldsymbol{\mu}_{21}(\boldsymbol{b}_1) + \boldsymbol{\delta}_2(\boldsymbol{b}_2),\\
&\dot{\boldsymbol{a}}_3 \leq \boldsymbol{\mu}_{31}(\boldsymbol{b}_1) + \boldsymbol{\mu}_{32}(\boldsymbol{b}_2) + \boldsymbol{\delta}_3(\boldsymbol{b}_3),
\end{aligned}$$

which contains all the ingredients necessary for the general proof. Define Γ_l as the matrix corresponding to the test matrix Q_l attached to the vector $\boldsymbol{\delta}_l$ and $\mathbf{1}_l$ as the row vector with s_l elements equal to 1. Let also ε be an arbitrarily chosen strictly positive real number. In a way similar to the one followed to get (9.16), we obtain

$$\boldsymbol{\delta}_l(\boldsymbol{b}_l) \;\leq\; -\Gamma_l \boldsymbol{\alpha}(\boldsymbol{b}_l)$$

and therefore $\Gamma_l^{-1}\boldsymbol{\delta}_l(\boldsymbol{b}_l) \leq -\boldsymbol{\alpha}(\boldsymbol{b}_l)$. This leads to

$$\begin{aligned}
&\Gamma_1^{-1}\dot{\boldsymbol{a}}_1 \leq -\boldsymbol{\alpha}_1(\boldsymbol{b}_1),\\
&\Gamma_2^{-1}\dot{\boldsymbol{a}}_2 \leq \Gamma_2^{-1}\boldsymbol{\mu}_{21}(\boldsymbol{b}_1) - \boldsymbol{\alpha}_2(\boldsymbol{b}_2),\\
&\Gamma_3^{-1}\dot{\boldsymbol{a}}_3 \leq \Gamma_3^{-1}\boldsymbol{\mu}_{31}(\boldsymbol{b}_1) + \Gamma_3^{-1}\boldsymbol{\mu}_{32}(\boldsymbol{b}_2) - \boldsymbol{\alpha}_3(\boldsymbol{b}_3).
\end{aligned}$$

To deal with the terms in $\boldsymbol{b}_2$, we define s_2 functions $\varphi_i : [1, +\infty) \to \mathbb{R}$ as

$$\begin{aligned}
\varphi_{r_1+1}(\rho) &= \max_{b_{r_1+1}\in[0,\bar{b}]} (\mathbf{1}_3\Gamma_3^{-1}\boldsymbol{\mu}_{32}(b_{r_1+1}) - (\rho-1)\alpha_{r_1+1}(b_{r_1+1})),\\
&\vdots\\
\varphi_{r_2}(\rho) &= \max_{b_{r_2}\in[0,\bar{b}]} (\mathbf{1}_3\Gamma_3^{-1}\boldsymbol{\mu}_{32}(b_{r_2}) - (\rho-1)\alpha_{r_2}(b_{r_2})).
\end{aligned} \tag{9.28}$$

By Lemma 25, the functions φ_i are non-increasing and $\lim_{\rho\to+\infty}\varphi_i(\rho)=0$. As a result, we can select a vector $\boldsymbol{\psi}_2(\varepsilon)$ of size s_2 with components $\psi_i(\varepsilon)$ in $[1,+\infty)$ such that

$$\sum_{i=r_1+1}^{r_2}\varphi_i(\psi_i(\varepsilon))\leq\frac{\varepsilon}{2}.$$

This gives

$$\mathbf{1}_3\Gamma_3^{-1}\boldsymbol{\mu}_{32}(\boldsymbol{b}_2)-\boldsymbol{\psi}_2(\varepsilon)\boldsymbol{\alpha}_2(\boldsymbol{b}_2)\leq\frac{\varepsilon}{2}-\mathbf{1}_2\boldsymbol{\alpha}_2(\boldsymbol{b}_2).$$

Similarly, to deal with the terms in $\boldsymbol{b}_1$, we define s_1 functions $\varphi_i:[1,+\infty)\to\mathbb{R}$ as

$$\begin{aligned}\varphi_1(\rho) &= \max_{b_1\in[0,\bar{b}]}(\boldsymbol{\psi}_2(\varepsilon)\Gamma_2^{-1}\boldsymbol{\mu}_{21}(b_1)+\mathbf{1}_3\Gamma_3^{-1}\boldsymbol{\mu}_{31}(b_1)-(\rho-1)\alpha_1(b_1)),\\ &\vdots\\ \varphi_{r_1}(\rho) &= \max_{b_{r_1}\in[0,\bar{b}]}(\boldsymbol{\psi}_2(\varepsilon)\Gamma_2^{-1}\boldsymbol{\mu}_{21}(b_{r_1})+\mathbf{1}_3\Gamma_3^{-1}\boldsymbol{\mu}_{31}(b_{r_1})-(\rho-1)\alpha_{r_1}(b_{r_1})).\end{aligned}\tag{9.29}$$

Again, by Lemma 25, the functions φ_i are non-increasing and $\lim_{\rho\to+\infty}\varphi_i(\rho)=0$. So we can select a vector $\boldsymbol{\psi}_1(\varepsilon)$ of size s_1 with components $\psi_i(\varepsilon)$ in $[1,+\infty)$ such that

$$\sum_{i=1}^{r_1}\varphi_i(\psi_i(\varepsilon))\leq\frac{\varepsilon}{2}.$$

This gives

$$\boldsymbol{\psi}_2(\varepsilon)\Gamma_2^{-1}\boldsymbol{\mu}_{21}(\boldsymbol{b}_1)+\mathbf{1}_3\Gamma_3^{-1}\boldsymbol{\mu}_{31}(\boldsymbol{b}_1)-\boldsymbol{\psi}_1(\varepsilon)\boldsymbol{\alpha}_1(\boldsymbol{b}_1)\leq\frac{\varepsilon}{2}-\mathbf{1}_1\boldsymbol{\alpha}_1(\boldsymbol{b}_1).$$

So we have obtained

$$\mathbf{1}_3\Gamma_3^{-1}\dot{\boldsymbol{a}}_3+\boldsymbol{\psi}_2(\varepsilon)\Gamma_2^{-1}\dot{\boldsymbol{a}}_2+\boldsymbol{\psi}_1(\varepsilon)\Gamma_1^{-1}\dot{\boldsymbol{a}}_1\leq\varepsilon-\mathbf{1}_3\boldsymbol{\alpha}_3(\boldsymbol{b}_3)-\mathbf{1}_2\boldsymbol{\alpha}_2(\boldsymbol{b}_2)-\mathbf{1}_1\boldsymbol{\alpha}_1(\boldsymbol{b}_1)\leq\varepsilon-\sum_{i=1}^{r_3}\alpha_i.\tag{9.30}$$

The claim follows by integrating both sides of (9.30) from 0 to t and applying Lemmas 22 and 24. □

9.4 On the linear small-gain-like condition

In this section we discuss the linear small-gain condition and explain why it is necessary to use this in the assumptions of Proposition 12 instead of the nonlinear condition. The

discussion, for simplicity, is limited to the case $p = 2$, in which (9.9) yields

$$\beta_{12}(b_2)\beta_{21}(b_1) \leq (1-\varepsilon)\alpha_2(b_2)\alpha_1(b_1), \quad \forall (b_1, b_2) \in [0, \bar{b}]^2. \tag{9.31}$$

To simplify the discussion we restrict ourselves to consider the case in which the functions α_i are invertible and the above inequality (9.31) holds for all non-negative real numbers b_1 and b_2. From the theory of interconnected nonlinear systems we would expect that stability properties be related to the nonlinear small-gain condition

$$\beta_{21} \circ \alpha_1^{-1} \circ \beta_{12} \circ \alpha_2^{-1}(s) \; < \; s, \qquad \forall s > 0. \tag{9.32}$$

Lemma 26. If β_{12} (β_{21} respectively) is positive definite, condition (9.31) implies, but it is not implied by, condition (9.32).

Proof. We first show that the linear condition implies the nonlinear one. Pick any pair (b_1, b_2) in $]0, \bar{b}]^2$ and note that the linear condition (9.31) yields

$$\left[\beta_{12} \circ \alpha_2^{-1}(b_2)\right] \left[\beta_{21} \circ \alpha_1^{-1}(b_1)\right] \; \leq \; (1-\varepsilon)\, b_1\, b_2 \; .$$

In particular, the selection

$$s > 0, \qquad b_2 = s, \qquad b_1 = \beta_{12} \circ \alpha_2^{-1}(s),$$

yields

$$b_1 \; \beta_{21} \circ \alpha_1^{-1} \circ \beta_{12} \circ \alpha_2^{-1}(s) \; \leq \; (1-\varepsilon)\, b_1\, s,$$

which implies condition (9.32).

To show that the converse is not true, let $\alpha_1(s) = s$, $\beta_{12}(s) = s^2$, $\alpha_2(s) = s$ and $\beta_{21}(s) = \gamma\sqrt{s}$. The nonlinear small-gain reduces to $\gamma s \leq (1-\varepsilon)s$ which holds for all $0 \leq \gamma < 1$, whereas the linear condition reduces to $\gamma \dfrac{b_2}{\sqrt{b_1}} < (1-\varepsilon)$ which does not hold whatever the positive value of γ is. □

As usual for small-gain conditions it is difficult to establish the true necessity of (9.31). We now show that violation of the non-strict inequality yields the existence of functions a_i and b_i such that the convergence result of Lemma 21 does not hold.

Lemma 27. Assume there exist strictly positive real numbers b_{1a}, b_{2b} and b_{2c} such that

$$\frac{\beta_{12}(b_{2b})\beta_{21}(b_{1a})}{\alpha_2(b_{2b})\alpha_1(b_{1a})} > 1, \qquad \frac{\beta_{12}(b_{2c})\beta_{21}(b_{1a})}{\alpha_2(b_{2c})\alpha_1(b_{1a})} < 1. \tag{9.33}$$

Then there exist functions a_i and b_i such that the convergence result in Lemma 21 does not hold.

Remark 79. Condition (9.33) says that, with $b_1 = b_{1a}$, the inequality (9.31) holds for $b_2 = b_c$ but does not for $b_2 = b_{2b}$. ■

Proof. Assume for the time being that we can find strictly positive real numbers ε_1 and ε_2 such that there exist strictly positive real numbers T_b and T_c satisfying the linear equations

$$\begin{bmatrix} \beta_{12}(b_{2b}) - \alpha_1(\varepsilon_1) & \beta_{12}(b_{2c}) - \alpha_1(\varepsilon_1) \\ \alpha_2(b_{2b}) - \beta_{21}(\varepsilon_1) & \alpha_2(b_{2c}) - \beta_{21}(\varepsilon_1) \end{bmatrix} \begin{bmatrix} T_b \\ T_c \end{bmatrix} = - \begin{bmatrix} \beta_{12}(\varepsilon_2) - \alpha_1(b_{1a}) \\ \alpha_2(\varepsilon_2) - \beta_{21}(b_{1a}) \end{bmatrix}. \tag{9.34}$$

Then, let b_1 and b_2 be piecewise constant and $(1 + T_b + T_c)$-periodic functions defined as

$$b_1(t) = \begin{cases} b_{1a} & \text{if} \quad t \in [0, 1], \\ \varepsilon_1 & \text{if} \quad t \in]1, 1 + T_b + T_c[, \end{cases} \qquad b_2(t) = \begin{cases} \varepsilon_2 & \text{if} \quad t \in [0, 1], \\ b_{2b} & \text{if} \quad t \in]1, 1 + T_b], \\ b_{2c} & \text{if} \quad t \in]1 + T_b, 1 + T_b + T_c[, \end{cases}$$

As a result,

$$\begin{aligned} a_1(1 + T_b + T_c) - a_1(0) \leq & - [\alpha_1(b_{1a}) - \beta_{12}(\varepsilon_2)] + [-\alpha_1(\varepsilon_1) + \beta_{12}(b_{2b})] T_b \\ & + [-\alpha_1(\varepsilon_1) + \beta_{12}(b_{2c})] T_c = 0, \end{aligned}$$

and

$$\begin{aligned} a_2(1 + T_b + T_c) - a_2(0) \leq & - [\alpha_2(\varepsilon_2) - \beta_{21}(b_{1a})] + [-\alpha_2(b_{2b}) + \beta_{21}(\varepsilon_1)] T_b \\ & + [-\alpha_2(b_{2c}) + \beta_{21}(\varepsilon_1)] T_c = 0. \end{aligned}$$

Therefore the result holds with a_i any constant function.

Now to prove that T_b, T_c do exist we note that when ε_1 and ε_2 are both zero, the solution of the equations (9.34) is

$$T_b = \frac{\alpha_2(b_{2c})\alpha_1(b_{1a}) - \beta_{12}(b_{2c})\beta_{21}(b_{1a})}{\alpha_2(b_{2c})\beta_{12}(b_{2b}) - \alpha_2(b_{2b})\beta_{12}(b_{2c})}, \qquad T_c = \frac{\beta_{12}(b_{2b})\beta_{21}(b_{1a}) - \alpha_2(b_{2b})\alpha_1(b_{1a})}{\alpha_2(b_{2c})\beta_{12}(b_{2b}) - \alpha_2(b_{2b})\beta_{12}(b_{2c})}.$$

By condition (9.33) T_b and T_c are strictly positive if the denominator is strictly positive. This is indeed the case since, multiplying the inequalities in (9.33), yields

$$\beta_{12}(b_{2b})\beta_{21}(b_{1a})\alpha_2(b_{2c})\alpha_1(b_{1a}) > \alpha_2(b_{2b})\alpha_1(b_{1a})\beta_{12}(b_{2c})\beta_{21}(b_{1a}),$$

where $\alpha_1(b_{1a}) > 0$, since $b_{1a} > 0$ and $\beta_{12}(b_{2b}) > 0$ because of (9.33). Therefore, by continuity, T_b and T_c are strictly positive when ε_1 and ε_2 are strictly positive, and sufficiently small. □

Thus, (9.31) is necessary to guarantee that there do not exist functions b_1 and b_2 such that the convergence result of Lemma 21 does not hold.

9.5 "lim inf" asymptotic properties in dynamical systems

Proposition 12 can be applied to study asymptotic properties of the solutions of dynamical systems. In particular the following theorem solve Problem 1 and gives conditions to establish the "lim inf" or "lim" convergence of such solutions.

Theorem 30. Let $i \in \{1, \dots, p\}$ and $j \in \{1, \dots, p\}$. Consider system (9.4) and let $V_i : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable functions and $h_i : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ be continuous functions. Consider continuous, positive definite functions $\alpha_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, continuous functions $\beta_{ij} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, satisfying $\beta_{ij}(0) = 0$. Let $\boldsymbol{V}_l$, $\boldsymbol{h}_l$, $\boldsymbol{\delta}_l$ and $\boldsymbol{\mu}_{lm}$ be vectors of dimension s_l, with components obtained from the V_i's, h_i's, α_i's and β_{ij}'s[4], such that the following hold.

1. Along the solutions of system (9.4), we have

$$\begin{aligned} \dot{\boldsymbol{V}}_1(x) &\leq \boldsymbol{\delta}_1(\boldsymbol{h}_1(x)), \\ \dot{\boldsymbol{V}}_2(x) &\leq \boldsymbol{\mu}_{21}(\boldsymbol{h}_1(x)) + \boldsymbol{\delta}_2(\boldsymbol{h}_2(x)), \qquad \forall x \in \mathbb{R}^n, \\ &\vdots \\ \dot{\boldsymbol{V}}_q(x) &\leq \boldsymbol{\mu}_{q1}(\boldsymbol{h}_1(x)) + \boldsymbol{\mu}_{q2}(\boldsymbol{h}_2(x)) + \cdots + \boldsymbol{\delta}_q(\boldsymbol{h}_q(x)), \end{aligned} \tag{9.35}$$

[4] In particular

$$\begin{aligned} \boldsymbol{V}_l &= \begin{bmatrix} V_{(r_{l-1}+1)} & V_{(r_{l-1}+2)} & \dots & V_{r_l} \end{bmatrix}^T, \\ \boldsymbol{h}_l &= \begin{bmatrix} h_{(r_{l-1}+1)} & h_{(r_{l-1}+2)} & \dots & h_{r_l} \end{bmatrix}^T. \end{aligned}$$

with $r_q = p$.

2. The matrix Q_l for each diagonal element $\boldsymbol{\delta}_l$ is irreducible as a function and satisfies the linear small-gain-like condition (9.9).

Then, for any bounded solution $t \mapsto X(x,t)$ of (9.4),

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \sum_{i=1}^{p} h_i(X(x,s))ds = 0,$$

and therefore

$$\liminf_{t\to\infty} \sum_{i=1}^{p} h_i(X(x,t)) = 0. \tag{9.36}$$

Moreover,

3*a*. if all functions $\dfrac{\beta_{ij}}{\alpha_j}$ of all off-diagonal elements $\boldsymbol{\mu}_{lm}$ are bounded,

or

3*b*. if the largest invariant set $\mathcal{H}_p$ contained in the set

$$\Omega_{h_1,\dots,h_p} = \{x \in \mathbb{R}^n : h_1(x) = h_2(x) = \cdots = h_p(x) = 0\},$$

is stable,

then

$$\lim_{t\to\infty} \sum_{i=1}^{p} h_i(X(x,t)) = 0. \tag{9.37}$$

Proof. Property (9.36) follows directly from Proposition 12 with $h_i(X(x,t))$ playing the role of $b_i(t)$.

If 3*a*) holds, (9.37) follows directly from Remark 78 with $h_i(X(x,t))$ playing the role of $b_i(t)$. Note that the uniform continuity of $t \mapsto h_i(X(x,t))$ follows from the continuity of h_i, and the boundedness of the locally Lipschitz function $t \mapsto X(x,t)$ (since f is continuous).

If 3*b*) holds suppose that all the blocks have dimension one and $p = 3$. This contains all ingredients necessary for the general case.
Since V_1 is bounded and decreasing along all the trajectories of the system by assumption, the first inequality in (9.35), namely $\dot{V}_1 \leq -\alpha_1(h_1)$ implies that

$$\lim_{t\to\infty} h_1(X(x,t)) = 0.$$

Since the solution is bounded, $X(x,t)$ has an ω-limit set $\Omega(x)$ which is invariant and compact, the previous limit implies

$$\Omega(x) \subset \Omega_{h_1} = \{x \in \mathbb{R}^n : h_1(x) = 0\}.$$

For every $x_{\omega_1} \in \Omega(x)$, $h_1(x_{\omega_1}) = 0$ which, by the second inequality in (9.35), implies

$$\lim_{t\to\infty} h_2(X(x_{\omega_1}, t)) = 0,$$

and similarly to the previous discussion, this implies

$$\Omega(x_{\omega_1}) \subset \Omega_{h_1,h_2} = \{x \in \mathbb{R}^n : h_1(x) = h_2(x) = 0\}.$$

For every $x_{\omega_2} \in \Omega(x_{\omega_1})$, $h_1(x_{\omega_2}) = h_2(x_{\omega_2}) = 0$ which, by the third inequality in (9.35), implies

$$\lim_{t\to\infty} h_3(X(x_{\omega_2}, t)) = 0,$$

and again, this implies

$$\Omega(x_{\omega_2}) \subset \Omega_{h_1,h_2,h_3} = \{x \in \mathbb{R}^n : h_1(x) = h_2(x) = h_3(x) = 0\}.$$

This proves that, if the differential inequalities (9.35) are in triangular form and $\Omega_{h_1,\ldots,h_p}$ is stable, then (9.37) holds. Note that, if the first block of the differential inequalities (9.35) has dimension greater than one, then (9.37) follows directly from Lemma 21 applied to that block. The proof of the general triangular block case can be derived from this last fact and the discussion carried out for the triangular case.

Now we prove that it is sufficient that the largest invariant set $\mathcal{H}_p$ contained in $\Omega_{h_1,\ldots,h_p}$ is stable. Again, for simplicity, consider the case $p = 3$. Assume, by contradiction, that there exist x in $\mathbb{R}^n$, ε strictly positive and a sequence t_m going to infinity with m such that

$$d(X(x, t_m), \mathcal{H}_3) > \varepsilon.$$

Since $\mathcal{H}_3$ is stable there exists δ strictly positive such that, for any χ in $\mathbb{R}^n$ satisfying

$$d(\chi, \mathcal{H}_3) \leq \delta, \tag{9.38}$$

we have

$$d(X(\chi, s), \mathcal{H}_3) \leq \varepsilon \qquad \forall s \geq 0. \tag{9.39}$$

Then, since $\Omega(x)$ is a closed invariant set we have $\Omega(x_{\omega_1}) \subset \Omega(x)$ and since $\mathcal{H}_3$ is the largest invariant set contained in the set Ω_{h_1,h_2,h_3} one has $\Omega(x_{\omega_2}) \subset \mathcal{H}_3$. Now because of the convergence of $X(x_{\omega_2}, t)$ to its ω-limit set $\Omega(x_{\omega_2})$, hence there exists T_2 such that

$$d(X(x_{\omega_2}, T_2), \mathcal{H}_3) \leq d(X(x_{\omega_2}, T_2), \Omega(x_{\omega_2})) \leq \frac{\delta}{2}$$

and

$$X(x_{\omega_2}, T_2) \in \Omega(x_{\omega_1}) \subset \Omega(x).$$

This means that $X(x_{\omega_2}, T_2)$ is an ω-limit point of $X(x, t)$, there exists T such that

$$|X(x, T) - X(x_{\omega_2}, T_2)| \leq \frac{\delta}{2}.$$

As a result the triangular inequality yields

$$d(X(x, T), \mathcal{H}_3) \leq |X(x, T) - X(x_{\omega_2}, T_2)| + d(X(x_{\omega_2}, T_2), \mathcal{H}_3) \leq \delta.$$

Therefore $\chi = X(x, T)$ satisfies (9.38), which by (9.39) yields a contradiction. □

Remark 80. The fact that (9.37) is implied by the stability of $\Omega_{h_1,\ldots,h_p}$ is a restatement of a well-known result, see for instance [230, Lemma I.4]. The fact that (9.37) is implied by the stability of the largest invariant set $\mathcal{H}_p$ contained in $\Omega_{h_1,\ldots,h_p}$ is a new result. ■

Remark 81. If $3b$) holds then (9.37) implies that $\mathcal{H}_p$ is asymptotically stable. ■

9.5.1 Example: Duffing oscillator

We present an elementary example which gives a simple illustration of how the results of the chapter can be used. Note that the convergence properties we obtain could be established with classical tools. A more involved example follows.

Consider the 2-dimensional system describing the Duffing oscillator, namely

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= \alpha x_1 - \beta x_2 - \gamma x_1^3, \end{aligned} \tag{9.40}$$

with $(x_1,\, x_2) \in \mathbb{R}^2$, $\alpha > 0$, $\beta > 0$ and $\gamma > 0$. The equilibrium points are $(x_1, x_2) = (0, 0)$, $(x_1, x_2) = \left(\pm\sqrt{\dfrac{\alpha}{\gamma}}, 0\right)$.

Let

$$V_1(x_1, x_2) = \frac{1}{\beta}\left(\gamma\frac{x_1^4}{4} - \alpha\frac{x_1^2}{2} + \frac{x_2^2}{2}\right),$$

and

$$V_2(x_1, x_2, x_3) = -(\alpha x_1 - \gamma x_1^3)x_2.$$

Then

$$\begin{aligned} \dot{V}_1 &= -x_2^2, \\ \dot{V}_2 &= -(\alpha - 3\gamma x_1^2)x_2^2 - (\alpha x_1 - \gamma x_1^3)^2 + \beta(\alpha x_1 - \gamma x_1^3)x_2. \end{aligned} \tag{9.41}$$

Since V_1 is radially unbounded, the first equality in (9.41) implies that all trajectories are bounded. Then, selecting

$$c \geq \sup_t |X_1(x_1, x_2, t)|,$$

(9.41) yields

$$\dot{V}_2 \leq -\frac{1}{2}(\alpha x_1 - \gamma x_1^3)^2 + \left(3\gamma c^2 - \alpha + \frac{\beta^2}{2}\right)x_2^2,$$

which motivates the choice

$$\begin{aligned} &\alpha_1(s) = s, \quad \beta_{12}(s) = 0, && b_1 = x_2^2, \\ &\alpha_2(s) = \frac{1}{2}s, \quad \beta_{21}(s) = \left(3\gamma c^2 - \alpha + \frac{\beta^2}{2}\right)s, && b_2 = \left(\alpha x_1 - \gamma x_1^3\right)^2. \end{aligned}$$

Note that we have a triangular structure and that

$$\sup_{s>0} \frac{\beta_{21}(s)}{\alpha_1(s)} = \frac{6\gamma c^2 - 2\alpha + \beta^2}{2},$$

is finite, hence Theorem 30 yields

$$\lim_{t\to+\infty} X_2(x_1, x_2, t)^2 + \left(\alpha X_1(x_1, x_2, t) - \gamma X_1(x_1, x_2, t)^3\right)^2 = 0,$$

which implies that the solutions of the system are converging to at least one equilibrium point.

9.5.2 Example: a more rich dynamics

Consider the class of systems described by the differential equations

$$\begin{aligned} \dot{x}_1 &= \eta\, x_1(1-x_1^2-x_2^2)x_3^r + k_1 x_2[\Psi(x_{1+}) + x_3^p], \\ \dot{x}_2 &= \eta\, x_2(1-x_1^2-x_2^2)x_3^r - k_1 x_1[\Psi(x_{1+}) + x_3^p], \\ \dot{x}_3 &= -k_2 x_3^q, \end{aligned} \tag{9.42}$$

where $(x_1,\, x_2,\, x_3) \in \mathbb{R}^3$, $x_{1+} = \max\{x_1, 0\}$, Ψ is a positive definite function, p and $r \geq p$ are positive even integers, q is a positive odd integer, k_1, k_2 are positive and $\eta \geq 0$. The set of equilibrium points is given by $\{(x_1, x_2, x_3) : x_{1+} = x_3 = 0\}$. Note that

$$\dot{\overbrace{x_1^2 + x_2^2 + x_3^2}} \;=\; -\eta(x_1^2+x_2^2)(x_1^2+x_2^2-1)x_3^r - k_2 x_3^{q+1}.$$

This shows that all solutions are bounded. Let

$$V_1(x_1, x_2, x_3) \;=\; \frac{x_3^2}{2k_2}$$

and

$$V_2(x_1, x_2, x_3) \;=\; \frac{x_2}{k_1}.$$

Then

$$\begin{aligned} \dot{V}_1 &= -x_3^{q+1}, \\ \dot{V}_2 &= \frac{\eta}{k_1}\, x_2(1-x_1^2-x_2^2)x_3^r - x_1[\Psi(x_{1+}) + x_3^p] \\ &= -x_1\Psi(x_{1+}) + x_3^p[-x_1 + \frac{\eta}{k_1} x_2(1-x_1^2-x_2^2)x_3^{r-p}] \\ &\leq -x_{1+}\Psi(x_{1+}) + x_3^p[x_{1-} + \frac{\eta}{k_1}(x_2(1-x_1^2-x_2^2))_+ x_3^{r-p}], \end{aligned} \tag{9.43}$$

where $x_{1-} = -\min\{x_1, 0\}$. These inequalities motivate the choice

$$\boldsymbol{\delta}_1(s) = \alpha_1(s) = s^{\frac{q+1}{2}}, \quad \boldsymbol{\delta}_2(s) = \alpha_2(s) = s, \qquad \boldsymbol{\mu}_{21}(s) = \beta_{21}(s) = cs^{\frac{p}{2}},$$

$$\boldsymbol{h}_1(x) = b_1 = x_3^2, \qquad \boldsymbol{h}_2(x) = b_2 = x_{1+}\Psi(x_{1+}),$$

where

$$c \geq x_{1-} + \frac{\eta}{k_1}(x_2(1-x_1^2-x_2^2))_+ x_3^{r-p}.$$

Note that we have a triangular structure and that, for any strictly positive $\bar{b}$, we have

$$\sup_{b\in(0,\bar{b}]} \frac{\beta_{21}(b)}{\alpha_1(b)} = \sup_{b\in(0,\bar{b}]} \frac{c}{b^{\frac{q-p+1}{2}}},$$

which is finite for $p \geq q+1$ and infinite otherwise. By Theorem 30

$$\lim_{t\to+\infty} \frac{1}{t}\int_0^t [X_3(x_1,x_2,x_3,t)^2 + X_{1+}(x_1,x_2,x_3,t)\Psi(X_{1+}(x_1,x_2,x_3,t))] = 0.$$

Since

$$\begin{aligned}
&\lim_{t\to+\infty} \int_0^t X_3(x_1,x_2,x_3,t)^{q+1}(\tau)d\tau \ < +\infty,\\
&\lim_{t\to+\infty} X_3(x_1,x_2,x_3,t) = 0,
\end{aligned}$$

then

$$\begin{aligned}
&\lim_{t\to+\infty} X_{1+}(x_1,x_2,x_3,t)\Psi(X_{1+}(x_1,x_2,x_3,t)) = 0, \quad \text{if } p \geq q+1,\\
&\liminf_{t\to+\infty} X_{1+}(x_1,x_2,x_3,t)\Psi(X_{1+}(x_1,x_2,x_3,t)) = 0, \quad \text{otherwise.}
\end{aligned} \tag{9.44}$$

In what follows we focus on the case $p = q-1$ and we show that the asymptotic property expressed by the second of equations (9.44) cannot be improved. To this end re-write the system using polar coordinates (θ,ρ) in the (x_1,x_2)-plane, *i.e.*

$$\begin{aligned}
\dot{\rho} &= \eta\rho(1-\rho^2)x_3^r,\\
\dot{\theta} &= -k_1(\Psi(\rho(\cos\theta)_+) + x_3^{q-1}),\\
\dot{x}_3 &= -k_2 x_3^q.
\end{aligned}$$

From the first equation, we obtain

$$\rho(t) = \frac{\rho(0)\exp(\eta\int_0^t x_3(s)^r ds)}{\sqrt{1-\rho(0)^2+\rho(0)^2\exp(2\eta\int_0^t x_3(s)^r ds)}}.$$

This implies

$$\begin{aligned}
&\eta = 0 && \Rightarrow \quad \rho(t) = \rho_0\\
&\eta > 0,\ x_3(t)^r \text{ integrable} && \Rightarrow \quad \min(1,\rho_0) \leq \rho(t) \leq \max(1,\rho_0)\\
&\eta > 0,\ x_3(t)^r \text{ not integrable} && \Rightarrow \quad \rho(t) \to 1.
\end{aligned}$$

From the second equation

$$\theta(t) \;\leq\; \theta(0) \;-\; k_1 \int_0^t x_3(s)^{q-1} ds,$$

and from the third equation

$$\frac{dx_3}{x_3} = -k_2 x_3^{q-1} dt,$$

and then

$$\frac{1}{k_2}\left(\log(x_3(t)) - \log(x_3(0))\right) = -\int_0^t x_3(s)^{q-1} ds,$$

that can be directly substituted in the right hand side of $\theta(t)$. As a result

$$\theta(t) \;\leq\; \left[\theta(0) - \frac{k_1}{k_2}\log(x_3(0))\right] \;+\; \frac{k_1}{k_2}\log(x_3(t)),$$

and, since $\lim_{t\to\infty} x_3(t) = 0$, $\theta(t)$ tends to $-\infty$ modulo 2π, *i.e.* $\theta(t)$ does not converge. Hence the vector $(x_1(t), x_2(t))$ does not stop turning around the origin. This implies that

$$\limsup_{t\to+\infty} x_{1+}(t) \neq 0,$$

for all $(x_1(0), x_2(0)) \in \mathbb{R}^2/\{0\}$. This last equation shows that the asymptotic property expressed by the second of equations (9.44) cannot be improved.

"lim inf" convergence case

Let $p = 2$, $q = 3$, $k_1 = k_2 = 1$, $\Psi(s) = |s|$ and consider the three cases $\eta = 0$; $\eta = 1$ with $r = 2$; and $\eta = 1$ with $r = 4$. Fig. 9.1 shows the trajectory of the system with initial condition $x(0) = \begin{bmatrix} 0.5 & 0 & 1 \end{bmatrix}'$ for the three cases, whereas Fig. 9.2 shows the time histories of the states x_1, x_2 and x_3. Note that the time axis is in log-scale. Fig. 9.2 highlights that all trajectories with initial condition off the (x_1, x_2)-plane have an oscillatory behavior with a period that tends to infinity. Note that trajectories with initial conditions such that $x_3(0) = 0$ converge to the set

$$\{(x_1, x_2) \mid x_1^2 + x_2^2 = x_1(0)^2 + x_2(0)^2, x_1 \leq 0\},$$

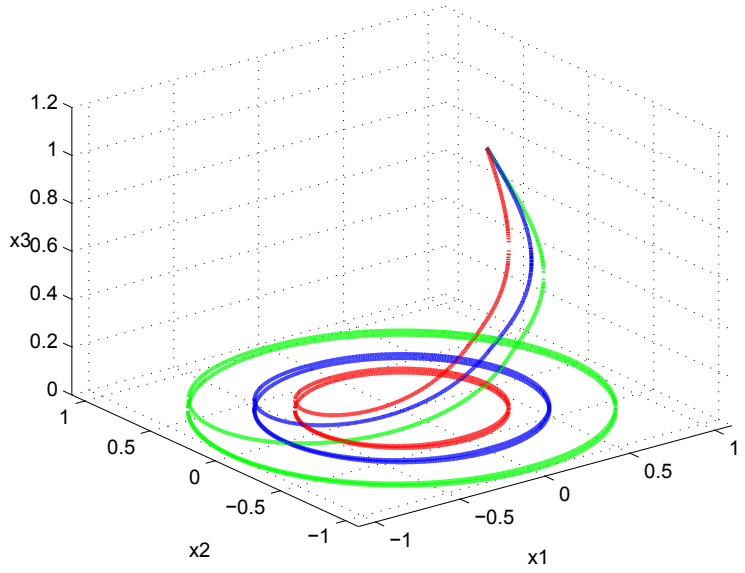

Figure 9.1: The trajectory of system (9.42), with $p = 2$, $q = 3$, $k_1 = k_2 = 1$, $\Psi(s) = |s|$, $x(0) = [0.5\ 0\ 1]'$, for the three considered cases. The trajectory converges to the circle of radius $\rho_0 = 0.5$ for $\eta = 0$ (red/dotted); of radius $\min(1, \rho_0) \leq \rho \leq \max(1, \rho_0)$ for $\eta = 1$ and $r = 4$ (blue/solid); of radius 1 for $\eta = 1$ and $r = 2$ (green/dashed).

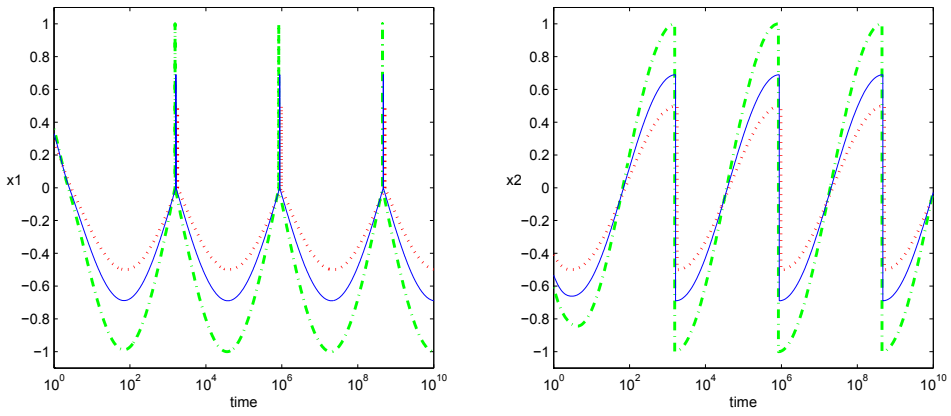

Figure 9.2: Time histories of the states of system (9.42) with $p = 2$, $q = 3$, $k_1 = k_2 = 1$, $\Psi(s) = |s|$ and $x(0) = [0.5\ 0\ 1]'$: $\eta = 0$ (red/dotted); $\eta = 1$ and $r = 4$ (blue/solid); $\eta = 1$ and $r = 2$ (green/dashed).

i.e. to a semi-circle centered at the origin, the size of which depends upon the initial conditions. This set is not stable, hence condition 3*b*) does not hold.

Remark 82. The ω-limit set of the trajectories of the system starting off the (x_1, x_2)-

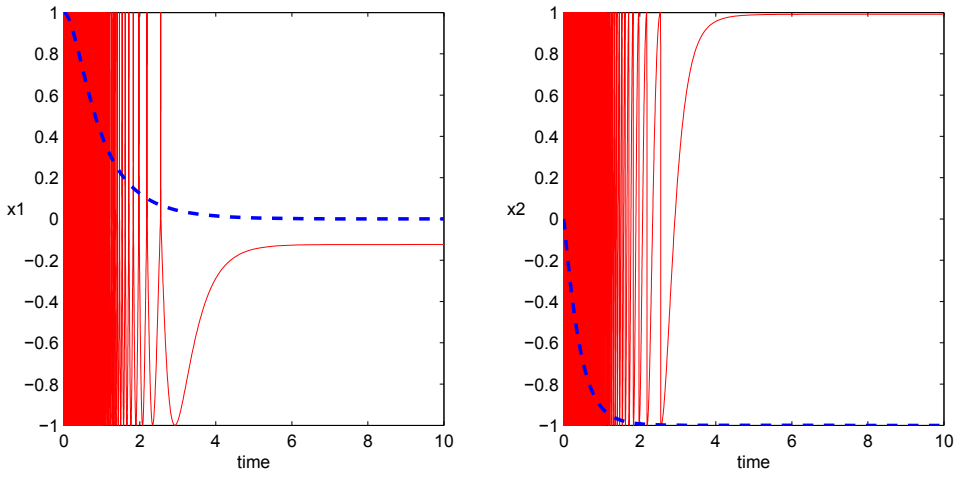

Figure 9.3: Time histories of the states of the system (9.42) with $p = 2$, $q = 1$, $\Psi(s) = |s|$, $\eta = 0$ and $x(0) = [1\ 0\ 1]'$: $k_1 = k_2 = 1$ (blue/dashed); $k_1 = 1000$, $k_2 = 1$ (red/solid).

plane is, as detailed in [239], a chain recurrent set, which strictly contains the ω-limit set of the trajectories of the system starting in the (x_1, x_2)-plane, consistently with the results in [239] and [240] on asymptotically autonomous semiflows. ■

Remark 83. As a consequence of the discussion in this section, the (x_1, x_2)-subsystem of system (9.42), with $p = 2$ and $q = 3$, and x_3 regarded as an input, does not possess the converging-input converging-state property, see [241], [242] and [243]. This does not contradict the result in [241], which highlights (among other things, and similarly to what is done in this chapter) the importance of asymptotic stability (of an equilibrium, or of a set) to establish asymptotic properties of solutions. ■

"lim" convergence case

Let $p = 2$, $q = 3$, $\Psi(s) = |s|$, $\eta = 0$ and consider the two cases $k_1 = k_2 = 1$; and $k_1 = 1000$, $k_2 = 1$. Figure 9.3 shows a trajectory with initial state $x(0) = [1\ 0\ 1]'$. Unlike the previous case $x_1(t)$ and $x_2(t)$ converge to a point such that $x_{1+} = 0$. Note that this is the case also if x_1 and x_2 undergo fast transient (solid line). Finally, Figure 9.4 shows the projection of the phase portrait on the (x_1, x_2)-plane for $p = 2$, $q = 3$, $k_1 = k_2 = 1$, $\Psi(s) = |s|$ and $\eta = 0$ and the set of initial conditions $\{(x_1(0), x_2(0), x_3(0)) : x_1(0)^2 + x_2(0)^2 \leq 1,\ x_3(0) = 1\}$.

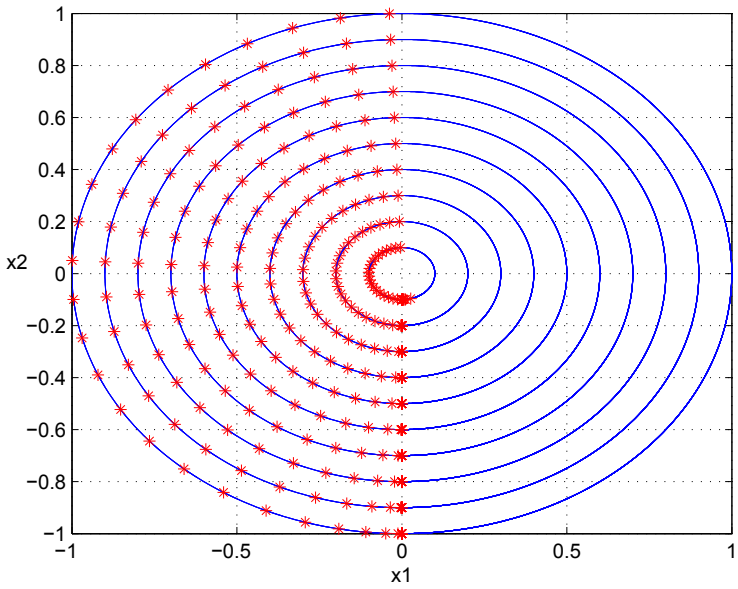

Figure 9.4: Projection of the phase portrait on the (x_1, x_2)-plane for $p = 2$, $q = 3$, $k_1 = k_2 = 1$, $\Psi(s) = |s|$ and $\eta = 0$ and the set of initial conditions $\{(x_1(0), x_2(0), x_3(0)) : x_1(0)^2 + x_2(0)^2 \leq 1,\, x_3(0) = 1\}$. The final states are represented by stars.

9.6 Conclusion

A class of theorems inspired by the Krasovskii-LaSalle invariance principle has been presented in a unified framework. The contribution of the chapter is a tool to study "lim inf" convergence properties of solutions of dynamical systems. In particular the theorems give sufficient conditions to determine the convergence in the mean and the "lim inf" convergence. These theorems are derived by a relaxation of Matrosov and Small-gain Theorems, and they are based on a "lim inf" Barbalat's Lemma (Lemma 22 and 24). Additional technical assumptions to have "lim" convergence are given. The "lim inf" / "lim" relation and the role of some of the assumptions are illustrated by means of examples.

Part III

Control

Chapter 10

Approximate finite-horizon optimal control for input-affine nonlinear systems with input constraints

10.1 Introduction

Given a system with known initial state and with hard constraints on the amplitude of the control input, the aim of the finite-horizon optimal control problem is to determine the control signal which minimizes the cost functional while satisfying the control constraints. The general problem is hard to solve due to several difficulties such as complicated dynamics of the system, the "shape" of the cost functional and the type of constraints. In this chapter the analysis is focused on a particular class of problems: time optimal control problems for input affine systems and some related problems, as for instance the maximum range optimal control problem. For linear systems this class of problems has a well-known solution which relies upon the use of bang-bang controls, see *e.g.* the book [244]. For nonlinear systems the solution can be found solving the associated Hamilton-Jacobi-Bellman (HJB) partial differential equation (pde). However, it might be difficult or even impossible to solve the equation analytically.

In [245] and [246] a new method to solve, approximatively, classes of optimal control prob-

lems has been developed. The method relies upon the use of dynamic state feedback and does not require the solution of any pde. In particular the problem therein studied is the linear quadratic regulator problem of an input-affine nonlinear system without constraints. In this chapter we extend the results in [246] providing approximate solutions to a class of finite-horizon optimal control problems with locally non-quadratic cost functionals and in the presence of input constraints. In particular we focus our analysis on the minimum time optimal control problem and the maximum range optimal control problem, which have been widely studied and are of interest for both theory and application. Note that differently from most of the literature on approximate solutions of optimal control problems, this method does not rely upon a discretization of the problem (for other continuous time approaches see *e.g.* [247]).

The extension of the ideas in [246] is not straightforward. Since locally non-quadratic costs are considered, it is not possible to exploit the solution of the algebraic Riccati equation (ARE) associated to the linearized problem in the construction of an approximate solution of the problem as done in [246]. The consequences of this fact are threefold. First, there is no a priori information on the shape of the approximate solution and, second, there is the additional difficulty of enforcing that the value function is positive definite. Finally the non-differentiability of the value function which is inherent to problems in which there is a hard constraint on the input leads to a feedback that may not be everywhere differentiable. We illustrate the theory solving the approximate minimum time optimal control problem for a bioreactor [248] and the approximate maximum range optimal control problem for a rocket [249], the so-called Goddard problem.

The chapter is organized as follows. In Section 10.2 the formulation of the problem is given together with some additional definitions. In Section 10.3 the main result is presented, *i.e.* a dynamic control law that approximatively solves the optimal control problem. In Section 10.4, the two case studies are discussed, and in Section 10.5 conclusions are drawn.

All the results of this chapter are original contributions developed in fulfillment of my PhD course and they have been published in the conference paper [23] and in the journal paper [24].

10.2 Preliminaries

10.2.1 Problem formulation

Consider an input-affine, nonlinear, system described by the equation[1]

$$\dot{x} = f(x) + g(x)u, \tag{10.1}$$

where $x(t) \in \mathbb{R}^n$ is the state of the system, the mappings $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are sufficiently smooth and $u(t) \in \mathbb{R}^m$ is the control input subject to the constraints

$$u_{min} \leq u_j(t) \leq u_{max}, \quad j = 1, 2, ..., m, \quad \forall t \in \mathbb{R}, \tag{10.2}$$

with $u_{min} < u_{max}$. For convenience, to simplify the theoretical analysis, we assume $u_{min} = -1$ and $u_{max} = 1$. Finally, consider the additional equation

$$Cx(T) = 0, \tag{10.3}$$

with $C \in \mathbb{R}^{r\times n}$ constant, which is used to model r constraints on the final state. Note, however, that the proposed method does not enforce this type of constraints, and that condition (10.3) has to be checked *a posteriori*.

The aim of the minimum time optimal control problem is to find a control input u that minimizes the cost functional

$$J(x_0, u) = \int_0^T dt,$$

where T is not *a priori* assigned, while satisfying the constraints (10.1), (10.2) and (10.3). To include in our study a wider class of constrained optimal control problems, as for instance the maximum distance optimal control problem, we consider more general cost functionals. Thus, we want to find a control input u that minimizes the cost functional

$$J(x_0, u) = \int_0^T [\phi(x(t)) + \gamma(x(t))u(t)]dt, \tag{10.4}$$

[1]For simplicity, the arguments of the functions are dropped whenever this does not cause confusion.

with $\phi : \mathbb{R}^n \to \mathbb{R}$ and $\gamma : \mathbb{R}^n \to \mathbb{R}^{1\times m}$ smooth mappings, while satisfying the constraints (10.1), (10.2) and (10.3). Hence, the optimal control problem has running cost

$$\mathcal{L}(x, u) = \phi(x) + \gamma(x)u, \tag{10.5}$$

and terminal cost

$$\mathcal{T}(x(T), u(T)) = 0. \tag{10.6}$$

The dynamic programming solution of this problem is based on the solution $V : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ of the HJB pde

$$\phi(x) + V_t + V_x f(x) - |\gamma(x) + V_x g(x)| = 0, \tag{10.7}$$

where the notation

$$|\gamma(x) + V_x g(x)| = \sum_{j=1}^{m} |\gamma_j(x) + \sum_{i=1}^{n} V_{x_i} g_{ij}(x)| \tag{10.8}$$

has been used, subject to the condition $V(T, x) = 0$.

We can now formalize the optimal control problem under analysis.

Problem 2. Consider the system (10.1), the constraints (10.2) and (10.3), and the cost (10.4). The *regional dynamic constrained finite-horizon optimal control* problem consists in finding an integer $\tilde{n} \geq 0$, a dynamic control law described by equations of the form

$$\dot{\xi} = \alpha(x, \xi, t), \qquad u = \beta(x, \xi, t), \tag{10.9}$$

with $\xi(t) \in \mathbb{R}^{\tilde{n}}$, $\alpha : \mathbb{R}^n \times \mathbb{R}^{\tilde{n}} \times \mathbb{R} \to \mathbb{R}^{\tilde{n}}$ and $\beta : \mathbb{R}^n \times \mathbb{R}^{\tilde{n}} \times \mathbb{R} \to \mathbb{R}^m$ smooth mappings with $|\beta_i(x, \xi, t)| \leq 1$ for all $i \in [1, m]$, and a set $\bar{\Omega} \subset \mathbb{R}^n \times \mathbb{R}^{\tilde{n}}$ such that the closed-loop system

$$\dot{x} = f(x) + g(x)\beta(x, \xi, t), \qquad \dot{\xi} = \alpha(x, \xi, t), \tag{10.10}$$

satisfies the condition

$$J(x_0, \beta) \leq J((x_0, \tilde{u}),$$

for any $\tilde{u}$ satisfying the constraint (10.2), and any (x_0, ξ_0) for which the trajectory of the system (10.10) remains in $\bar{\Omega}$ for all $t \in [0, T]$.

The solution of this problem is still hard to determine. Hence, we define an approximate version of the regional dynamic finite-horizon optimal control problem.

Problem 3. Consider the system (10.1), the constraints (10.2) and (10.3), and the cost (10.4). The *approximate regional dynamic constrained finite-horizon optimal control* problem consists in finding an integer $\tilde{n} \geq 0$, a dynamic control law described by the equations (10.9), a set $\bar{\Omega} \subset \mathbb{R}^n \times \mathbb{R}^{\tilde{n}}$ and functions $\rho_1 : \mathbb{R}^n \times \mathbb{R}^{\tilde{n}} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ and $\rho_2 : \mathbb{R}^n \times \mathbb{R}^{\tilde{n}} \to \mathbb{R}_{\geq 0}$ such that the regional dynamic constrained finite-horizon optimal control problem is solved with respect to the running cost

$$\mathcal{L}(x, \xi, u) = \phi(x) + \gamma(x)u + \rho_1(x, \xi, t), \tag{10.11}$$

and the terminal cost

$$\mathcal{T}(x(T), \xi(T)) = \rho_2(x(T), \xi(T)). \tag{10.12}$$

Remark 84. Ideally the function ρ_1 and ρ_2 should be as small as possible to guarantee that the solution of the approximate problem is "close" to the solution of the original problem. ■

Remark 85. The non-negativity of ρ_2 is required to avoid that the terminal cost may become unbounded. This assumption can be relaxed requiring that ρ_2 be bounded from below. ■

To simplify the forthcoming development we discuss the underlying linearized problem. Let x_ℓ and u_ℓ be the linearization point and define $A = \dfrac{\partial f}{\partial x}(x_\ell)$, $F = \dfrac{\partial \phi}{\partial x}(x_\ell)$, $h_{s\ell} = f(x_\ell) + g(x_\ell)u_\ell$, $h_{c\ell} = \phi(x_\ell) + \gamma(x_\ell)u_\ell$, $B = g(x_\ell)$ and $G = \gamma(x_\ell)$. Suppose without loss of generality that $x_\ell = 0$ and $u_\ell = 0$. Consider the linear system described by the equation

$$\dot{x} = Ax + Bu + h_{s\ell}, \tag{10.13}$$

and the cost functional

$$J_\ell(x_0, u(t)) = \int_0^T [Fx(t) + Gu(t) + h_{c\ell}]dt, \tag{10.14}$$

subject to the constraints (10.2) and (10.3). The dynamic programming solution of this

problem relies on the solution $V_\ell : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ of the HJB pde

$$Fx + h_{c\ell} + V_{\ell\tau} + V_{\ell x}(Ax + h_\ell) - |V_{\ell x}B + G| = 0,$$

subject to $V_\ell(T, x) = 0$, where $V_{\ell x}$ and $V_{\ell\tau}$ are the partial derivatives of the value function V_ℓ. Note that V_ℓ may not be everywhere differentiable. For the following it is enough assuming that V_ℓ is differentiable almost everywhere.

Remark 86. In what follows we assume that the underlying linearized problem has been solved and V_ℓ has been computed. Although solving the linearized problem is in general difficult for the cost functional (10.4), there are some significant cases in which this is possible, *e.g.* the minimum time optimal control problem. As shown in the examples, with the method proposed we aim to outperform the solution of the linearized problem when this is known. ■

Remark 87. Note that the choice of the linearization point (x_ℓ, u_ℓ) may be arbitrary. A possibly good selection is $(x(T), u(T))$, when known (*e.g.* in the bioreactor problem in Section 10.4.1), or an approximation of $(x(T), u(T))$ (*e.g.* in the Goddard problem in Section 10.4.2). Alternatively, one may use the final values obtained solving the linearized problem. ■

10.2.2 Algebraic solution and value function

Consider the extended state $(x^T, \tau)^T$, with $\dot{\tau} = 1$ and assume that the HJB equation (10.7) with $t = \tau$ can be solved *algebraically*, as detailed hereafter.

Definition 33. Let $\Sigma : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n\times n}$, with $x^T\Sigma(x,\tau)x$ non-negative for all $(x, \tau) \in \mathbb{R}^n \times \mathbb{R}$, and $\sigma : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}_{\geq 0}$. A mapping $W(x,\tau) = [V_{\ell x} + \Delta_x(x,\tau), V_{\ell\tau} + \Delta_\tau(x,\tau)]$, $\Delta_x : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{1\times n}$, $\Delta_\tau : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, is an *algebraic* $\widehat{W}$ *solution* of (10.7) if it is almost everywhere differentiable and

$$0 = \phi + [V_{\ell\tau} + \Delta_\tau(x,\tau)] + [V_{\ell x} + \Delta_x(x,\tau)]f - |[V_{\ell x} + \Delta_x(x,\tau)]g + \gamma| + x^T\Sigma x + \tau^2\sigma. \tag{10.15}$$

Using the algebraic $\widehat{W}$ solution of equation (10.15), define the function

$$V(x,\tau,\xi,s) = V_\ell(x,\tau) + \Delta_x(\xi,s)x + \Delta_\tau(\xi,s)\tau + \frac{1}{2}||x-\xi||_R^2 + \frac{1}{2}b||\tau - s||^2, \tag{10.16}$$

where $\xi \in \mathbb{R}^n$, $s \in \mathbb{R}$, $b > 0$, $R = R^T > 0$ is a matrix of positive weights and $||v||_R^2 = v^T R v$.

Remark 88. We use a different notion of algebraic solution than the one in [246]. Therein, the solution of the ARE associated to the linear problem is exploited to ensure that the function V is locally positive definite. Since in our case the cost functional is not quadratic, it is not possible to use the ARE. To guarantee that the function V is locally positive definite we also exploit the solution of the associated linearized problem, but compute the "correction" terms in a different way. ■

10.3 Dynamic control law

Consider the nonlinear system (10.1), the constraints (10.2) and (10.3), and the cost functional (10.4). Let

$$\Lambda(\xi, s) = \Psi(\xi, s) R^{-1},$$

$$\lambda(\xi, s) = \psi(\xi, s) R^{-1},$$

where $\Psi(\xi, s) \in \mathbb{R}^{n \times n}$ and $\psi(\xi, s) \in \mathbb{R}^{1 \times n}$ are the Jacobian matrices of the mapping $\Delta_x(\xi, s)$ and $\Delta_\tau(\xi, s)$ with respect to ξ. Let $P(x)$, $\Phi(x, \xi, s)$, $\ell(x, \tau, s)$, $H(x, \xi, s)$, $\Pi(x, \tau, s)$, $W_1(x, \xi, s)$, $W_2(x, s)$, $D_1(x, \xi, s)$ and $D_2(x, s)$ be almost everywhere continuous mappings such that

$$f(x) = P(x)x,$$

$$\Delta_x(x, s) - \Delta_x(\xi, s) = (x - \xi)^T \Phi(x, \xi, s)^T,$$

$$\Delta_\tau(x, s) - \Delta_\tau(x, \tau) = \ell(x, \tau, s)(s - \tau),$$

$$\Delta_\tau(\xi, s) - \Delta_\tau(x, s) = (x - \xi)^T H(x, \xi, s)(x - \xi),$$

$$\Delta_x(x, s) - \Delta_x(x, \tau) = x^T \Pi(x, \tau, s),$$

$$\frac{\partial \Delta_x(\xi, s)}{\partial s} - \frac{\partial \Delta_x(x, s)}{\partial s} = W_1(x, \xi, s)(x - \xi),$$

$$\frac{\partial \Delta_x(x, s)}{\partial s} = W_2(x, s)x,$$

$$\frac{\partial \Delta_\tau(\xi, s)}{\partial s} - \frac{\partial \Delta_\tau(x, s)}{\partial s} = D_1(x, \xi, s)(x - \xi),$$

$$\frac{\partial \Delta_\tau(x,s)}{\partial s} = D_2(x,s)x.$$

We are now ready to state the main result of the chapter.

Proposition 13. Consider the system (10.1), the constraints (10.2) and (10.3), and the cost (10.4). Let W be an algebraic $\widehat{W}$ solution of (10.15). Let $R = R^T > 0$ and $b > 0$ be such that the function V defined in (10.16) is positive definite. In addition suppose there exists a set $\Omega \subseteq \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$ in which

$$\begin{bmatrix} L_1 & L_2 \\ L_2^T & L_3 \end{bmatrix} < \begin{bmatrix} \Sigma & 0 \\ 0 & \sigma \end{bmatrix}, \tag{10.17}$$

and

$$|L_4| \leq |L_4 + (x-\xi)^T R g|, \tag{10.18}$$

for all $(x, \tau, \xi, s) \in \Omega$, where

$$L_1 = P^T \Pi^T + \eta W_2 + \Lambda Y + Y^T \Lambda^T + \Lambda H \Lambda^T,$$

$$L_2 = Y\lambda^T + \Lambda H \lambda^T + \frac{\eta}{2} D_2^T + \frac{\eta}{2} \Lambda D_1^T,$$

$$L_3 = \lambda H \lambda^T + \frac{\eta}{2} D_1 \lambda^T + \frac{\eta}{2} \lambda D_1^T,$$

$$L_4 = (V_{\ell x} + \Delta_x(x,\tau))g + \gamma,$$

$$Y = \frac{1}{2}(R - \Phi)^T P + \frac{\eta}{2} W_1,$$

and $\eta = 1 - \dfrac{\ell}{b}$. Suppose finally that

$$\frac{\partial}{\partial u_i}\left[\frac{d}{dt}\left(V_x g_i + \gamma_i\right)\right] \neq 0, \quad \text{if } \{V_x g_i + \gamma_i\} = 0. \tag{10.19}$$

Then there exists $\bar{k}$ such that for all $k > \bar{k}$ the function V satisfies the HJB inequality

$$\mathcal{HJBI} \triangleq \phi + V_x f + V_\tau + V_\xi \dot{\xi} + V_s \dot{s} - |V_x g + \gamma| \leq 0, \tag{10.20}$$

for all $(x, \tau, \xi, s) \in \Omega$, with $\dot{\xi} = -kV_\xi^T$ and $\dot{s} = \eta$. Hence

$$\begin{aligned} \dot{s} &= \eta, \\ \dot{\xi} &= -k(\Psi^T x - R(x-\xi) + \psi^T \tau), \\ u_i &= \begin{cases} -\operatorname{sign}\{V_x g_i + \gamma_i\}, & \text{if } \{V_x g_i + \gamma_i\} \neq 0, \\ u_{i_{singular}}, & \text{if } \{V_x g_i + \gamma_i\} = 0, \end{cases} \end{aligned} \tag{10.21}$$

with $u_{i_{singular}}$ such that the derivative with respect to time of $\{V_x g_i + \gamma_i\}$ is identically equal to zero, solves Problem 3 with $\rho_1 = -\mathcal{HJBI} \geq 0$ and $\rho_2 = V(x(T), \tau(T), \xi(T), s(T))$.

Proof. Consider the partial derivatives of the function V defined in (10.16), namely

$$\begin{aligned} V_x &= V_{\ell x} + \Delta_x(\xi, s) + (x-\xi)^T R^T = V_{\ell x} + \Delta_x(x, \tau) + (x-\xi)^T (R-\Phi)^T + x^T \Pi, \\ V_\tau &= V_{\ell \tau} + \Delta_\tau(\xi, s) + b(\tau - s) = V_{\ell \tau} + \Delta_\tau(x, s) + b(\tau - s) + (x-\xi)^T H (x-\xi), \\ V_\xi &= x^T \Psi - (x-\xi)^T R + \tau \psi, \\ V_s &= x^T \frac{\partial \Delta_x}{\partial s}(\xi, s) + \tau \frac{\partial \Delta_\tau}{\partial s}(\xi, s) - b(\tau - s), \end{aligned}$$

and substitute them in (10.20). Substituting (10.15) into (10.20), yields

$$\mathcal{HJBI} = \mu_1(x, \tau, \xi, s) + \mu_2(x, \tau, s) + \mu_3(x, \tau),$$

where μ_1 is a quadratic form in the variables x, $(x-\xi)$ and τ, *i.e.*

$$\mu_1 \triangleq -\begin{bmatrix} x^T & (x-\xi)^T & \tau \end{bmatrix} \begin{bmatrix} M + kC^T C \end{bmatrix} \begin{bmatrix} x \\ (x-\xi) \\ \tau \end{bmatrix}, \tag{10.22}$$

with $C = \begin{bmatrix} \Psi^T & -R & \psi^T \end{bmatrix}$,

$$M = \begin{bmatrix} M_{11} & -Y^T & -\frac{\eta}{2} D_2^T \\ -Y & -H & -\frac{\eta}{2} D_1^T \\ -\frac{\eta}{2} D_2 & -\frac{\eta}{2} D_1 & \sigma \end{bmatrix},$$

and $M_{11} = \Sigma - \eta W_2 - P^T \Pi^T$. The term μ_2 is defined as

$$\mu_2 \triangleq \Delta_\tau(x,s) - \Delta_\tau(x,\tau) + b(\tau - s) - b(\tau - s)\eta. \tag{10.23}$$

From the definition of η it follows that μ_2 is identically equal to zero for all $(x, \xi, \tau, s) \in \Omega$. In fact, $\eta = 1 - \dfrac{\ell}{b}$ and $\Delta_\tau(x,s) - \Delta_\tau(x,\tau) - \ell(x,\tau,s)(s-\tau) = 0$ by definition. Finally, the term μ_3 is defined as

$$\mu_3 \triangleq -|V_x g + \gamma| + |(V_{\ell x} + \Delta_x(x,\tau))g + \gamma|, \tag{10.24}$$

which is nonpositive by condition (10.18).

Note now that the matrix $C \in \mathbb{R}^{n \times 2n+1}$ has constant rank n and the columns of the matrix

$$Z = \begin{bmatrix} I_{n\times n} & 0_{n\times 1} \\ R^{-1}\Psi^T & R^{-1}\psi^T \\ 0_{1\times n} & 1 \end{bmatrix},$$

span the kernel of the matrix C. Condition (10.17) implies that the matrix $Z^T M Z$ is positive definite, hence exploiting the results of [250] there exists $\bar{k}$ such that, for all $k > \bar{k}$, $\mu_1 \leq 0$.

This proves that the HJB inequality (10.20) is satisfied for all $(x, \xi, \tau, s) \in \Omega$. □

Remark 89. The gain k in the $\dot{\xi}$ equation in (10.20) can be selected as a function of (x, ξ, τ, s) to reduce the absolute value of μ_1, hence the additional running cost ρ_1 in (10.11). ■

Remark 90. Note that condition (10.18) is not necessary. Suppose that there exists a set Ω in which $V_\xi V_\xi^T$ is nonzero, then selecting

$$k(x,\tau,\xi,s) = (V_\xi V_\xi^T)^{-1} \begin{bmatrix} x^T & (x-\xi)^T & \tau \end{bmatrix} M \begin{bmatrix} x \\ (x-\xi) \\ \tau \end{bmatrix} - \mu_3 \tag{10.25}$$

yields $\mu_1 = -\mu_3$, thus providing a solution to Problem 3 with the additional running cost $\rho_1 = 0$ without the need to satisfy (10.18). Note that there is no condition on the sign of the gain k, since in this context stability requirements are not imposed. ■

Remark 91. The additional terminal cost ρ_2 depends upon the gain k and the initial condition of ξ and s. If k is chosen to minimize ρ_1 as described in Remark 90, ρ_2 can be minimized selecting the initial condition of ξ and s. ■

Remark 92. The left hand side of (10.17) is zero at zero. Then, if $\Sigma(0,0) = \bar{\Sigma} > 0$ and $\sigma(0,0) = \bar{\sigma} > 0$, by continuity, there exists a region Ω containing the origin such that (10.17) is satisfied for all $(x, \tau, \xi, s) \in \Omega$. ■

10.4 Application of the approximate solution

In this section the theory is illustrated by means of two classic examples: the minimum time optimal control problem for a bioreactor and the maximum range optimal control problem for a rocket, the so-called Goddard problem. Proposition 13 is used to design a dynamic control law solving an approximate problem. For both examples we compare the optimal control law (which we denote u_*, and which can be explicitly computed) with the control law obtained from the optimal solution of the linearized problem (which we denote u_ℓ) and the dynamic control law obtained from the solution of the approximate problem (which we denote u_d).

10.4.1 Example: minimum time optimal control of a bioreactor

Several results have been obtained in the control of bioreactors, see *e.g.* [248], [251], [252], [253] and references therein. In [248] an optimal policy for a class of bioreactors has been proposed, while in [253] the optimal policy has been developed for a more general class of bioreactors and it has been demonstrated that the optimal policy may contain singular arcs.

A simple model that globally describes the dynamic behavior of a bioreactor is given by the differential equations, see [248],

$$\dot{z} = \theta(z, v) + \pi(z, v)u, \qquad \dot{v} = u, \tag{10.26}$$

where

$$\theta(z(t), v(t)) = -\mu(z(t)) \left[z_{in} - z(t) + \frac{\rho(w_0)}{y v(t)} \right],$$
$$\pi(z(t), v(t)) = \frac{z_{in} - z(t)}{v(t)},$$

$$\mu(z) = \frac{\mu_0 z}{k_s + z + \dfrac{z^2}{k_i}}, \tag{10.27}$$

and $\rho(w_0) = v_0(x_0 + yz_0) - yz_{in}v_0$. The variable $z(t)$ describes the substrate concentration in the tank, $v(t)$ the volume of water in the tank and $u(t)$ the input water flow. In addition x_0 is the initial biomass concentration, y is the constant yield coefficient, z_{in} the constant substrate concentration in the input flow, k_s the affinity constant, k_i the inhibition constant and μ_0 the constant maximum specific growth rate. The function $\mu(z)$ models the microbiological growth rate, which is described using the Haldane law (10.27). The control u takes values in the interval $[0, u_{max}]$, with $u_{max} > 0$, and the variables z and v take non-negative values. Furthermore, when the maximum level of water in the tank v_{max} is reached the control u is set to zero. Differently from [248], z_{in} is assumed constant.

The purpose of the bioreactor is to control the concentration z of the substrate in the tank below a specified level z_{min}, before the volume reaches v_f, where $0 < v(0) < v_f \leq v_{max}$. The cost functional to be minimized is the reaction time T, *i.e.* in (10.4) we have $\phi(x) = 1$ and $\gamma(x) = 0$. Thus, this is a minimum time optimal control problem.

The first step to develop the dynamic control law consists in computing the solution of the minimum time optimal control problem for the linearized system. The linearized system around (z_f, v_f) is described by equations of the form

$$\begin{aligned} \dot{x}_1 &= -ax_1 + bx_2 + cu + d, \\ \dot{x}_2 &= u, \end{aligned} \tag{10.28}$$

where $x_1 = z - z_f$, $x_2 = v - v_f$ and a, b, c and d are constants. Note that d is zero if the state (z_f, v_f) is an equilibrium point. To compute the value function for the linear problem, the system is integrated and the time t is eliminated, since the value function is exactly the time needed to reach the final state. From general results on linear systems it is known that the optimal control has at most one switching time and the control input assumes only the extreme values. With these observations only two non-trivial situations are possible: the control input is 0 until the switching time and u_{max} until the final time,

or, the control input is u_{max} until the switching time and 0 thereafter. A simple analysis reveals that the first candidate optimal control does not yield positive switching times, whereas the second candidate optimal control yields a positive switching time for all states. Let V_ℓ be the value function of the linearized problem and $V_{\ell z}$ and $V_{\ell v}$ its partial derivatives with respect to z and v, respectively.

Note that

$$1 + V_{\ell v} u_{max} + V_{\ell z} \dot{z}_\ell = 0, \tag{10.29}$$

where $\dot{z}_\ell$ is the right hand side of the first of the equations (10.28) with $u = u_{max}$. Consistently with Definition 33 the solution for this problem relies on the solution of the algebraic equation

$$1 + (V_{\ell v} + \Delta_v) u_{max} + (V_{\ell z} + \Delta_z)[\theta + \pi u_{max}] = 0. \tag{10.30}$$

One $\widehat{W}$ solution is given by

$$\Delta_z = -V_{\ell z}, \qquad \Delta_v = \frac{V_{\ell z} \dot{z}_\ell}{u_{max}}.$$

According to (10.16) the modified value function is

$$V(z, v) = V_\ell(z, v) + \frac{V_{\ell z}(\xi_1, \xi_2) \dot{z}_\ell(\xi_1, \xi_2)}{u_{max}} v - V_{\ell z}(\xi_1, \xi_2) z + \frac{1}{2} R_1 (z - \xi_1)^2 + \frac{1}{2} R_2 (v - \xi_2)^2.$$

Assuming that there exist $R_1 > 0$ and $R_2 > 0$ such that $V(z, v)$ is locally positive definite, we compute the partial derivatives of V, namely V_{ξ_1}, V_{ξ_2}, V_z and V_v. Proposition 13 yields the dynamic control law[2]

$$\begin{aligned}
\dot{\xi}_1 &= -k V_{\xi_1}, \\
\dot{\xi}_2 &= -k V_{\xi_2}, \\
u &= \begin{cases} [1 - \operatorname{sign}(V_z \pi + V_v)] \dfrac{u_{max}}{2}, & \text{if } V_z \pi + V_v \neq 0, \\ u_{singular}, & \text{if } V_z \pi + V_v = 0, \end{cases}
\end{aligned} \tag{10.31}$$

[2]The different definition of u is to adapt the proposition, developed when there is a symmetric bound on the control, to this problem in which the bound is asymmetric.

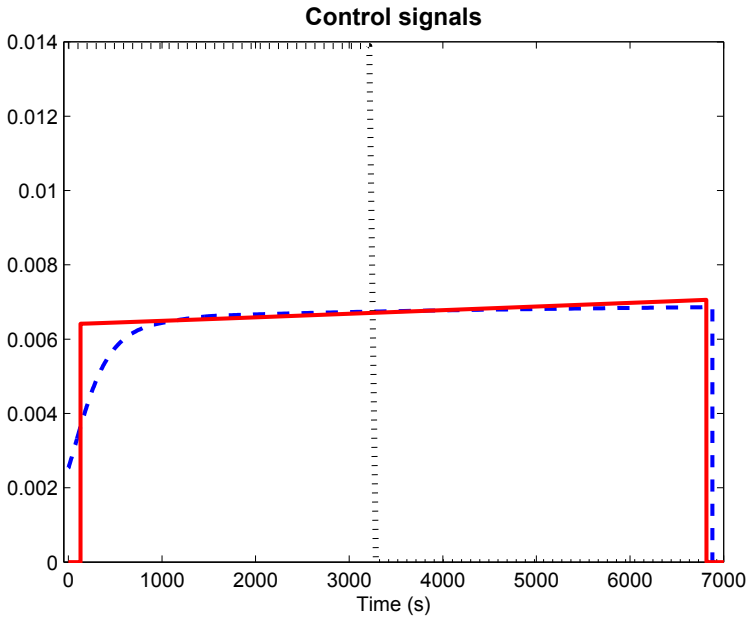

Figure 10.1: Time histories of the control signals u_* (solid line), u_ℓ (dotted line) and u_d (dashed line) for $(z_0, v_0) = (50, 5)$.

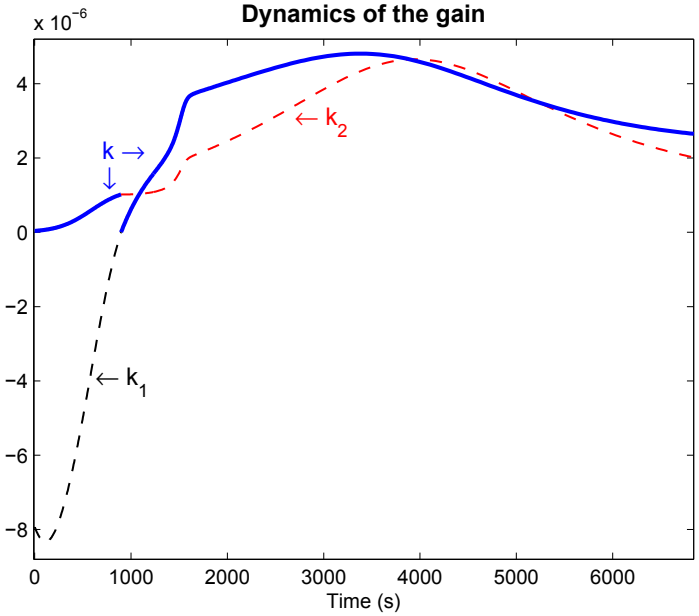

Figure 10.2: Time history of the gain k (solid line) for the initial condition $(z_0, v_0) = (50, 5)$. k_1 and k_2 are as in equation (10.32).

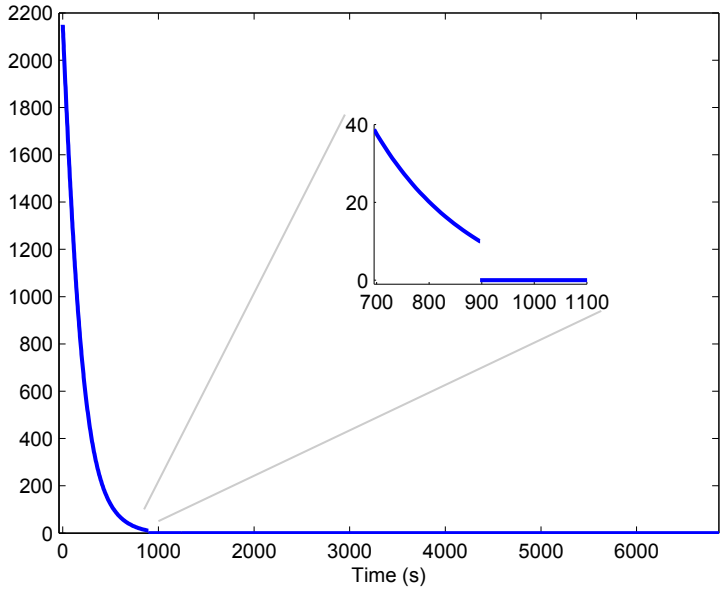

Figure 10.3: Time history of the additional running cost ρ_1 for the initial condition $(z_0, v_0) = (50, 5)$. The discontinuity is due to the switch in the gain k.

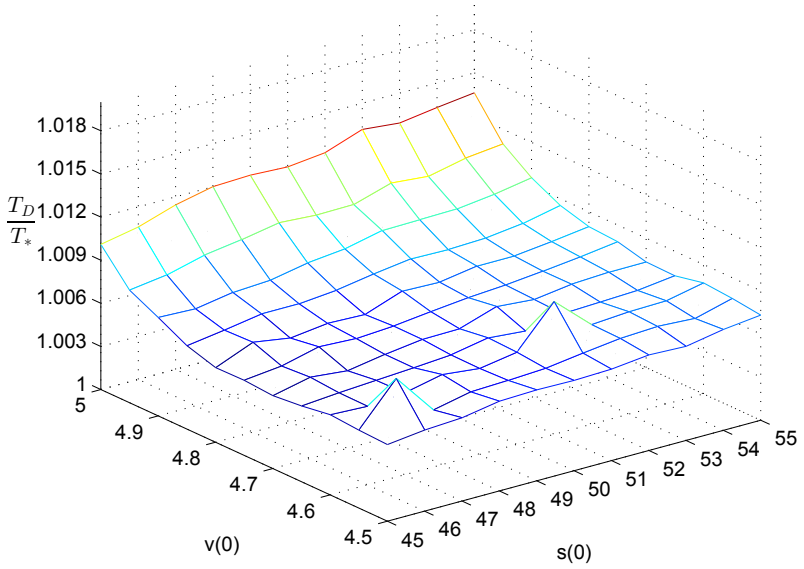

Figure 10.4: Ratio between the final time T_D, resulting from the application of the control law u_d, and the optimal final time T_* for $z_0 \in [45, 55]$ and $v_0 \in [4.5, 5]$.

with $u_{singular}$ such that the derivative with respect time of $V_z\pi + V_v$ is identically equal to zero.

The closed-loop system (10.26)-(10.31) has been simulated with the parameters (see [248]): $v_f = 50m^3$, $z_f = 1mgl^{-1}$, $u_{max} = 0.013888889m^3s^{-1}$, $k_s = 2mgl^{-1}$, $k_i = 50mgl^{-1}$, $\mu_0 = 0.00002s^{-1}$, $y = 0.5$ and $z_{in} = 300mgl^{-1}$. The initial conditions (z_0, v_0) have been selected in the regions $45 \leq z_0 \leq 55mgl^{-1}$ and $4.5 \leq v_0 \leq 5m^3$. The biomass concentration x_0 has been selected as $x_0 = 13000+3400(5-v_0)mgl^{-1}$ to recover the case studied in [248]. The gain of the dynamic control law has been defined as

$$k = \begin{cases} k_1 = \dfrac{1 + V_z\theta + (V_z\pi + V_v)u}{V_{\xi_1}^2 + V_{\xi_2}^2}, & \text{if } V_{\xi_1}^2 + V_{\xi_2}^2 \neq 0, \\ & \text{and } k_1 > 0, \\ k_2 = \dfrac{10}{1 + V_{\xi_1}^2 + V_{\xi_2}^2}, & \text{otherwise.} \end{cases} \tag{10.32}$$

The selection $k = k_1$ is in accordance with Remark 90 and guarantees that Proposition 13 holds with $\rho_1 = 0$. Finally, the initial condition for the dynamic extension has been set to $(\xi_1(0), \xi_2(0)) = (1, v_0)$ and $R_1 = R_2 = 340$.

Simulations have been carried out with the optimal control law u_*, the optimal control law for the linearized system u_ℓ and the dynamic control law u_d. As already discussed the optimal control law for the linearized problem gives a *batch-strategy*: the reactor is filled as fast as possible and the reaction phase is stopped when $z(t)$ reaches z_f.

Fig. 10.1 shows the time histories of the control signals $u_*(t)$, $u_\ell(t)$ and $u_d(t)$ for the initial state $(z_0, v_0) = (50, 5)$. Note that the dynamic control approximately recovers the optimal strategy. The final time for the optimal policy is 7071 seconds, for the dynamic control is 7154 seconds, and for the batch-strategy is 18524 seconds. Fig. 10.2 displays the time history of the gain k (solid line). Note that k is initially equal to k_2 and at approximatively 900 seconds it switches to k_1. This behavior is consistent with the one shown in Fig. 10.3 in which the value of ρ_1 is positive when $k = k_2$ and identically equal to zero when $k = k_1$. Fig. 10.4 shows the ratio $\dfrac{T_D}{T_*}$ for a range of initial conditions. Note that the dynamic control law yields performance similar to the optimal one and that it significantly outperforms the batch-strategy which gives an average ratio of 2.6.

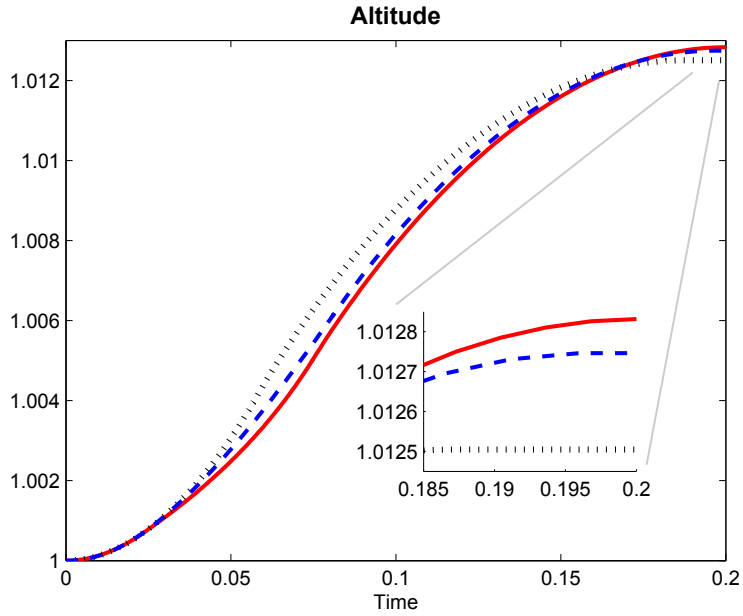

Figure 10.5: Time histories of the altitude resulting from the application of u_* (solid line), u_ℓ (dotted line) and u_d (dashed line).

10.4.2 Example: maximum range optimal control of a rocket

The Goddard problem consists in the maximization of the altitude of a rocket in vertical flight when atmospheric drag and gravity are considered. The model has been taken from [249] and [254] in which optimal solutions have been proposed. The rocket is regarded as a point of variable mass forced by a gravitational field with inverse square law. The aerodynamic drag is described by an exponential law, namely

$$D(v,h) = C_D b v|v| e^{\beta(1-h)},$$

where b and β are constants, and C_D is the zero-lift drag coefficient. The dynamics of the rocket are described by the nondimensionalized equations

$$\dot{h} = v, \qquad \dot{v} = -\frac{D(v,h)}{m} - \frac{1}{h^2} + \frac{u}{m}, \qquad \dot{m} = -\frac{u}{c}. \tag{10.33}$$

The altitude h is the distance from the Earth's center, v the velocity of the rocket, m the mass of the rocket, c the effective exhaust velocity of the gas flow and u the thrust of the rocket that is bounded, *i.e.* $0 \leq u \leq u_{max}$. The initial state (h_0, v_0, m_0) and the final mass

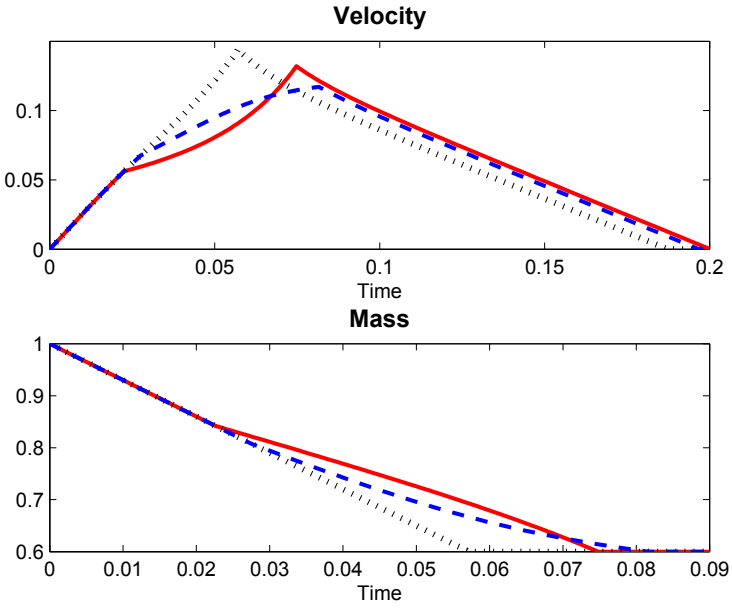

Figure 10.6: Time histories of the mass and the velocity resulting from the application of u_* (solid line), u_ℓ (dotted line) and u_d (dashed line).

m_f are known. The problem consists in the determination of the optimal policy of the thrust that maximizes the final altitude h_f. Thus, the cost functional to be minimized is

$$J = -(h(t_f) - h(0)) = -\int_0^{t_f} \dot{h}(t)dt = -\int_0^{t_f} v(t)dt,$$

i.e. in (10.4) we have $\phi(x) = -v$ and $\gamma(x) = 0$. Note that a singular arc in the optimal control can be present.

In the simulation the initial state has been taken as $(1, 0, 1)$, with an unknown final state of $(1 + \delta, 0, 0.6)$ and a linearization point of $(1 + \varepsilon, 0, 0.6)$. To compute the value function for the linearized problem, we integrate the linearized system noting that the only possible optimal control policy is to give full thrust until the mass reaches the final value and after to wait until the velocity reaches zero. To simplify the computation we solve the linearized problem with respect to the minimum time needed to finish the fuel. Let V_ℓ be the value function of this linearized problem and $V_{\ell h}$, $V_{\ell v}$ and $V_{\ell m}$ its partial derivatives with respect

to h, v and m, respectively. Note that

$$1 + V_{\ell h} v + V_{\ell v} \dot{x}_\ell - \frac{u_{max}}{c} V_{\ell m} = 0, \tag{10.34}$$

where $\dot{x}_\ell$ is given by the second equation of the linearized system computed for $u = u_{max}$. Consistently with Definition 33 the solution for this problem relies on the solution of the algebraic equation

$$(V_{\ell h} + \Delta_h)v - v - (V_{\ell v} + \Delta_v)\left(\frac{D(v,h)}{m} + \frac{1}{h^2}\right) + \left((V_{\ell v} + \Delta_v)\frac{1}{m} - \frac{1}{c}(V_{\ell m} + \Delta_m)\right)u_{max} = 0.$$

One $\widehat{W}$ solution is given by

$$\Delta_h = \alpha, \qquad \Delta_v = -V_{\ell v}, \qquad \Delta_m = -\frac{c}{u_{max}}(V_{\ell v}\dot{x}_\ell + (1-\alpha)v + 1).$$

According to (10.16) the modified value function is

$$\begin{aligned} V(h,v,m) = V_\ell(h,v,m) + \alpha h - V_{\ell v}(\xi)v - \frac{c}{u_{max}}(V_{\ell v}(\xi)\dot{x}_\ell(\xi) + (1-\alpha)\xi_2 + 1)m \\ + \frac{1}{2}R_1||h-\xi_1||^2 + \frac{1}{2}R_2||v-\xi_2||^2 + \frac{1}{2}R_3||m-\xi_3||^2. \end{aligned}$$

Assuming that there exist $R_1 > 0$, $R_2 > 0$ and $R_3 > 0$ such that $V(h,v,m)$ is positive definite, we compute the partial derivatives of V, namely V_{ξ_1}, V_{ξ_2}, V_{ξ_3}, V_h, V_v and V_m. Proposition 13 yields the dynamic control law

$$\begin{aligned} \dot{\xi_1} &= -kV_{\xi_1}, \\ \dot{\xi_2} &= -kV_{\xi_2}, \\ \dot{\xi_3} &= -kV_{\xi_3}, \\ u &= \begin{cases} \left[1 - \text{sign}\left(\dfrac{V_v}{m} - \dfrac{V_m}{c}\right)\right]\dfrac{u_{max}}{2}, & \text{if } \dfrac{V_v}{m} \neq \dfrac{V_m}{c}, \\ u_{singular}, & \text{if } \dfrac{V_v}{m} = \dfrac{V_m}{c}, \end{cases} \end{aligned} \tag{10.35}$$

with $u_{singular}$ such that the derivative with respect to time of $\dfrac{V_v}{m} - \dfrac{V_m}{c}$ is identically equal to zero.

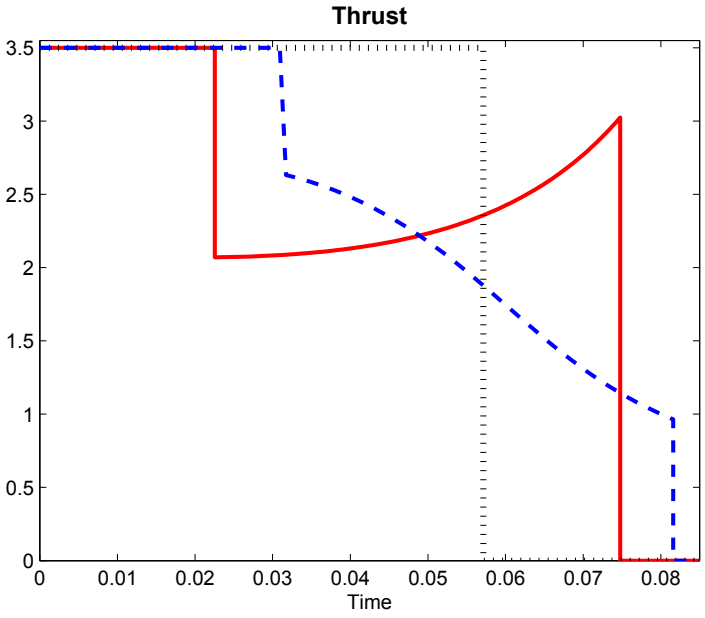

Figure 10.7: Time histories of the thrust u_* (solid line), u_ℓ (dotted line) and u_d (dashed line).

The closed-loop system (10.33)-(10.35) has been simulated with the parameters (see [255] and [254]): $C_D = 0.05$, $b = 6200$, $\beta = 500$, $u_{max} = 3.5$, $c = 0.5$ and final mass $m_f = 0.6$. The gain of the dynamic control law has been defined as

$$k = \frac{-v + V_h v + V_v + W_v \left(-\dfrac{D}{m} - \dfrac{1}{h^2} \right) + \left(\dfrac{V_v}{m} - \dfrac{V_m}{c} \right) u}{V_{\xi_1}^2 + V_{\xi_2}^2 + V_{\xi_3}^2}. \tag{10.36}$$

This selection of k is in accordance with Remark 90 and guarantees that Proposition 13 holds with $\rho_1 = 0$. The parameters of the dynamic control law have been selected as $\varepsilon = -0.2$, $\alpha = 1$, $R_1 = 180$, $R_2 = 106$ and $R_3 = 64$. Finally, the initial condition for the dynamic extension has been set to $(\xi_1(0), \xi_2(0), \xi_3(0)) = (1, 0.18, 0.92)$. Simulations have been carried out with the optimal control law $u_*(t)$, the optimal control law for the linearized problem $u_\ell(t)$ and the dynamic control law $u_d(t)$.

Fig. 10.5, Fig. 10.6 and Fig. 10.7 display the time histories of the altitude, the mass, the velocity and the thrust resulting from the application of u_*, u_ℓ and u_d. Note that the dynamic control law approximately recovers the optimal strategy as demonstrated by the presence of a singular arc. Finally, taking as performance index the relative distance

between the final altitudes, namely

$$r_i = \frac{h_*(T) - h_i(T)}{h_*(T) - h_*(0)}, \qquad \text{with } i \in \{D, \ell, *\},$$

yields

$$r_\ell > r_D > r_* = 0,$$

with $r_D = 0.0067$ and $r_\ell = 0.0255$.

10.5 Conclusion

The finite-horizon optimal control problem with input constraints for input-affine nonlinear systems has been studied. The problem has been solved by means of a dynamic extension yielding a combination of bang-bang signals and singular arcs. Simulations on a model of an industrial wastewater treatment plant and of a rocket in vertical flight have shown the performance of the dynamic control law, which are remarkably close to the optimal ones, although they do not rely on the solution of any HJB.

Conclusion

Summary of Contribution

On the topic of approximation of dynamical systems, the Thesis sheds new light on the problem of model reduction of time-delay systems. Novelties of the proposed approach include the possibility of obtaining reduced order models for linear time-delay systems with an arbitrary number and types of delay. The possibility of increasing the number of interpolation points by exploiting the delays is another unique result, for which no equivalent method exists in the literature. The extension of model reduction by moment matching to nonlinear time-delay systems is a new contribution.

Data-driven model reduction by moment matching is another novel result which has some similarities to proper orthogonal decomposition, but differs in derivation, meaning and application. The algorithms resulting from this contribution offer a method for the estimation of moment with unparalleled speed. The estimation of the moments from the time history of the output of the system is on its own a new concept.

The validation of these results in the field of power systems is a contribution which departs from the usual area of control theory. In fact, this work has been presented in a conference and submitted to a journal of the IEEE Power and Energy Society, demonstrating the impact of the model reduction techniques developed in the Thesis beyond the field of linear algebra and control theory.

Even though model reduction for linear singular systems is a well-studied problem, the partition of the projector into a fast and a slow part is a new result. However, the importance of this result is diminished by the fact that a practical application and several direct extensions have not been investigated yet.

The idea of characterizing “moments” for input signals which consist of an infinite number

of frequencies is another totally novel idea which allowed to extend the moment matching techniques to switching systems.

In addition to this direct consequence, this work laid the foundations to develop the concept of discontinuous phasor transform which is the first contribution in the area of analysis. Thus, a result that was born in the realm of model reduction had an impact on a totally different field. In fact, the work on the discontinuous phasor has been published in a journal belonging to the IEEE Power Electronics Society. This is another work which will impact a field beyond the one of control and systems.

The ideas behind the results presented in Chapter 9 are due to A. Astolfi and L. Praly. My contribution, developed in collaboration with them, has been to solve the irreducible case, to recognize the link to the Barbalat's lemma, to solve the triangular reducible block case and to present the results in a fairly comprehensible and unified way.

Finally, the last original contribution of the Thesis consists of a technique to determine approximate solutions of the HJB pde when the input of the system is constrained. Moreover, the distance from the optimal solution can be computed and a method to minimize this distance is given.

Future Research Directions

Given a set of interpolation points, the problem of determining the remaining of the parameters of the reduced order model to obtain an optimal model with respect to some error norm has eluded any attempt to a solution for a long time. Recent developments carried out at New York University in collaboration with Z.-P. Jiang have shed some light on this problem and work in this direction is underway.

The problem of model reduction for nonlinear singular systems is another research direction that must be investigated to extend and conclude the research started on linear singular systems.

Model reduction by moment matching can still be extended to several other classes of systems, such as hybrid systems and stochastic systems. Probably the most interesting extension is to hybrid systems, for which early contributions have already been submitted to two conferences [25, 27].

Another interesting and fundamental problem is the study, development and validation of reduced order models for control purposes. Results in this direction would further increase the impact of model reduction and may solve many problems for which the solution of the original control problem is difficult or computationally intractable (in a certain sense providing an alternative way to determine approximate solutions of optimal control problems).

Further extensions of the work regarding the discontinuous phasor transform can be easily imagined with the development of a phasor transform for nonlinear circuits powered first by sinusoidal sources and then by switching sources.

Bibliography

[1] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Englewood Cliffs: Prentice Hall, 2001.

[2] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, ser. Texts in Applied Mathematics. Springer New York, 1998.

[3] A. Astolfi, "Model reduction by moment matching for linear and nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2321–2336, 2010.

[4] T. Penzl, "Algorithms for model reduction of large dynamical systems," *Linear Algebra and its Applications*, vol. 415, no. 2-3, pp. 322 – 343, 2006, special Issue on Order Reduction of Large-Scale Systems.

[5] A. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA: SIAM Advances in Design and Control, 2005.

[6] J. P. Richard, "Time-delay systems: an overview of some recent advances and open problems," *Automatica*, vol. 39, no. 10, pp. 1667–1694, 2003.

[7] G. Scarciotti and A. Astolfi, *A review on model reduction by moment matching for nonlinear systems*. Invited book chapter. Book in honour of Laurent Praly (TBA), 2016.

[8] ——, "Model reduction by moment matching for linear time-delay systems," *19th IFAC World Congress, Cape Town, South Africa, August 24-29*, 2014.

[9] ——, "Model reduction by moment matching for nonlinear time-delay systems," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, 2014.

[10] ——, "Model reduction of neutral linear and nonlinear time-invariant time-delay systems with discrete and distributed delays," *To appear on IEEE Transactions on Automatic Control*, vol. 61, no. 6, 2016.

[11] ——, "Model reduction for linear systems and linear time-delay systems from input/output data," in *2015 European Control Conference, Linz, July*, 2015.

[12] ——, "Model reduction for nonlinear systems and nonlinear time-delay systems from input/output data," in *Proceedings of the 54th IEEE Conference on Decision and Control (to appear)*, 2015.

[13] ——, "Data-driven model reduction for linear and nonlinear, possibly time-delay, systems," *Submitted to Automatica*, 2016.

[14] G. Scarciotti, "Model reduction of power systems with preservation of slow and poorly damped modes," in *IEEE Power & Energy Society General Meeting, Denver, Colorado, July 26-30*, 2015.

[15] ——, "Low computational complexity model reduction of power systems with preservation of physical characteristics," *Submitted to IEEE Transactions on Power Systems*, 2016.

[16] G. Scarciotti and A. Astolfi, "Characterization of the moments of a linear system driven by explicit signal generators," in *Proceedings of the 2015 American Control Conference, Chicago, IL, July*, 2015, pp. 589–594.

[17] ——, "Model reduction by matching the steady-state response of explicit signal generators," *To appear on IEEE Transactions on Automatic Control*, vol. 61, no. 7, 2016.

[18] G. Scarciotti, "Model reduction for linear singular systems," in *Proceedings of the 54th IEEE Conference on Decision and Control (to appear)*, 2015.

[19] G. Scarciotti and A. Astolfi, "A note on the electrical equivalent of the moment theory," in *Proceedings of the 2016 American Control Conference, Boston (to appear)*, 2016.

[20] ——, "Moment based discontinuous phasor transform and its application to the steady-state analysis of inverters and wireless power transfer systems," *IEEE Transactions on Power Electronics (to appear)*, 2016.

[21] G. Scarciotti, L. Praly, and A. Astolfi, "A small-gain-like theorem for large-scale systems," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, Dec 2013, pp. 6593–6596.

[22] ——, "Invariance-like theorems and weak convergence properties," *IEEE Transactions on Automatic Control*, 2015.

[23] G. Scarciotti and A. Astolfi, "Approximate finite-horizon optimal control for input-affine nonlinear systems with input constraints," *IFAC Symposium Nonlinear Control Systems, Toulouse, France*, vol. 9, no. 1, pp. 199–204, 2013.

[24] ——, "Approximate finite-horizon optimal control for input-affine nonlinear systems with input constraints," *Journal of Control and Decision*, vol. 1, no. 2, pp. 149–165, 2014.

[25] ——, "Moments at "discontinuous signals" with applications: model reduction for hybrid systems and phasor transform for switching circuits," in *2nd International Symposium on Mathematical Theory of Networks and Systems, Minneapolis, MN, USA (submitted to)*, 2016.

[26] A. Padoan, G. Scarciotti, and A. Astolfi, "A geometric characterisation of persistently exciting signals generated by autonomous systems," in *IFAC Symposium Nonlinear Control Systems, Monterey, CA, USA (submitted to)*, 2016.

[27] G. Scarciotti and A. Astolfi, "Model reduction for hybrid systems with state-dependent jumps," in *IFAC Symposium Nonlinear Control Systems, Monterey, CA, USA (submitted to)*, 2016.

[28] K. J. Åström and P. R. Kumar, "Control: A perspective," *Automatica*, vol. 50, no. 1, pp. 3–43, 2014.

[29] V. M. Adamjan, D. Z. Arov, and M. G. Krein, "Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem," *Mathematics of the USSR Sbornik*, vol. 15, pp. 31–73, 1971.

[30] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds," *International Journal of Control*, vol. 39, no. 6, pp. 1115–1193, 1984.

[31] M. G. Safonov, R. Y. Chiang, and D. J. N. Limebeer, "Optimal Hankel model reduction for nonminimal systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 4, pp. 496–502, 1990.

[32] B. C. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Transactions on Automatic Control*, vol. 26, no. 1, pp. 17–32, 1981.

[33] D. G. Meyer, "Fractional balanced reduction: model reduction via a fractional representation," *IEEE Transactions on Automatic Control*, vol. 35, no. 12, pp. 1341–1345, 1990.

[34] W. S. Gray and J. Mesko, "General input balancing and model reduction for linear and nonlinear systems," in *European Control Conference, Brussels, Belgium*, 1997.

[35] S. Lall and C. Beck, "Error bounds for balanced model reduction of linear time-varying systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 946–956, 2003.

[36] H. Kimura, "Positive partial realization of covariance sequences," *Modeling, Identification and Robust Control*, pp. 499–513, 1986.

[37] C. I. Byrnes, A. Lindquist, S. V. Gusev, and A. S. Matveev, "A complete parameterization of all positive rational extensions of a covariance sequence," *IEEE Transactions on Automatic Control*, vol. 40, pp. 1841–1857, 1995.

[38] T. T. Georgiou, "The interpolation problem with a degree constraint," *IEEE Transactions on Automatic Control*, vol. 44, pp. 631–635, 1999.

[39] A. C. Antoulas, J. A. Ball, J. Kang, and J. C. Willems, "On the solution of the minimal rational interpolation problem," *Linear Algebra and Its Applications, Special Issue on Matrix Problems*, vol. 137-138, pp. 511–573, 1990.

[40] C. I. Byrnes, A. Lindquist, and T. T. Georgiou, "A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint," *IEEE Transactions on Automatic Control*, vol. 46, pp. 822–839, 2001.

[41] K. Gallivan, A. Vandendorpe, and P. Van Dooren, "Sylvester equations and projection-based model reduction," *Journal of Computational and Applied Mathematics*, vol. 162, no. 1, pp. 213–229, 2004.

[42] K. A. Gallivan, A. Vandendorpe, and P. Van Dooren, "Model reduction and the solution of Sylvester equations," in *MTNS, Kyoto*, 2006.

[43] C. A. Beattie and S. Gugercin, "Interpolation theory for structure-preserving model reduction," in *Proceedings of the 47th IEEE Conference on Decision and Control, Cancun, Mexico*, 2008.

[44] S. Gugercin, A. C. Antoulas, and C. Beattie, "$\mathcal{H}_2$ model reduction for large-scale linear dynamical systems," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 2, pp. 609–638, 2008.

[45] S. A. Al-Baiyat, M. Bettayeb, and U. M. Al-Saggaf, "New model reduction scheme for bilinear systems," *International Journal of Systems Science*, vol. 25, no. 10, pp. 1631–1642, 1994.

[46] S. Lall, P. Krysl, and J. Marsden, "Structure-preserving model reduction for mechanical systems," *Physica D*, vol. 184, pp. 304–318, 2003.

[47] J. Soberg, K. Fujimoto, and T. Glad, "Model reduction of nonlinear differential-algebraic equations," *IFAC Symposium Nonlinear Control Systems, Pretoria, South Africa*, vol. 7, pp. 712–717, 2007.

[48] K. Fujimoto, "Balanced realization and model order reduction for port-Hamiltonian systems," *Journal of System Design and Dynamics*, vol. 2, no. 3, pp. 694–702, 2008.

[49] J. M. A. Scherpen and W. S. Gray, "Minimality and local state decompositions of a nonlinear state space realization using energy functions," *IEEE Transactions on Automatic Control*, vol. 45, no. 11, pp. 2079–2086, Nov 2000.

[50] W. S. Gray and J. M. A. Scherpen, "Nonlinear Hilbert adjoints: properties and applications to Hankel singular value analysis," in *Proceedings of the 2001 American Control Conference*, vol. 5, 2001, pp. 3582–3587.

[51] E. Verriest and W. Gray, "Dynamics near limit cycles: Model reduction and sensitivity," in *Symposium on Mathematical Theory of Networks and Systems, Padova, Italy*, 1998.

[52] W. S. Gray and E. I. Verriest, "Balanced realizations near stable invariant manifolds," *Automatica*, vol. 42, no. 4, pp. 653–659, 2006.

[53] K. Kunisch and S. Volkwein, "Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition," *Journal of Optimization Theory and Applications*, vol. 102, no. 2, pp. 345–371, 1999.

[54] M. Hinze and S. Volkwein, "Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control," in *Dimension Reduction of Large-Scale Systems*, ser. Lecture Notes in Computational and Applied Mathematics. Springer, 2005, pp. 261–306.

[55] K. Willcox and J. Peraire, "Balanced model reduction via the proper orthogonal decomposition," *AIAA Journal*, vol. 40, no. 11, pp. 2323–2330, 2002.

[56] K. Kunisch and S. Volkwein, "Proper orthogonal decomposition for optimality systems," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 42, no. 01, pp. 1–23, 2008.

[57] P. Astrid, S. Weiland, K. Willcox, and T. Backx, "Missing point estimation in models described by proper orthogonal decomposition," *IEEE Transactions on Automatic Control*, vol. 53, no. 10, pp. 2237–2251, Nov 2008.

[58] S. Lall, J. E. Marsden, and S. Glavaski, "A subspace approach to balanced truncation for model reduction of nonlinear control systems," *International Journal on Robust and Nonlinear Control*, vol. 12, pp. 519–535, 2002.

[59] K. Fujimoto and D. Tsubakino, "Computation of nonlinear balanced realization and model reduction based on Taylor series expansion," *Systems & Control Letters*, vol. 57, no. 4, pp. 283–289, 2008.

[60] A. Astolfi, "A new look at model reduction by moment matching for linear systems," in *Proceedings of the 46th IEEE Conference on Decision and Control*, Dec 2007, pp. 4361–4366.

[61] W. Dib, A. Astolfi, and R. Ortega, "Model reduction by moment matching for switched power converters," in *Proceedings of the 48th IEEE Conference on Decision and Control, held jointly with the 28th Chinese Control Conference*, Dec 2009, pp. 6555–6560.

[62] T. C. Ionescu and A. Astolfi, "Families of reduced order models that achieve nonlinear moment matching," in *Proceedings of the 2013 American Control Conference, Washington, DC, USA, June 17-19*, 2013, pp. 5518–5523.

[63] T. C. Ionescu, A. Astolfi, and P. Colaneri, "Families of moment matching based, low order approximations for linear systems," *Systems & Control Letters*, vol. 64, pp. 47–56, 2014.

[64] A. C. Antoulas, "A new result on passivity preserving model reduction," *Systems & Control Letters*, vol. 54, no. 4, pp. 361–374, 2005.

[65] D. C. Sorensen, "Passivity preserving model reduction via interpolation of spectral zeros," *Systems & Control Letters*, vol. 54, no. 4, pp. 347–360, 2005.

[66] C. Hespel and G. Jacob, "Approximation of nonlinear dynamic systems by rational series," *Theoretical Computer Science*, vol. 79, no. 1, pp. 151–162, 1991.

[67] C. Hespel, "Truncated bilinear approximants: Carleman, finite Volterra, Padé-type, geometric and structural automata," in *Algebraic Computing in Control*, ser. Lecture

Notes in Control and Information Sciences, G. Jacob and F. Lamnabhi-Lagarrigue, Eds. Springer, 1991, vol. 165, pp. 264–278.

[68] P. Feldmann and R. W. Freund, "Efficient linear circuit analysis by Pade approximation via the Lanczos process," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 14, no. 5, pp. 639–649, May 1995.

[69] E. Grimme, D. Sorensen, and P. van Dooren, "Model reduction of state space systems via an implicitly restarted Lanczos method," *Numer. Algorithms*, vol. 12, pp. 1–31, 1995.

[70] A. C. Antoulas, "Polplatzierung bei der modellreduktion (on pole placement in model reduction)," *Automatisierungstechnik*, vol. 55, no. 9, pp. 443–448–374, 2009.

[71] T. T. Georgiou, "Partial realization of covariance sequences," Ph.D. dissertation, University of Florida, Gainesville, 1983.

[72] H. Kimura, "A canonical form for partial realization of covariance sequences," Technical Report 83-01, 1983.

[73] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory.* New York: Macmillan, 1992.

[74] S. Gugercin and K. Willcox, "Krylov projection framework for Fourier model reduction," *Automatica*, vol. 44, no. 1, pp. 209–215, 2008.

[75] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*, ser. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2011.

[76] L. Farina, "Is a system representable as a compartmental system?" in *1997 European Control Conference*, July 1997, pp. 18–20.

[77] A. Astol and P. Colaneri, "A note on the existence of positive realizations," *Linear Algebra and its applications*, vol. 390, pp. 329–343, 2004.

[78] A. Isidori, *Nonlinear Control Systems*, 3rd ed., ser. Communications and Control Engineering. Springer, 1995.

[79] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, ser. AIMS Series on Applied Mathematics. Philadelphia, PA: SIAM, 2007, vol. 2.

[80] J. Huang, *Nonlinear Output Regulation: Theory and Applications*, ser. International series in pure and applied mathematics. Philadelphia, PA: SIAM Advances in Design and Control, 2004.

[81] J. K. Hale, *Theory of functional differential equations*, ser. Applied Mathematical Sciences Series. Springer Verlag Gmbh, 1977.

[82] G. Stépán, *Retarded Dynamical Systems: Stability and Characteristic Functions*, ser. Pitman research notes in mathematics series. Longman Scientific & Technical, 1989.

[83] J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional Differential Equations*, ser. Applied Mathematical Sciences Series. New York:Springer, 1993, vol. 99.

[84] W. Michiels and S. I. Niculescu, *Stability and Stabilization of Time-Delay Systems: An Eigenvalue-Based Approach.* Philadelphia: SIAM, 2007.

[85] S. I. Niculescu, *Delay Effects on Stability.* Heidelberg: Springer, 2001.

[86] Q. C. Zhong, *Robust Control of Time-delay Systems.* Germany: Springer, 2006.

[87] N. Bekiaris-Liberis and M. Krstic, *Nonlinear Control Under Nonconstant Delays*, ser. Advances in Design and Control. SIAM, 2013.

[88] V. Kharitonov, "Robust stability analysis of time delay systems: A survey," *4th IFAC Conference on System Structure and Control, Nantes, France, 8-10 July*, vol. Penary lecture, pp. 1–12, 1998.

[89] V. B. Kolmanovskii, S. I. Niculescu, and K. Gu, "Delay effects on stability: A survey," *Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, AZ, December*, pp. 1993–1998, 1999.

[90] V. B. Kolmanovskii and A. Myshkis, *Introduction to the Theory and Applications of Functional Differential Equations*, ser. Mathematics and Its Applications. Springer, 1999.

[91] J. K. Hale and S. M. Verduyn Lunel, "Effects of small delays on stability and control," *Operator Theory: Advances and Applications*, vol. 122, pp. 275–301, 2001.

[92] S. I. Niculescu, A. Trofino Nito, J. M. Dion, and L. Dugard, "Delay-dependent stability of linear systems with delayed state: an LMI approach," in *Proceedings of the 34th IEEE Conference on Decision and Control*, vol. 2, Dec 1995, pp. 1495–1496.

[93] J. Zhang, C. R. Knospe, and P. Tsiotras, "New results for the analysis of linear systems with time-invariant delays," *International Journal of Robust and Nonlinear Control*, vol. 13, no. 12, pp. 1149–1175, 2003.

[94] A. W. Olbrot, "A sufficiently large time delay in feedback loop must destroy exponential stability of any decay rate," *IEEE Transactions on Automatic Control*, vol. 29, pp. 367–368, 1984.

[95] R. Datko, "A paradigm of ill-posedness with respect to time delays," *IEEE Transactions on Automatic Control*, vol. 43, no. 7, pp. 964–967, 1998.

[96] N. MacDonald, "Two delays may not destabilize although either delay can," *Math Biosciences*, vol. 82, pp. 127–140, 1986.

[97] J. Beddington and R. M. May, "Time lags are not necessarily destabilizing," *Math. Biosciences*, vol. 27, pp. 109–117, 1986.

[98] G. Abdallah, P. Dorato, J. Benitez-Read, and R. Byrne, "Delayed positive feedback can stabilize oscillatory systems," *Proceedings of the 1993 American Control Conference, San Francisco*, pp. 3106–3107, 1993.

[99] A. Goubet, M. Dambrine, and J. P. Richard, "An extension of stability criteria for linear and nonlinear time delay systems," *IFAC Conference on System Structure and Control, Nantes, France*, pp. 278–283, 1995.

[100] V. D. Blondel and A. Megretski, *Unsolved Problems in Mathematical Systems and Control Theory*. Princeton University Press, 2004.

[101] P. M. Mäkilä and J. R. Partington, "Laguerre and Kautz shift approximations of delay systems," *International Journal of Control*, vol. 72, pp. 932–946, 1999.

[102] ——, "Shift operator induced approximations of delay systems," *SIAM Journal of Control and Optimization*, vol. 37, no. 6, pp. 1897–1912, 1999.

[103] J. Zhang, C. R. Knospe, and P. Tsiotras, "Stability of linear time-delay systems: A delay-dependent criterion with a tight conservatism bound," *Proceedings of the 2000 American Control Conference, Chicago, IL, June*, pp. 1458–1462, 2000.

[104] S. H. Al-Amer and F. M. Al-Sunni, "Approximation of time-delay systems," *Proceedings of the 2000 American Control Conference, Chicago, IL, June*, pp. 2491–2495, 2000.

[105] H. T. Banks and F. Kappel, "Spline approximations for functional differential equations," *Journal of Differential Equations*, vol. 34, pp. 496–522, 1979.

[106] G. Gu, P. P. Khargonekar, and E. B. Lee, "Approximation of infinite-dimensional systems," *IEEE Transactions on Automatic Control*, vol. 34, no. 6, 1992.

[107] K. Glover, J. Lam, and J. R. Partington, "Rational approximation of a class of infinite dimensional system i: Singular value of hankel operator," *Mathematics of Control Circulation and Systems*, vol. 3, pp. 325–344, 1990.

[108] C. Glader, G. Hognas, P. M. Mäkilä, and H. T. Toivonen, "Approximation of delay systems: A case study," *International Journal of Control*, vol. 53, no. 2, pp. 369–390, 1991.

[109] Y. Ohta and A. Kojima, "Formulas for Hankel singular values and vectors for a class of input delay systems," *Automatica*, vol. 35, pp. 201–215, 1999.

[110] M. G. Yoon and B. H. Lee, "A new approximation method for time-delay systems," *IEEE Transactions on Automatic Control*, vol. 42, no. 7, pp. 1008–1012, 1997.

[111] W. Michiels, E. Jarlebring, and K. Meerbergen, "Krylov-based model order reduction of time-delay systems," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 4, pp. 1399–1421, 2011.

[112] E. Jarlebring, T. Damm, and W. Michiels, "Model reduction of time-delay systems using position balancing and delay Lyapunov equations," *Mathematics of Control, Signals, and Systems*, vol. 25, no. 2, pp. 147–166, 2013.

[113] Q. Wang, Y. Wang, E. Y. Lam, and N. Wong, "Model order reduction for neutral systems by moment matching," *Circuits, Systems, and Signal Processing*, vol. 32, no. 3, pp. 1039–1063, 2013.

[114] T. C. Ionescu and O. V. Iftime, "Moment matching with prescribed poles and zeros for infinite-dimensional systems," *American Control Conference, June, Montreal, Canada*, pp. 1412–1417, 2012.

[115] O. V. Iftime, "Block circulant and block Toeplitz approximants of a class of spatially distributed systems-An LQR perspective," *Automatica*, vol. 48, no. 12, pp. 3098–3105, dec 2012.

[116] P. Lancaster, *Theory of matrices*. New York: Academic, 1969.

[117] ——, "Explicit solutions of linear matrix equations," *SIAM Review, Vol 12, No. 4, October*, 1970.

[118] Y. Halevi, "Reduced-order models with delay," *International Journal of Control*, vol. 64, pp. 733–744, 1996.

[119] J. Carr, *Applications of Centre Manifold Theory*, ser. Applied Mathematical Sciences Series. Springer-Verlag, 1981, no. v. 35.

[120] T. C. Ionescu and A. Astolfi, "Families of moment matching based, structure preserving approximations for linear port Hamiltonian systems," *ArXiv e-prints*, Apr 2013.

[121] Y. Yamamoto, "Minimal representations for delay systems," in *Proceedings of the 17th IFAC World Congress, Seul, Korea, July 6-11*, 2008, pp. 1249–1254.

[122] C. I. Byrnes, M. W. Spong, and T. J. Tarn, "A several complex variables approach to feedback stabilization of linear neutral delay-differential systems," *Mathematical systems theory*, vol. 17, no. 1, pp. 97–133, 1984.

[123] J. K. Hale and S. M. Verduyn Lunel, "Strong stabilization of neutral functional differential equations," *IMA Journal of Mathematical Control and Information*, vol. 19, no. 1-2, pp. 5–23, 2002.

[124] H. Mounier, P. Rouchon, and J. Rudolph, "Some examples of linear systems with delays," *European Journal of Automatic Systems*, vol. 31, no. 6, pp. 911–926, 1997.

[125] J. K. Hale, "Behavior near constant solutions of functional differential equations," *Journal of differential equations*, vol. 15, pp. 278–294, 1974.

[126] V. B. Kolmanovskii and V. R. Nosov, *Stability of Functional Differential Equations*, ser. Mathematics in science and engineering. Elsevier Science, 1986.

[127] R. Hotzel and M. Fliess, "On linear systems with a fractional derivation: Introductory theory and examples," *Mathematics and Computers in Simulation*, vol. 45, no. 3-4, pp. 385–395, 1998.

[128] R. Curtain, O. Iftime, and H. Zwart, "A comparison between LQR control for a long string of SISO systems and LQR control of the infinite spatially invariant version," *Automatica*, vol. 46, no. 10, pp. 1604–1615, 2010.

[129] Z. Artstein, "Linear systems with delayed controls: A reduction," *IEEE Transactions on Automatic Control*, vol. 27, no. 4, pp. 869–879, Aug 1982.

[130] P. A. Ioannou and C. C. Chien, "Autonomous intelligent cruise control," *IEEE Transactions on Vehicular Technology*, vol. 42, pp. 657–672, 1993.

[131] S. Huang, "Automatic vehicle following with integrated throttle and brake control," *International Journal of Control*, vol. 72, pp. 45–83, 1999.

[132] R. H. Middleton and J. H. Braslavsky, "String instability in classes of linear time invariant formation control with limited communication range," *IEEE Transactions on Automatic Control*, vol. 55, no. 7, pp. 1519–1530, 2010.

[133] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 2, pp. 131–140, 1990.

[134] B. Saldivar, S. Mondié, J. J. Loiseau, and V. Rasvan, "Stick-slip oscillations in oillwell drilstrings: distributed parameter and neutral type retarded model approaches," *18th IFAC World Congress, Milano, Italy*, pp. 284–289, 2011.

[135] ——, "Exponential stability analysis of the drilling system described by a switched neutral type delay equation with nonlinear perturbations," *Proceedings of the 50th IEEE Conference on Decision and Control, and European Control Conference*, pp. 4164–4169, 2011.

[136] B. Saldivar and S. Mondié, "Drilling vibration reduction via attractive ellipsoid method," *Journal of the Franklin Institute*, vol. 350, no. 3, pp. 485–502, 2013.

[137] B. Salvidar, I. Boussaada, H. Mounier, S. Mondié, and S. I. Niculescu, "An overview on the modeling of oilwell drilling vibrations," *19th IFAC World Congress, Cape Town, South Africa, August 24-29*, 2014.

[138] N. Bekiaris-Liberis and M. Krstic, "Compensation of wave actuator dynamics for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1555–1570, June 2014.

[139] E. M. Navarro-López and D. Cortés, "Avoiding harmful oscillations in a drillstring through dynamical analysis," *Journal of Sound and Vibration*, vol. 307, no. 1-2, pp. 152–171, 2007.

[140] I. Boussaada, H. Mounier, S. I. Niculescu, and A. Cela, "Analysis of drilling vibrations: a time-delay system approach," *20th Mediterranean Conference on Control & Automation*, pp. 610–614, 2012.

[141] K. H. Lundberg, H. R. Miller, and D. L. Trumper, "Initial conditions, generalized functions, and the laplace transform - troubles at the origin," *Control Systems, IEEE*, vol. 27, no. 1, pp. 22–35, Feb 2007.

[142] Y. Jiang and Z. P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.

[143] T. Bian, Y. Jiang, and Z. P. Jiang, "Adaptive dynamic programming and optimal control of nonlinear nonaffine systems," *Automatica*, vol. 50, no. 10, pp. 2624–2632, 2014.

[144] Y. Jiang and Z. P. Jiang, "Robust adaptive dynamic programming and feedback stabilization of nonlinear systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 882–893, May 2014.

[145] L. C. Baird, "Reinforcement learning in continuous time: advantage updating," vol. 4, pp. 2448–2453, Jun 1994.

[146] D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F. L. Lewis, "Adaptive optimal control for continuous-time linear systems based on policy iteration," *Automatica*, vol. 45, no. 2, pp. 477–484, 2009.

[147] E. N. Lorenz, *Empirical Orthogonal Functions and Statistical Weather Prediction*, ser. Scientific report 1, Statistical Forecasting Project. MIT, Department of Meteorology, 1956.

[148] C. W. Rowley, T. Colonius, and R. M. Murray, "Model reduction for compressible flows using POD and Galerkin projection," *Physica D: Nonlinear Phenomena*, vol. 189, no. 1-2, pp. 115–129, 2004.

[149] B. R. Noack, K. Afanasiev, M. Morzynski, G. Tadmor, and F. Thiele, "A hierarchy of low-dimensional models for the transient and post-transient cylinder wake," *Journal of Fluid Mechanics*, vol. 497, pp. 335–363, 12 2003.

[150] K. J. Åström and B. Wittenmark, *Adaptive Control*, ser. Addison-Wesley series in electrical engineering: control engineering, 1995.

[151] T. Chen and B. A. Francis, *Optimal Sampled-data Control Systems*. Springer, 1995.

[152] T. N. E. Greville, "Some applications of the pseudoinverse of a matrix," *SIAM Rev. 2*, pp. 15–22, 1960.

[153] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, ser. CMS Books in Mathematics. Springer, 2003.

[154] Q. Wang and L. Zhang, "Online updating the generalized inverse of centered matrices," in *Proceedings of the 25th AAAI Conference on Artificial Intelligence*, 2011, pp. 1826–1827.

[155] F. Le Gall, "Powers of tensors and fast matrix multiplication," in *International Symposium on Symbolic and Algebraic Computation, Kobe, Japan, July 23-25*, 2014, pp. 296–303.

[156] H. Rodriguez, R. Ortega, and A. Astolfi, "Adaptive partial state feedback control of the DC-to-DC Ćuk converter," *Proceedings of the 2005 American Control Conference*, vol. 7, pp. 5121–5126, June 2005.

[157] J. H. Chow, *Power System Coherency and Model Reduction*, ser. Power Electronics and Power Systems. Springer, 2013.

[158] A. J. Germond and R. Podmore, "Dynamic aggregation of generating unit models," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-97, no. 4, pp. 1060–1069, July 1978.

[159] R. Podmore, "Identification of coherent generators for dynamic equivalents," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-97, no. 4, pp. 1344–1354, July 1978.

[160] B. Avramovic, P. V. Kokotovic, J. R. Winkelman, and J. H. Chow, "Area decomposition for electromechanical models of power systems," *Automatica*, vol. 16, no. 6, pp. 637–648, 1980.

[161] P. V. Kokotovic, B. Avramovic, J. H. Chow, and J. R. Winkelman, "Coherency based decomposition and aggregation," *Automatica*, vol. 18, no. 1, pp. 47–56, 1982.

[162] D. Chaniotis and M. A. Pai, "Model reduction in power systems using Krylov subspace methods," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 888–894, May 2005.

[163] C. Sturk, L. Vanfretti, Y. Chompoobutrgool, and H. Sandberg, "Coherency-independent structured model reduction of power systems," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2418–2426, Sept 2014.

[164] A. K. Singh and B. C. Pal, "Report on the 68-bus, 16-machine, 5-area system," IEEE PES Task Force on Benchmark Systems for Stability Controls. Ver. 3.3, Dec [Online] http://www.sel.eesc.usp.br/ieee/, 2013.

[165] K. A. Gallivan, A. Vandendorpe, and P. Van Dooren, "Model reduction of MIMO systems via tangential interpolation," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 2, pp. 328–349, 2004.

[166] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed. Society for Industrial and Applied Mathematics, 2003.

[167] P. Sauer and A. Pai, *Power System Dynamics and Stability*. Prentice Hall, 1998.

[168] P. M. Anderson and A. A. A. Fouad, *Power system control and stability*, 2nd ed., ser. Power Engineering Series. IEEE Press, 2003.

[169] L. A. Zadeh and C. A. Desoer, *Linear system theory: The state space approach*, ser. McGraw-Hill series in System Science. McGraw-Hill, 1963.

[170] R. E. Kalman, P. L. Falb, and M. A. Arbib, *Topics in mathematical system theory*, ser. International series in pure and applied mathematics. McGraw-Hill, 1969.

[171] S. H. Wu, R. B. Sherwood, and R. Covey, "Block II Apollo digital reaction control systems study," TRW Space Technical Laboratory Report 373D-6004-RU000, June, 1965.

[172] A. Khayatian and D. G. Taylor, "Multirate modeling and control design for switched-mode power converters," *IEEE Transactions on Automatic Control*, vol. 39, no. 9, pp. 1848–1852, 1994.

[173] H. Sira-Ramirez and L. S. Orestes, "On the dynamical pulse-width-modulation control of robotic manipulator systems," *International Journal of Robust Nonlinear Control*, vol. 6, no. 6, 1996.

[174] D. G. Taylor, "Pulse-width modulated control of electromechanical systems," *IEEE Transactions on Automatic Control*, vol. 37, no. 4, pp. 524–528, 1992.

[175] S. Bittanti and P. Colaneri, *Periodic Systems: Filtering and Control*, ser. Communications and Control Engineering. Springer, 2008.

[176] R. W. Brockett, *Finite dimensional linear systems*, ser. Series in Decision and Control. Wiley, 1970.

[177] O. M. Grasselli, L. Menini, and S. Galeani, *Sistemi dinamici. Introduzione all'analisi e primi strumenti di controllo.* Milano: Hoepli, 2008.

[178] A. Saberi, P. Sannuti, and A. Stoorvogel, *Control of Linear Systems with Regulation and Input Constraints.* London, UK: Springer-Verlag, 2000.

[179] B. A. Francis, "The linear multivariable regulator problem," in *IEEE Conference on Decision and Control including the 15th Symposium on Adaptive Processes*, Dec 1976, pp. 873–878.

[180] L. Dai, *Singular Control Systems*, ser. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, 1989.

[181] P. Kunkel and V. L. Mehrmann, *Differential-algebraic Equations: Analysis and Numerical Solution*, ser. EMS textbooks in mathematics. European Mathematical Society, 2006.

[182] F. L. Lewis, M. A. Christodoulou, B. G. Mertzios, and K. Ozcaldiran, "Chained aggregation of singular systems," *IEEE Transactions on Automatic Control*, vol. 34, no. 9, pp. 1007–1012, Sep 1989.

[183] K. Perev and B. Shafai, "Balanced realization and model reduction of singular systems," *International Journal of Systems Science*, vol. 25, no. 6, pp. 1039–1052, 1994.

[184] W. Q. Liu and V. Sreeram, "Model reduction of singular systems," in *Proceedings of the 39th IEEE Conference on Decision and Control*, vol. 3, 2000, pp. 2373–2378.

[185] P. J. Kootsookos and R. R. Bitmead, "The Nehari shuffle and minimax FIR filter design," *Control and Dynamic Systems series*, vol. 64, no. 1, pp. 239–298, 1994.

[186] J. Wang, Q. Zhang, W. Liu, and V. Sreeram, "Model reduction of singular systems via covariance approximation," in *Proceedings of the 2004 American Control Conference, Boston, MA, June*, vol. 1, 2004, pp. 90–95.

[187] A. B. H. Adamou-Mitiche, L. Mitiche, and V. Sima, "Model reduction for descriptor systems," in *First International Symposium on Control, Communications and Signal Processing*, March 2004, pp. 827–830.

[188] S. Gugercin, T. Stykel, and S. Wyatt, "Model reduction of descriptor systems by interpolatory projection methods," *SIAM Journal on Scientific Computing*, vol. 35, no. 5, pp. B1010–B1033, 2013.

[189] M. K. Kazimierczuk and D. Czarkowski, *Resonant Power Converters*, 2nd ed. Wiley, 2012.

[190] M. K. Kazimierczuk, *Pulse-width Modulated DC-DC Power Converters*. Wiley, 2008.

[191] R. Middlebrook and S. Cuk, "A general unified approach to modelling switching-converter power stages," in *IEEE Power Electronics Specialists Conference*, June 1976, pp. 18–34.

[192] R. W. Erickson and D. Maksimović, *Fundamentals of Power Electronics*, ser. Electrical science series. Springer US, 2001.

[193] V. Vorperian, "Simplified analysis of PWM converters using model of PWM switch. part I: Continuous conduction mode," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 3, pp. 490–496, May 1990.

[194] ——, "Simplified analysis of PWM converters using model of PWM switch. part II: Discontinuous conduction mode," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 3, pp. 497–505, May 1990.

[195] H. Mao, D. Boroyevich, and F. C. Y. Lee, "Novel reduced-order small-signal model of a three-phase PWM rectifier and its application in control design and system analysis," *IEEE Transactions on Power Electronics*, vol. 13, no. 3, pp. 511–521, May 1998.

[196] A. M. Davis, *Linear Circuit Analysis*, ser. Electrical Engineering Series. PWS Pub., 1998.

[197] J. W. Nilsson and S. A. Riedel, *Electric Circuits*, 10th ed. Pearson/Prentice Hall, 2008.

[198] C. T. Rim and G. H. Cho, "Phasor transformation and its application to the DC/AC analyses of frequency phase-controlled series resonant converters (SRC)," *IEEE Transactions on Power Electronics*, vol. 5, no. 2, pp. 201–211, Apr 1990.

[199] Y. Yin, R. Zane, J. Glaser, and R. W. Erickson, "Small-signal analysis of frequency-controlled electronic ballasts," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 50, no. 8, pp. 1103–1110, Aug 2003.

[200] C. T. Rim, "Unified general phasor transformation for AC converters," *IEEE Transactions on Power Electronics*, vol. 26, no. 9, pp. 2465–2475, Sept 2011.

[201] S. Ben-Yaakov, S. Glozman, and R. Rabinovici, "Envelope simulation by SPICE-compatible models of linear electric circuits driven by modulated signals," *IEEE Transactions on Industry Applications*, vol. 37, no. 2, pp. 527–533, Mar 2001.

[202] S. Y. R. Hui, Wenxing Zhong, and C. K. Lee, "A critical review of recent progress in mid-range wireless power transfer," *IEEE Transactions on Power Electronics*, vol. 29, no. 9, pp. 4500–4511, Sept 2014.

[203] J. Yin, D. Lin, C. K. Lee, T. Parisini, and S. Y. R. Hui, "Front-end monitoring of multiple loads in wireless power transfer systems without wireless communication systems," *IEEE Transactions on Power Electronics*, 2015.

[204] U. A. Bakshi and A. P. Godse, *Analog Electronic Circuits*. Technical Publications, 2009.

[205] W. M. Moussa and J. E. Morris, "Comparison between state space averaging and PWM switch for switch mode power supply analysis," in *Proceedings of the 1990 IEEE Southern Tier Technical Conference*, Apr 1990, pp. 15–21.

[206] C. T. Rim, D. Y. Hu, and G. H. Cho, "Transformers as equivalent circuits for switches: general proofs and D-Q transformation-based analyses," *IEEE Transactions on Industry Applications*, vol. 26, no. 4, pp. 777–785, Jul 1990.

[207] Diodes Inc., "BAS19 datasheet," http://www.diodes.com/_files/datasheets/ds12004.pdf, accessed: 2015-11-30.

[208] M. Pinuela, D. C. Yates, S. Lucyszyn, and P. D. Mitcheson, "Maximizing DC-to-load efficiency for inductive power transfer," *IEEE Transactions on Power Electronics*, vol. 28, no. 5, pp. 2437–2447, May 2013.

[209] W. Hahn and A. P. Baartz, *Stability of motion*, ser. Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete. Springer, 1967.

[210] N. Rouche, P. Habets, and M. Laloy, *Stability Theory by Liapunov's Direct Method*. New York-Heidelberg-Berlin: Springer-Verlag, 1977.

[211] J. P. La Salle and Z. Artstein, *The Stability of Dynamical Systems*, ser. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1987.

[212] A. Lyapunov, *The General Problem of the Stability of Motion*, ser. Control Theory and Applications Series. Taylor & Francis Group, 1992.

[213] J. P. La Salle, "Some extensions of liapunov's second method," *IRE Trans. Circuit Theory*, vol. CT-7, pp. 520–527, 1960.

[214] J. P. La Salle and S. Lefschetz, *Stability by Lyapunov's Direct Method*. New York: Academic, 1961.

[215] J. P. La Salle, "An invariance principle in the theory of stability," in *Differential Equations and Dynamical Systems*, J. Hale and J. La Salle, Eds. New York: Academic, 1967, pp. 277–286.

[216] F. Mazenc and M. Malisoff, "Strict Lyapunov function constructions under LaSalle conditions with an application to Lotka-Volterra systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 841–854, 2010.

[217] V. Matrosov, "On the stability of motion," *Journal of Applied Mathematics and Mechanics*, vol. 26, no. 5, pp. 1337–1353, 1962.

[218] A. Loria, E. Panteley, D. Popovic, and A. R. Teel, "A nested Matrosov theorem and persistency of excitation for uniform convergence in stable nonautonomous systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 2, pp. 183–198, 2005.

[219] F. Mazenc and D. Nesic, "Lyapunov functions for time-varying systems satisfying generalized conditions of Matrosov theorem," *Mathematics in Control Signals and Systems*, vol. 19, pp. 151–182, 2007.

[220] A. Astolfi and L. Praly, "A LaSalle version of Matrosov theorem," *Proceedings of the 50th IEEE Conference on Decision and Control, and European Control Conference*, pp. 320–324, 2011.

[221] I. Karafyllis and Z. P. Jiang, *Stability and Stabilization of Nonlinear Systems*, ser. Communications and Control Engineering. Springer, 2011.

[222] Z. P. Jiang, I. M. Y. Mareels, and Y. Wang, "A Lyapunov formulation of the nonlinear small-gain theorem for interconnected ISS systems," *Automatica*, vol. 32, no. 8, pp. 1211–1215, 1996.

[223] M. Vidyasagar, *Input-Output Analysis of Large-Scale Interconnected Systems.* Springer-Verlag, 1981.

[224] C. A. Desoer and M. Vidyasagar, *Feedback systems: input-output properties*, ser. Electrical science series. Academic Press, 1975.

[225] I. Y. Tyukin, E. Steur, H. Nijmeijer, and C. van Leeuwen, "Non-uniform small-gain theorems for systems with unstable invariant sets." *SIAM Journal on Control and Optimization*, vol. 47, no. 2, pp. 849–882, 2008.

[226] A. J. Van der Schaft, *$\mathcal{L}_2$-Gain and Passivity Techniques in Nonlinear Control.* Springer-Verlag, 1996.

[227] A. Isidori, *Nonlinear Control Systems II.* Springer-Verlag, 1999.

[228] D. Hill and P. Moylan, "The stability of nonlinear dissipative systems," *IEEE Transactions on Automatic Control*, vol. 21, no. 5, pp. 708–711, 1976.

[229] E. D. Sontag, "Smooth stabilization implies coprime factorization," *IEEE Transactions on Automatic Control*, vol. 34, no. 4, pp. 435–443, 1989.

[230] E. D. Sontag and Y. Wang, "New characterizations of input-to-state stability," *IEEE Transactions on Automatic Control*, vol. 41, no. 9, pp. 1283–1294, 1996.

[231] H. Ito, "Stability criteria for interconnected iISS systems and ISS systems using scaling of supply rates," *Proceedings of the 2004 American Control Conference*, vol. 2, pp. 1055–1060, 2004.

[232] ——, "State-dependent scaling problems and stability of interconnected iISS and ISS systems," *IEEE Transactions on Automatic Control*, vol. 51, no. 10, pp. 1626–1643, 2006.

[233] D. Angeli and A. Astolfi, "A tight small gain theorem for not necessarily ISS systems," *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference*, pp. 5427–5431, 2005.

[234] S. N. Dashkovskiy, B. S. Rüffer, and F. R. Wirth, "An ISS small-gain theorem for general networks," *Mathematics of Control, Signals, and Systems*, vol. 19, no. 2, pp. 93–122, 2007.

[235] ——, "Small gain theorems for large scale systems and construction of ISS Lyapunov functions," *SIAM J. Control and Optimization*, vol. 48, no. 6, pp. 4089–4118, 2010.

[236] A. Astolfi and L. Praly, "A weak version of the small-gain theorem," *Proceedings of the 51st IEEE Conference on Decision and Control*, pp. 4586–4590, 2012.

[237] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, ser. Classics in Applied Mathematics. SIAM, 1994, vol. 9.

[238] E. Cesàro, "Sur la multiplication des series," *Bull. Sci. Math*, vol. 14, no. 2, pp. 114–120, 1890.

[239] K. Mischaikow, H. Smith, and H. R. Thieme, "Asymptotically autonomous semiflows: chain recurrence and Lyapunov functions," *Transactions of the American Mathematical Society*, vol. 347, no. 5, pp. 1669–1685, 1995.

[240] H. R. Thieme, "Asymptotically autonomous differential equations in the plane," *Rocky Mountain Journal of Mathematics*, vol. 24, no. 1, pp. 351–380, 1994.

[241] E. D. Sontag, "A remark on the converging-input converging-state property," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 313–314, 2003.

[242] A. R. Teel and J. Hespanha, "Examples of GES systems that can be driven to infinity by arbitrarily small additive decaying exponentials," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, pp. 1407–1410, 2004.

[243] A. Astolfi, "A remark on an example by Teel-Hespanha with applications to cascaded systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 2, pp. 289–293, 2007.

[244] D. Liberzon, *Calculus of Variations and Optimal Control Theory: A Concise Introduction.* Princeton University Press, 2012.

[245] M. Sassano and A. Astolfi, "Dynamic approximate solutions of the HJ inequality and of the HJB equation for input-affine nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 10, pp. 2490–2503, 2012.

[246] ——, "Approximate finite-horizon optimal control without PDE's," *Systems & Control Letters*, vol. 62, no. 2, pp. 97–103, 2013.

[247] R. W. Beard, G. N. Saridis, and J. T. Wen, "Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation," *Automatica*, vol. 33, no. 12, pp. 2159–2177, 1997.

[248] J. A. Moreno, "Optimal time control of bioreactors for the wastewater treatment," *Optimal Control Applications and Methods*, vol. 20, no. 3, pp. 145–164, 1999.

[249] P. Tsiotras and H. J. Kelley, "Goddard problem with constrained time of flight," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 15, no. 2, pp. 289–296, 1992.

[250] K. M. Anstreicher and M. H. Wright, "A note on the augmented hessian when the reduced hessian is semidefinite," *SIAM Journal on Optimization*, vol. 11, no. 1, pp. 243–253, 2000.

[251] R. Antonelli and A. Astolfi, "Nonlinear proportional integral controls robustly stabilise biological reactors," *IFAC Symposium on Advanced Control of Chemical Processes, Pisa, Italy*, 2000.

[252] M. J. Betancur, J. A. Moreno, I. Moreno-Andrade, and G. Buitrón, "Practical optimal control of fed-batch bioreactors for the waste water treatment," *International Journal of Robust and Nonlinear Control*, vol. 16, no. 3, pp. 173–190, 2006.

[253] A. Rapaport and D. Dochain, "Minimal time control of fed-batch processes with growth functions having several maxima," *IEEE Transactions on Automatic Control*, vol. 56, no. 11, pp. 2671–2676, 2011.

[254] H. Munick, "Goddard problem with bounded thrust," *AIAA Journal*, vol. 3, no. 7, pp. 1283–1285, 1964.

[255] K. Graichen and N. Petit, "Solving the goddard problem with thrust and dynamic pressure constraints using saturation functions," *Proceedings of the 17th IFAC World Congress, Seoul, South Korea, July 6-11*, pp. 14 301–14 306, 2008.