Chapter 2
A Review on Model Reduction by Moment Matching for Nonlinear Systems

Giordano Scarciotti and Alessandro Astolfi

Abstract The model reduction problem for nonlinear systems and nonlinear time-delay systems based on the steady-state notion of moment is reviewed. We show how this nonlinear description of moment is used to pose and solve the model reduction problem by moment matching for nonlinear systems, to develop a notion of frequency response for nonlinear systems, and to solve model reduction problems in the presence of constraints on the reduced order model. Model reduction of nonlinear time-delay systems is then discussed. Finally, the problem of approximating the moment of nonlinear, possibly time-delay, systems from input/output data is briefly illustrated.

2.1 Introduction

The model reduction problem has been widely studied for 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 [1]. The model reduction problem consists in finding a simplified description of a dynamical system maintaining at the same time specific properties. For linear system, the problem has been extensively studied exploiting a variety of techniques, some of them based on the singular value decomposition, see, e.g., [2–4] which make use of Hankel operators or, e.g., [5–8] which exploit balanced realizations, and some based on the Krylov projec-

---

Dedicated to Laurent: a pioneer in the land of control

G. Scarciotti · A. Astolfi (✉)
Imperial College London, London SW7 2AZ, UK
e-mail: a.astolfi@ic.ac.uk

G. Scarciotti
e-mail: gs3610@ic.ac.uk

A. Astolfi
University of Rome "Tor Vergata", Via Del Politecnico 1, 00133 Rome, Italy

tion matrices, see, e.g., [9–15], also called moment matching methods. 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 [16–19]. Energy-based methods have been proposed in [7, 20, 21]. Other techniques, based on the reduction around a limit cycle or a manifold, have been presented in [22, 23]. Model reduction methods based on proper orthogonal decomposition have been developed for linear and nonlinear systems, see, e.g., [24–28]. Finally, note that some computational aspects have been investigated in [23, 26, 29, 30]. In addition, 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 [31] and several results have been given using rational interpolations, see, e.g., [32–34], see also [35–41]. Recent results include model order reduction techniques for linear time-delay systems, see, e.g., [42–44], and for infinite dimensional systems, see, e.g., [45, 46] in which operators are used to provide reduced order models for linear systems.   The goal of this chapter is to review the model reduction techniques for nonlinear, possibly time-delay, systems based on the "steady-state" notion of moment. We start introducing the interpolation approach to moment matching, which is how moment matching has been classically interpreted and applied to linear systems. We then move to the steady-state approach introduced in [47]. We present some results on the model reduction problem by moment matching for nonlinear systems, as given in [48], and develop a notion of frequency response for nonlinear systems. These techniques are extended to nonlinear time-delay systems [49] and the problem of obtaining a family of reduced order models matching two (nonlinear) moments is solved for a special class of signal generators. Finally the problem of approximating the moment of nonlinear (time-delay) systems, without solving the partial differential equation that defines it, is presented and solved [50, 51].

Notation. We use standard notation. $\mathbb{R}_{>0}$ denotes the set of positive real numbers; $\mathbb{C}_{<0}$ denotes the set of complex numbers with negative real part; $\mathbb{D}_{<1}$ denotes the set of complex numbers with modulo smaller than one; ι denotes the imaginary unit. Given a set of delays $\{\tau_j\}$, the symbol $\mathfrak{R}_T^n = \mathfrak{R}_T^n([-T, 0], \mathbb{R}^n)$, 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 [52]. The symbol I denotes the identity matrix, $\sigma(A)$ denotes the spectrum of the matrix $A \in \mathbb{R}^{n \times n}$ and $\otimes$ indicates the Kronecker product. The vectorization of a matrix $A \in \mathbb{R}^{n \times m}$, denoted by vec($A$), is the $nm \times 1$ vector obtained by stacking the columns of the matrix A one on top of the other, namely 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 complex conjugate transposition operator. 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 $\det(\bar{s}I - A(\bar{s})) = 0$, $\bar{s} \in \mathbb{C}_{<0}$. $L_f h$ denotes the Lie derivative of the smooth function h along the smooth vector field f, as defined in [53, Chapter 1].

2.2   The Interpolation Approach

In this section we briefly recall the notion of moment and the related model reduction techniques as presented in [1]. 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 in the complex plane, which has been solved by the Nevanlinna-Pick theory (see, e.g., [54]).

Definition 2.1 Let $\{s_i\}$ be a sequence of distinct points in $Z \subset \mathbb{C}$ and let $\{w_i\}$ be an arbitrary sequence of points in $\mathbb{C}$. Given a space $\mathcal{W}$ of functions on Z, the *interpolation problem* consists in determining a function $W : Z \mapsto \mathbb{C}$ such that $W(s_i) = w_i$, for all $i = 1, \ldots, \nu$.

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 [1, Chapter 11]), provided it exists.

Definition 2.2 Let $s_i \in \mathbb{C} \setminus \sigma(A)$. The 0-*moment of system* (2.1) *at* s_i is the complex number $\eta_0(s_i) = W(s_i)$. The k-*moment of system* (2.1) *at* s_i is the complex number

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

with $k \geq 1$ integer.

In the interpolation approach to moment matching, a reduced order model is such that its transfer function (and, possibly, derivatives of this) takes the same values of the transfer function (and, possibly, 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 $s_i = 30\iota$. Since a minimal system can be entirely described by its transfer 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 correct 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
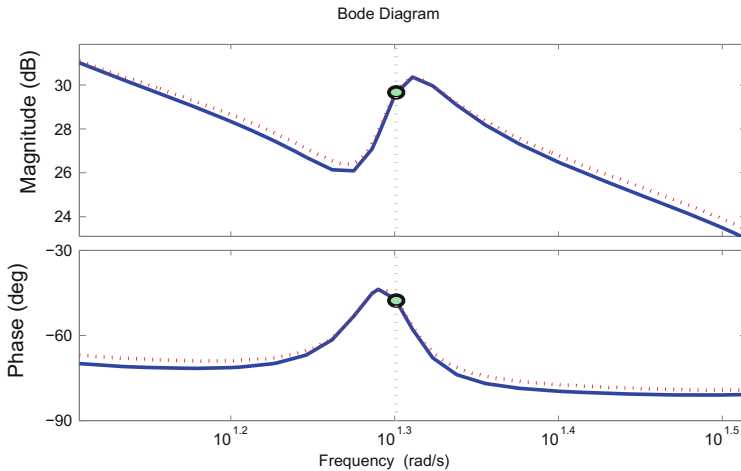
Fig. 2.1 Diagrammatic illustration of the interpolation approach. Magnitude (*top graph*) and phase (*bottom graph*) plot of a given system (*solid/blue line*) and of a reduced order model (*dashed/red line*). The green circle represents the interpolation point

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

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 system at a set of points s_i. The problem of model reduction by moment matching using the Petrov-Galerking projectors is thoroughly described in [1] and it is the subject of intensive research, see, e.g., [9–15]. Herein we report a few results which are instrumental for the aims of the chapter. We invite the reader to refer to [1] for additional detail.

Proposition 2.1 [1] *Consider ν distinct points $s_j \in \mathbb{C} \setminus \sigma(A)$, with $j = 1, \ldots, \nu$. The transfer function of the reduced order model (2.2), with*

$$V = \left[(s_1I - A)^{-1}B \cdots (s_\nu I - A)^{-1}B \right] \tag{2.4}$$

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 2.2 [1] *Consider the point $s_0 \in \mathbb{C} \setminus \sigma(A)$. The transfer function of the reduced order model (2.2), with*

$$V = \left[(s_0I - A)^{-1}B \, (s_0I - A)^{-2}B \cdots (s_0I - A)^{-\nu}B \right] \tag{2.5}$$

a generalized reachability matrix and W any left inverse of V, interpolates the trans-
fer function of system (2.1) and its $v - 1$ 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. 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 2.3 [1] *Consider $s_j \in \mathbb{C} \setminus \sigma(A)$, with $j = 1, \ldots, 2v$, the generalized reachability matrix*

$$\bar{V} = \left[(s_1 I - A)^{-1} B \cdots (s_v I - A)^{-1} B \right], \tag{2.6}$$

and the generalized observability matrix

$$\bar{W} = \left[(s_{v+1} I - A^*)^{-1} C^* \cdots (s_{2v} I - A^*)^{-1} C^* \right]. \tag{2.7}$$

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

Exploiting this result, the problem of preservation of passivity and stability has been solved in [55, 56], as reported here.

Lemma 2.1 [1] *If the interpolation points in Proposition 2.3 are chosen so that s_j, with $j = 1, \ldots, v$, are stable spectral zeros, i.e., they are such that $W^*(-s_i) + W(s_i) = 0$, and $s_{j+v} = -s_j$, with $j = 1, \ldots, v$, i.e., the interpolation points are chosen as zeros of the spectral factors and their mirror images, then the projected system is both stable and passive.*

We can now indicate the following drawbacks in the Krylov methods.

- 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.
- When a method capable of preserving some of these properties (such as stability and passivity) is presented, it usually implies that specific moments are matched. Hence, the designer cannot chose arbitrary moments. Moreover, there is a lack of system theoretic understanding behind why a particular interpolation point is related to a property like passivity.
- In Lemma 2.1 all the free parameters (the matrix W) are used and no additional property can be preserved.
- Finally, 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.

A possible solution to these issues is offered by the "steady-state-based" approach to moment matching. While the first three points are addressed in [48], we focus the rest of the chapter on the last problem: the model reduction of general classes of nonlinear systems.

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, [57, 58] for some results on the interpolation problem for nonlinear systems). In [48] (see also [14, 59]) a characterization of moment for system (2.1) has been given in terms of the solution of a Sylvester equation as follows.

Lemma 2.2 [48] *Consider system (2.1), $s_i \in \mathbb{C} \setminus \sigma(A)$, for all $i = 1, \dots, \eta$. There exists a one-to-one[1] 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 $C\Pi$, 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 has resulted in several developments in the area of model reduction by moment matching, see, e.g., [60, 61] and [49–51, 62–66], is that it establishes, through the Sylvester equation (2.8), a relation between the moments and the steady-state response of the output of the system. Before proceeding further we provide a formal definition of steady-state response. With abuse of notation, we indicate the state of a (linear, nonlinear, or more general) dynamical system as $x(t, x_0)$ to highlight the dependency on time and on the initial condition.

Definition 2.3 ([67, 68]) Let $\mathcal{B} \subset \mathbb{R}^n$ and suppose $x(t, x_0)$ is defined for all $t \geq 0$ and all $x_0 \in \mathcal{B}$. The *ω-limit set of the set* denoted by $w(\mathcal{B})$, is the set of all points x for which there exists a sequence of pairs $\{x_k, t_k\}$, with $x_k \in \mathcal{B}$ and $\lim_{k \to \infty} t_k = \infty$ such that $\lim_{k \to \infty} x(t_k, x_k) = x$.

---

[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.
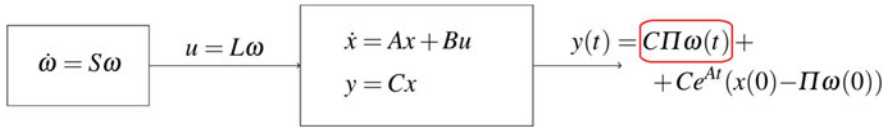
Fig. 2.2 Diagrammatic illustration of Theorem 2.1. The term denoting the steady-state response is circled

Definition 2.4 ([67, 68]) Suppose the responses of the system, with initial conditions in a closed and positively invariant set $\mathscr{X}$, are ultimately bounded. A *steady-state response* is any response with initial condition $x_0 \in w(\mathscr{B})$.

Exploiting the notion of steady-state response we can introduce the following result, which is illustrated in Fig. 2.2.

Theorem 2.1 [48] *Consider system (2.1), $s_i \in \mathbb{C} \setminus \sigma(A)$, for all $i = 1, \dots, \eta$, 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 $\eta_0(s_1)$, …, $\eta_{k_1-1}(s_1)$, …, $\eta_0(s_\eta)$, …, $\eta_{k_\eta-1}(s_\eta)$, and the steady-state response of the output y of such interconnected system.

Remark 2.1 [69] The minimality of the triple $(L, S, \omega(0))$ implies the observability of the pair (L, S) and the "controllability" of the pair $(S, \omega(0))$. This last condition, called *excitability* of the pair $(S, \omega(0))$, is a geometric characterization of the property that the signals generated by (2.10) are persistently exciting, see [70].

Remark 2.2 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 [50] the problem of computing the moments of an unknown linear systems from input/output data has been addressed. 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, [48, 49]. This fact also highlights how important for the moment matching techniques is to let the designer choose the interpolation points, which are related to the class of inputs to the system.

2.4 Model Reduction by Moment Matching for Nonlinear Systems

We can now extend the steady-state description of moment to nonlinear systems.[3]
Consider a nonlinear, single-input, single-output, continuous-time system described
by the equations

$$\dot{x} = f(x, u), \qquad y = h(x), \tag{2.11}$$

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

$$\dot{\omega} = s(\omega), \qquad u = l(\omega), \tag{2.12}$$

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

$$\dot{\omega} = s(\omega), \qquad \dot{x} = f(x, l(\omega)), \qquad y = h(x). \tag{2.13}$$

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 (2.11) with the signal generator cap-
tures 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 2.1 The signal generator (2.12) is observable, i.e., for any pair of ini-
tial 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[4] with $\omega(0) \neq 0$.

Assumption 2.2 The zero equilibrium of the system $\dot{x} = f(x, 0)$ is locally exponen-
tially stable.

Lemma 2.3 [48] *Consider system (2.11) and the signal generator (2.12). Suppose
Assumptions 2.1 and 2.2 hold. Then there is a unique mapping π, locally defined in
a neighborhood of $\omega = 0$, which solves the partial differential equation*

$$\frac{\partial \pi}{\partial \omega} s(\omega) = f(\pi(\omega), l(\omega)). \tag{2.14}$$

Remark 2.3 Lemma 2.3 implies that the interconnected system (2.13) possesses an
invariant manifold described by the equation $x = \pi(\omega)$.

Definition 2.5 Consider system (2.11) and the signal generator (2.12). Suppose
Assumption 2.1 holds. The function $h \circ \pi$, with π solution of equation (2.14), is the
moment of system (2.11) *at* (s, l).

---

[3]Note that the results of this section are local.

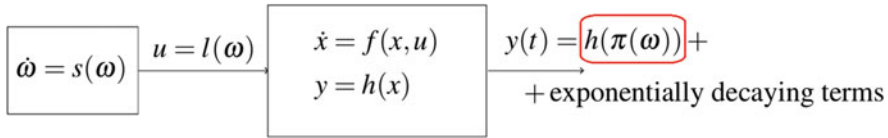[4]See [53, Chapter 8] for the definition of Poisson stability.

Fig. 2.3 Diagrammatic illustration of Theorem 2.2. The term denoting the steady-state response is circled

Theorem 2.2 [48] *Consider system (2.11) and the signal generator (2.12). Suppose Assumptions 2.1 and 2.2 hold. Then the moment of system (2.11) at (s, l) coincides with the steady-state response of the output of the interconnected system (2.13).*

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

Remark 2.4 [48] If the equilibrium $x = 0$ of the system $\dot{x} = f(x, 0)$ is unstable, it is still possible to define the moment of system (2.11) at (s, l) in terms of the function $h \circ \pi$, provided the equilibrium $x = 0$ is hyperbolic and the system (2.12) is Poisson stable, although it is not possible to establish a relation with the steady-state response of the interconnected system (2.13).

Remark 2.5 [48] 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 be difficult, or impossible, to provide general statements if the signal u, generated by system (2.12), is unbounded. Therefore, we assume that the signal generator generates bounded signals. For linear systems this assumption implies that we consider only points $s_i \in \mathbb{C}$ that are distinct and with zero real part.

2.4.1 The Frequency Response of a Nonlinear System

In [48], see also [71, 72], a nonlinear enhancement of the notion of frequency response of a linear system has been derived exploiting the steady-state description of moment. Note that this result is loosely related to the analysis in [66] where a generalization of the phasor transform based on the notion of moment is proposed. Consider system (2.11) and the signal generator (2.12). Let the signal generator (2.12) be such that

$$s(\omega) = \begin{bmatrix} 0 & \bar{\omega} \\ -\bar{\omega} & 0 \end{bmatrix} \omega, \qquad l(\omega) = \begin{bmatrix} L_1 & L_2 \end{bmatrix} \omega,$$

with $\omega(0) \neq 0$, $\bar{\omega} \neq 0$, and $L_1^2 + L_2^2 \neq 0$. Then, under Assumptions 2.1 and 2.2 the output of the interconnected system (2.13) converges toward a locally well-defined steady-state response, which, by definition, does not depend upon the initial condi-

tion $x(0)$. Moreover, such a steady-state response is periodic, hence, if it has the same period of $l(\omega(t))$, it can be written in Fourier series as $h(\pi(\omega(t))) = \sum_{k=-\infty}^{\infty} c_k e^{\iota k \bar{\omega} t}$, with $c_k \in \mathbb{C}$. Consider now the operator $\mathscr{P}_+$ which acts on a Fourier series as follows

$$\mathscr{P}_+\left(\sum_{k=-\infty}^{\infty} c_k e^{\iota k \bar{\omega} t}\right) = \sum_{k=0}^{\infty} \alpha_k e^{\iota k \bar{\omega} t},$$

with $\alpha_k \in \mathbb{C}$. With this operator we can define the frequency response of the nonlinear system (2.11) as

$$F(t, \omega(0), \bar{\omega}) = \frac{\mathscr{P}_+(h(\pi(\omega(t))))}{\mathscr{P}_+(l(\omega(t)))}.$$

This function depends upon the frequency $\bar{\omega}$, just as in the linear case, and, unlike the linear case, upon the initial condition $\omega(0)$ of the signal generator and time. Note finally that if the system (2.11) were linear, hence described by the Eq. (2.1), then $F(t, \omega(0), \bar{\omega})$ would be constant with respect to t and equal to $|W(\iota\bar{\omega})|e^{\iota\angle W(\iota\bar{\omega})}$, where $W(s) = C(sI - A)^{-1}B$, $|\cdot|$ indicates the absolute value operator and $\angle$ the phase operator.

2.4.2 Moment Matching

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

Definition 2.6 [48] Consider the signal generator (2.12). The system described by the equations

$$\dot{\xi} = \phi(\xi, u), \qquad \psi = \kappa(\xi), \tag{2.15}$$

with $\xi(t) \in \mathbb{R}^\nu$, is a *model at (s, l) of system* (2.11) if system (2.15) has the same moment at (s, l) as (2.11). In this case, system (2.15) is said to *match* the moment of system (2.11) at (s, l). Furthermore, system (2.15) is a reduced order model of system (2.11) if $\nu < n$.

Lemma 2.4 *Consider system (2.11), system (2.15) and the signal generator (2.12). Suppose Assumptions 2.1 and 2.2 hold. System (2.15) matches the moments of (2.11) at (s, l) if the equation*

$$\phi(p(\omega), l(\omega)) = \frac{\partial p}{\partial \omega} s(\omega) \tag{2.16}$$

has a unique solution p such that

$$h(\pi(\omega)) = \kappa(p(\omega)), \tag{2.17}$$

where π is the (unique) solution of equation (2.14).

In other words, we have to determine mappings ϕ, κ, and p such that Eqs. (2.16) and (2.17) hold. We introduce the following assumption to simplify the problem.

Assumption 2.3 There exist mappings κ and p such that $\kappa(0) = 0$, $p(0) = 0$, p is locally continuously differentiable, Eq. (2.17) 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 2.6 [48] Similar to the linear case, Assumption 2.3 holds selecting $p(\omega) = \omega$ and $k(\omega) = h(\pi(\omega))$.

Finally, as shown in [48], the system described by the equations

$$\dot{\xi} = s(\xi) - \delta(\xi)l(\xi) + \delta(\xi)u, \qquad \psi = h(\pi(\xi)), \qquad (2.18)$$

where δ is any mapping such that the equation

$$\frac{\partial p}{\partial \omega}s(\omega) = s(p(\omega)) - \delta(p(\omega))l(p(\omega)) + \delta(p(\omega))l(\omega), \qquad (2.19)$$

has the unique solution $p(\omega) = \omega$, is a family of *reduced order models of* (2.11) *at* (s, l).

2.4.3   Model Reduction by Moment Matching with Additional Properties

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 [48].

2.4.3.1   Matching with Asymptotic Stability

Consider the problem of determining a reduced order model (2.18) 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.

2.4.3.2   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 2.3 [48] *For all $r \in [1, \nu]$ there exists a δ such that system (2.18) has relative degree r at $\bar{\xi}$ if and only if the codistribution*

$$d\mathcal{O}_\nu(\xi) = \text{span}\{dh(\pi(\xi)), \cdots, dL_s^{\nu-1}h(\pi(\xi))\} \tag{2.20}$$

has dimension ν at $\bar{\xi}$.

2.4.3.3 Matching with Prescribed Zero Dynamics

Consider system (2.18) and the problem of determining the mapping δ such that the model has zero dynamics with specific properties. If $\bar{\xi}$ is an equilibrium of system (2.18), the problem is solved selecting δ such that the codistribution (2.20) has dimension ν at $\bar{\xi}$ [48]. Then there is a δ such that the zero dynamics of system (2.18) have a locally exponentially stable equilibrium and there is a coordinate transformation, locally defined around $\bar{\xi}$, such that the zero dynamics are described by the equations

$$\begin{aligned}
\dot{z}_1 &= z_2 + \hat{\delta}_1(z)z_1, \\
\dot{z}_2 &= z_3 + \hat{\delta}_2(z)z_1, \\
&\vdots \\
\dot{z}_{\nu-r} &= \hat{f}(z) + \hat{\delta}_{\nu-r}(z)z_1,
\end{aligned} \tag{2.21}$$

where the $\hat{\delta}_i$ are free functions and

$$\hat{f}(z) = \tilde{f}(\mathscr{Z})|_{\mathscr{Z}=[0,\ldots,0,z_1,\ldots,z_{\nu-r}]^\top},$$

with $\mathscr{Z} = \Xi(\xi)$ and $\tilde{f}(\mathscr{Z}) = L_s^\nu h(\pi(\Xi^{-1}(\mathscr{Z})))$.

2.4.3.4 Matching with a Passivity Constraint

Consider now the problem of selecting the mapping δ such that system (2.18) is lossless or passive. For such a problem the following fact holds.

Theorem 2.4 [48] *The family of reduced order models (2.18) contains, locally around $\bar{\xi}$, a lossless (passive, respectively) system with a differentiable storage function if there exists a differentiable function V, locally positive definite around $\bar{\xi}$, such that equation[5]*

$$V_\xi s(\xi) = h(\pi(\xi))l(\xi), \quad (V_\xi s(\xi) \leq h(\pi(\xi))l(\xi) \text{ respectively}), \tag{2.22}$$

holds locally around $\bar{\xi}$ and

---

[5]V_ξ and $V_{\xi\xi}$ denote, respectively, the gradient and the Hessian matrix of the scalar function V : $\xi \mapsto V(\xi)$.

$$V_{\xi\xi}(\xi) > 0. \tag{2.23}$$

2.4.3.5   Matching with L_2-gain

We now consider the problem of selecting the mapping δ such that system (2.18) has a given L_2-gain.

Theorem 2.5 *[48] The family of reduced order models (2.18) contains, locally around $\bar{\xi}$, a system with L_2-gain not larger than $\ell > 0$, and with a differentiable storage function if there exists a differentiable function V, locally positive definite around $\bar{\xi}$, such that Eq. (2.23) holds and*

$$V_{\xi}s(\xi) + (h(\pi(\xi)))^2 \le \ell^2 l^2(\xi), \tag{2.24}$$

holds locally around $\bar{\xi}$.

2.5   Model Reduction for Nonlinear Time-Delay Systems

Exploiting the steady-state notion of moment an extension of the model reduction method for nonlinear time-delay systems is given. 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.

2.5.1   Definition of π: Nonlinear Time-Delay Systems

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

$$\begin{aligned}
\dot{x} &= f(x_{\tau_0}, \dots, x_{\tau_\varsigma}, u_{\tau_{\varsigma+1}}, \dots, u_{\tau_\mu}), \qquad y = h(x), \\
x(\theta) &= \phi(\theta), \qquad\qquad\qquad -T \le \theta \le 0,
\end{aligned} \tag{2.25}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $\phi \in \mathfrak{R}_T^n$, $\tau_0 = 0$, $\tau_j \in \mathbb{R}_{>0}$ with $j = 1, \dots, \mu$ and f and h smooth mappings. Consider a signal generator (2.12) and the interconnected system

$$\dot{\omega} = s(\omega), \qquad \dot{x} = f(x_{\tau_0}, \dots, x_{\tau_\varsigma}, l(\omega_{\tau_{\varsigma+1}}), \dots, l(\omega_{\tau_\mu})), \qquad y = h(x). \tag{2.26}$$

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

Assumption 2.4 The zero equilibrium of the system $\dot{x} = f(x_{\tau_0}, \dots, x_{\tau_\varsigma}, 0, \dots, 0)$ is locally exponentially stable.

Lemma 2.5 [49, 53] *Consider system (2.25) and the signal generator (2.12). Suppose Assumptions 2.1 and 2.4 hold. Then there exists a unique mapping π, locally defined in a neighborhood of $\omega = 0$, which solves the partial differential equation*

$$\frac{\partial \pi}{\partial \omega} s(\omega) = f(\pi(\bar{\omega}_{\tau_0}), \dots, \pi(\bar{\omega}_{\tau_\varsigma}), l(\bar{\omega}_{\tau_{\varsigma+1}}), \dots, l(\bar{\omega}_{\tau_\mu})), \tag{2.27}$$

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

Remark 2.7 Lemma 2.5 implies that the interconnected system (2.26) possesses an invariant manifold, described by the equation $x = \pi(\omega)$. Note that the partial differential equation (2.27) is independent of time (as (2.14) in the delay-free case), e.g., if $s(\omega) = S\omega$ then $\bar{\omega}_{\tau_i} = e^{-S\tau_i}\omega$.

Definition 2.7 Consider system (2.25) and the signal generator (2.12). Suppose Assumption 2.1 holds. The function $h \circ \pi$, with π solution of equation (2.27), is the *moment of system* (2.25) *at* (s, l).

Theorem 2.6 [49] *Consider system (2.25) and the signal generator (2.12). Suppose Assumptions 2.1 and 2.4 hold. Then the moment of system (2.25) at (s, l) coincides with the steady-state response of the output of the interconnected system (2.26).*

2.5.2 Reduced Order Models for Nonlinear Time-Delay Systems

In this section two families of models achieving moment matching are given.

Definition 2.8 Consider system (2.25) and the signal generator (2.12). Suppose Assumption 2.1 and 2.4 hold. Then the system

$$\dot{\xi} = \phi(\xi_{\chi_0}, \dots, \xi_{\chi_{\hat{\rho}}}, u_{\chi_{\hat{\rho}+1}}, \dots, u_{\chi_\rho}), \qquad \psi = \kappa(\xi), \tag{2.28}$$

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 $j = 1, \dots, \rho$, and ϕ and κ smooth mappings, is a *model of system* (2.25) *at* (s, l) if system (2.28) has the same moment of system (2.25) at (s, l).

Lemma 2.6 *Consider system (2.25) and the signal generator (2.12). Suppose Assumption 2.1 and 2.4 hold. Then the system (2.28) is a model of system (2.25) at (s, l) if the equation*

$$\frac{\partial p}{\partial \omega} s(\omega) = \phi(p(\bar{\omega}_{\chi_0}), \dots, p(\bar{\omega}_{\chi_{\hat{\rho}}}), l(\bar{\omega}_{\chi_{\hat{\rho}+1}}), \dots, l(\bar{\omega}_{\chi_\rho})), \tag{2.29}$$

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

$$h(\pi(\omega)) = \kappa(p(\omega)), \tag{2.30}$$

where π is the unique solution of (2.27). System (2.28) is a reduced order model of system (2.25) at (s, l) if $\nu < n$, or if $\hat{\rho} < \varsigma$, or if $\rho < \mu$.

Similarly to the delay-free case we use part of the free mappings to obtain a simpler family of models.

Assumption 2.5 There exist mappings κ and p such that $\kappa(0) = 0$, $p(0) = 0$, p is locally continuously differentiable, Eq. (2.30) holds and p has a local inverse p^{-1}.

Consistently with Lemma 2.6, a family of models that achieves moment matching at (s, l) is described by

$$
\begin{aligned}
\dot{\xi} &= \Phi(\xi, \bar{\xi}_{\chi_1}, \dots, \bar{\xi}_{\chi_{\hat{\rho}}}) + \tfrac{\partial p(\omega)}{\partial \omega} \gamma(\xi_{\chi_1}, \dots, \xi_{\chi_{\hat{\rho}}}) + \tfrac{\partial p(\omega)}{\partial \omega} \sum_{j=\hat{\rho}+1}^{\rho} \delta_j(\xi) u_{\chi_j}, \\
\psi &= \kappa(\xi),
\end{aligned}
\tag{2.31}
$$

with

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

where $\bar{\xi}_{\chi_j} = \left[\bar{\omega}_{\chi_j} \right]_{\omega = p^{-1}(\xi)}$, κ and p are such that Assumption 2.5 holds, p is the unique solution of (2.29) and δ_j and γ are free mappings.

Assumption 2.5 holds with the selection $p(\omega) = \omega$ and $\kappa(\omega) = h(\pi(\omega))$. This yields a family of models described by the equations

$$
\begin{aligned}
\dot{\xi} &= s(\xi) - \sum_{j=\hat{\rho}+1}^{\rho} \delta_j(\xi) l(\bar{\xi}_{\chi_j}) - \gamma(\bar{\xi}_{\chi_1}, \dots, \bar{\xi}_{\chi_{\hat{\rho}}}) + \gamma(\xi_{\chi_1}, \dots, \xi_{\chi_{\hat{\rho}}}) + \sum_{j=\hat{\rho}+1}^{\rho} \delta_j(\xi) u_{\chi_j}, \\
\psi &= h(\pi(\xi)),
\end{aligned}
\tag{2.32}
$$

where δ_j and γ are arbitrary mappings such that Eq. (2.29), namely

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

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

The nonlinear model (2.32) has several free design parameters, namely δ_j, γ, χ_j, $\hat{\rho}$ and ρ. We note that selecting $\gamma \equiv 0$, $\hat{\rho} = 0$, $\rho = 1$ and $\chi_1 = 0$ (in this case we define $\delta = \delta_1$), yields a family of reduced order models with no delays. This family coincides with the family (2.18) and all results of Sect. 2.4.3 are directly applicable: the mapping δ can be selected to achieve matching with asymptotic stability, matching

with prescribed relative degree, etc. However, note that the choice of eliminating the delays may destroy some important dynamics of the model.

Remark 2.8 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

$$
\begin{aligned}
d(\dot{x}_{\tau_0}, \dots, \dot{x}_{\tau_{\varsigma_1}}) &= f(x_{\tau_{\varsigma_1+1}}, \dots, x_{\tau_{\varsigma_2}}, u_{\tau_{\varsigma_2+1}}, \dots, u_{\tau_\mu}), \\
y &= h(x),
\end{aligned}
\tag{2.33}
$$

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 $j = 1, \dots, \mu$ 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 [73–75] and references therein. Note, however, that for the simple case

$$
\begin{aligned}
\dot{x} + D\dot{x}_{\tau_1} &= f(x_{\tau_2}, \dots, x_{\tau_{\varsigma_1}}, u_{\tau_{\varsigma_1+1}}, \dots, u_{\tau_\mu}), \\
y &= h(x),
\end{aligned}
\tag{2.34}
$$

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 $\sigma(D) \subset \mathbb{D}_{<1}$.

2.5.3 Exploiting One Delay to Match $h \circ \pi_a$ and $h \circ \pi_b$

In this section 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 (2.25) and, to simplify the exposition, the signal generators described by the linear equation

$$
\dot{\omega} = S_a \omega, \qquad\qquad u = L_{ab}\omega, \tag{2.35}
$$

Note that, as highlighted in [48], considering the model reduction problem for non-linear systems when the signal generator is a linear system is of particular interest since the reduced order models have a very simple description, i.e., a family of reduced order models is described by a linear differential equation with a nonlinear output map. 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. This structure has two main advantages. Firstly, the selection of the free parameters that achieve additional goals, such as to assign the eigenvalues or the relative degree of the reduced order model, is remarkably simplified. Secondly, the computation of the reduced order model boils down to the computation of the output map $h \circ \pi$. A technique to approximate this mapping

is proposed in the next section. As a consequence of this discussion, a reduced order model of system (2.25) at (S_a, L_{ab}) is given by the family

$$
\begin{aligned}
\dot{\xi} &= F_0\xi + F_1\xi_\chi + G_2 u + G_3 u_\chi, \\
\psi &= \kappa_0(\xi) + \kappa_1(\xi_\chi),
\end{aligned}
\tag{2.36}
$$

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

$$
\begin{aligned}
F_0 P_a &+ F_1 P_a e^{-S_a\chi} - P_a S_a = -G_2 L_{ab} - G_3 L_{ab} e^{-S_a\chi}, \\
h(\pi_a(\omega)) &= \kappa_0(P_a\omega) + \kappa_1(P_a e^{-S_a\chi}\omega),
\end{aligned}
\tag{2.37}
$$

Consider now another signal generator described by the linear equation

$$
\dot{\omega} = S_b\omega, \qquad\qquad u = L_{ab}\omega,
\tag{2.38}
$$

and the problem of selecting F_0, F_1, G_2, G_3, κ_0, and κ_1 such that the reduced order model (2.36) matches the moments of system (2.25) at (S_a, L_{ab}) and (S_b, L_{ab}).

Proposition 2.4 *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_{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 (2.27), with $L = L_{ab}$ and $S = S_a$, and let $\pi_b(\omega) = \pi(\omega)$ be the unique solution of (2.27), with $L = L_{ab}$ and $S = S_b$. Then system (2.36) with the selection*

$$
\begin{aligned}
F_1 &= (S_b - S_a - G_3(e^{-S_b\chi} - e^{-S_a\chi_3}))(e^{-S_b\chi} - e^{-S_a\chi})^{-1}, \\
F_0 &= S_a - G_2 L_{ab} - G_3 L_{ab} e^{-S_a\chi} - F_1 e^{-S_a\chi}, \\
\kappa_0(\omega) &= h(\pi_a(\omega)) - \kappa_1(e^{-S_a\chi}\omega),
\end{aligned}
\tag{2.39}
$$

and k_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 (2.25) achieving moment matching at (S_a, L_{ab}) and (S_b, L_{ab}), for any G_2 and G_3 such that $s_i \notin \sigma(F_0 + F_1 e^{-s\chi})$, for all $s_i \in \sigma(S_a)$ and $s_i \in \sigma(S_b)$.

Proof As showed in the proof of Proposition 2.1 of [49], F_0 and F_1 solve the two Sylvester equations

$$
\begin{aligned}
F_0 P_a &+ F_1 P_a e^{-S_a\chi} - P_a S_a = -G_2 L_{ab} - G_3 L_{ab} e^{-S_a\chi}, \\
F_0 P_b &+ F_1 P_b e^{-S_b\chi} - P_b S_b = -G_2 L_{ab} - G_3 L_{ab} e^{-S_b\chi},
\end{aligned}
\tag{2.40}
$$

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

$$h(\pi_a(\omega)) = \kappa_0(\omega) + \kappa_1 \left(e^{-S_a \chi}\omega\right),$$
$$h(\pi_b(\omega)) = \kappa_0(\omega) + \kappa_1 \left(e^{-S_b \chi}\omega\right). \tag{2.41}$$

Solving the first equation with respect to δ_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 2.4 matches the moments $h \circ \pi_a$ and $h \circ \pi_b$ of the nonlinear system (2.25). Note that the matrices G_2 and G_3 remain free parameters and they can be used to achieve the properties discussed in Sect. 2.4.3. For instance, G_2 and G_3 can be used to set both the eigenvalues of F_0 and F_1.

Remark 2.9 Proposition 2.4 can be generalized to $\hat{\rho} > 1$ delays, obtaining a reduced order model that match $(\hat{\rho} + 1)\nu$ 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 $\gamma(\xi_{\chi_1}, \ldots, \xi_{\chi_{\hat{\rho}}})$ in (2.32) is replaced by $\hat{\gamma}_1(\xi_{\chi_1}) + \ldots + \hat{\gamma}_{\hat{\rho}}(\xi_{\chi_{\hat{\rho}}})$.

Remark 2.10 The number of delays in (2.25) does not play a role in Proposition 2.4. 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. This fact can be taken to the "limit" reducing a system which 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 $\hat{\rho}$ achieving moment matching at $(\hat{\rho} + 1)\nu$ moments.

2.6 Online Nonlinear Moment Estimation from Data

In this section we solve a fundamental problem for the theory we have presented, namely how to compute an approximation of the moment $h \circ \pi$ when the solution of the partial differential equation (2.13) or (2.27) is not known. Note, first of all, that the results of this section hold indiscriminately for delay-free and time-delay systems. In the following we do not even need to know the mappings f and h. In fact we are going to present a method to approximate the moment $h \circ \pi$ directly from input/output data, namely from $\omega(t)$ and $y(t)$. Note that given the exponential stability hypothesis on the system and Theorem 2.2 (Theorem 2.6 for time-delay systems), the equation

$$y(t) = h(\pi(\omega(t))) + \varepsilon(t), \tag{2.42}$$

where $\varepsilon(t)$ is an exponentially decaying signal, holds for the interconnections (2.13) and (2.26). We introduce the following assumption.

Assumption 2.6 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 $j = 1, \dots, M$ (M may be ∞), i.e., there exist $\pi_j \in \mathbb{R}$, with $j = 1, \dots, M$, such that

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

for any ω.

Let

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

with $N \leq M$. Using a weighted sum of basis functions, Eq. (2.42) can be written as

$$y(t) = \sum_{j=1}^{N} \pi_j \varphi_j(\omega(t)) + e(t) + \varepsilon(t) = \Gamma \Omega(\omega(t)) + e(t) + \varepsilon(t), \tag{2.43}$$

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

$$y(t) \approx \sum_{j=1}^{N} \widetilde{\pi}_j \varphi_j(\omega(t)) = \widetilde{\Gamma} \Omega(\omega(t)), \tag{2.44}$$

which neglects the approximation error $e(t)$ and the transient error $\varepsilon(t)$. Let $T_k^w = \{t_{k-w+1}, \dots, t_{k-1}, t_k\}$, with $0 \leq t_0 < t_1 < \dots < t_{k-w} < \dots < t_k < \dots < t_q$, with $w > 0$ and $q \geq w$, and 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 $\varepsilon(t)$ are known. Since this is not the case in practice, define $\widetilde{\Gamma}_k = \begin{bmatrix} \widetilde{\pi}_1 & \widetilde{\pi}_2 & \dots & \widetilde{\pi}_N \end{bmatrix}$ as the approximation, in the sense of (2.44), of the estimate Γ_k. Finally, we can compute this approximation as follows.

Theorem 2.7 [64] *Define the time-snapshots* $\widetilde{U}_k \in \mathbb{R}^{w \times N}$ *and* $\widetilde{Y}_k \in \mathbb{R}^w$ *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

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

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

$$\mathrm{vec}(\widetilde{\Gamma}_k) = (\widetilde{U}_k^\top \widetilde{U}_k)^{-1} \widetilde{U}_k^\top \widetilde{Y}_k, \tag{2.45}$$

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 2.7 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 2.7 holds (see [76] for a similar argument), $\widetilde{U}_k^\top \widetilde{U}_k$ is full rank. The next definition is a direct consequence of the discussion we have carried out.

Definition 2.9 The *estimated moment* of system (2.11) (or system (2.25)) is defined as

$$\widetilde{h \circ \pi}_{N,k}(\omega(t)) = \widetilde{\Gamma}_k \Omega(\omega(t)), \tag{2.46}$$

with $\widetilde{\Gamma}_k$ computed with (2.45).

Equation (2.45) is a classic least-square estimator and an efficient recursive formula can be easily derived.

Theorem 2.8 [64] *Assume that* $\Phi_k = (\widetilde{U}_k^\top \widetilde{U}_k)^{-1}$ *and* $\Psi_k = (\widetilde{U}_{k-1}^\top \widetilde{U}_{k-1} + \omega(t_k)\omega(t_k)^\top)^{-1}$ *are full rank for all* $t \geq t_r$ *with* $t_r \geq t_w$. *Given* $\mathrm{vec}(\widetilde{\Gamma}_r)$, Φ_r *and* Ψ_r, *the least-square estimation*

$$\mathrm{vec}(\widetilde{\Gamma}_k) = \mathrm{vec}(\widetilde{\Gamma}_{k-1}) + \Phi_k \omega(t_k)\Big(y(t_k) - \omega(t_k)^\top \mathrm{vec}(\widetilde{\Gamma}_{k-1})\Big) - \\ - \Phi_k \omega(t_{k-w})\Big(y(t_{k-w}) - \omega(t_{k-w})^\top \mathrm{vec}(\widetilde{\Gamma}_{k-1})\Big), \tag{2.47}$$

with

$$\Phi_k = \Psi_k - \Psi_k \omega(t_{k-w})(I + \omega(t_{k-w})^\top \Psi_k \omega(t_{k-w}))^{-1} \omega(t_{k-w})^\top \Psi_k \tag{2.48}$$

and

$$\Psi_k = \Phi_{k-1} - \Phi_{k-1}\omega(t_k)(I + \omega(t_k)^\top \Phi_{k-1}\omega(t_k))^{-1}\omega(t_k)^\top \Phi_{k-1}. \tag{2.49}$$

holds for all $t \geq t_r$.

Finally, the following result guarantees that the approximation converges to $h \circ \pi$.

Theorem 2.9 [64] *Suppose Assumptions 2.1 (2.1 for time-delay systems), 2.2 (2.4 for time-delay systems), 2.6 and 2.7 hold. Then*

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

2.7 Conclusion

In this chapter we have reviewed the model reduction technique for nonlinear, possibly time-delay, systems based on the "steady-state" notion of moment. We have firstly 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 nonlinear systems has been given and an enhancement of the notion of frequency response for nonlinear systems has been presented. Subsequently, these techniques have been extended to nonlinear time-delay systems and the problem of obtaining a family of reduced order models matching two moments has been solved for nonlinear time-delay systems. The review is concluded with a recently presented technique to approximate the moment of nonlinear, possibly time-delay, systems, without solving any partial differential equation.

References

1. Antoulas, A.: Approximation of Large-Scale Dynamical Systems. SIAM Advances in Design and Control, Philadelphia, PA (2005)
2. Adamjan, V.M., Arov, D.Z., Krein, M.G.: Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem. Math. USSR Sb. **15**, 31–73 (1971)
3. Glover, K.: All optimal Hankel-norm approximations of linear multivariable systems and their L∞-error bounds. Int. J. Control **39**(6), 1115–1193 (1984)
4. Safonov, M.G., Chiang, R.Y., Limebeer, D.J.N.: Optimal Hankel model reduction for nonminimal systems. IEEE Trans. Autom. Control **35**(4), 496–502 (1990)
5. Moore, B.C.: Principal component analysis in linear systems: controllability, observability, and model reduction. IEEE Trans. Autom. Control **26**(1), 17–32 (1981)
6. Meyer, D.G.: Fractional balanced reduction: model reduction via a fractional representation. IEEE Trans. Autom. Control **35**(12), 1341–1345 (1990)
7. Gray, W.S., Mesko, J.: General input balancing and model reduction for linear and nonlinear systems. In: European Control Conference. Brussels, Belgium (1997)
8. Lall, S., Beck, C.: Error bounds for balanced model reduction of linear time-varying systems. IEEE Trans. Autom. Control **48**(6), 946–956 (2003)
9. Kimura, H.: Positive partial realization of covariance sequences. In: Modeling, Identification and Robust Control, pp. 499–513 (1986)
10. Byrnes, C.I., Lindquist, A., Gusev, S.V., Matveev, A.S.: A complete parameterization of all positive rational extensions of a covariance sequence. IEEE Trans. Autom. Control **40**, 1841–1857 (1995)
11. Georgiou, T.T.: The interpolation problem with a degree constraint. IEEE Trans. Autom. Control **44**, 631–635 (1999)
12. Antoulas, A.C., Ball, J.A., Kang, J., Willems, J.C.: On the solution of the minimal rational interpolation problem. In: Linear Algebra and its Applications, Special Issue on Matrix Problems, vol. 137–138, pp. 511–573 (1990)
13. Byrnes, C.I., Lindquist, A., Georgiou, T.T.: A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint. IEEE Trans. Autom. Control **46**, 822–839 (2001)
14. Gallivan, K.A., Vandendorpe, A., Van Dooren, P.: Model reduction and the solution of Sylvester equations. In: MTNS, Kyoto (2006)

15. Beattie, C.A., Gugercin, S.: Interpolation theory for structure-preserving model reduction. In: Proceedings of the 47th IEEE Conference on Decision and Control, Cancun, Mexico (2008)
16. Al-Baiyat, S.A., Bettayeb, M., Al-Saggaf, U.M.: New model reduction scheme for bilinear systems. Int. J. Syst. Sci. **25**(10), 1631–1642 (1994)
17. Lall, S., Krysl, P., Marsden, J.: Structure-preserving model reduction for mechanical systems. Phys. D **184**, 304–318 (2003)
18. Soberg, J., Fujimoto, K., Glad, T.: Model reduction of nonlinear differential-algebraic equations. In: IFAC Symposium Nonlinear Control Systems, Pretoria, South Africa, vol. 7, pp. 712–717 (2007)
19. Fujimoto, K.: Balanced realization and model order reduction for port-Hamiltonian systems. J. Syst. Des. Dyn. **2**(3), 694–702 (2008)
20. Scherpen, J.M.A., Gray, W.S.: Minimality and local state decompositions of a nonlinear state space realization using energy functions. IEEE Trans. Autom. Control **45**(11), 2079–2086 (2000)
21. Gray, W.S., Scherpen, J.M.A.: Nonlinear Hilbert adjoints: properties and applications to Hankel singular value analysis. In: Proceedings of the 2001 American Control Conference, vol. 5, pp. 3582–3587 (2001)
22. Verriest, E., Gray, W.: Dynamics near limit cycles: model reduction and sensitivity. In: Symposium on Mathematical Theory of Networks and Systems, Padova, Italy (1998)
23. Gray, W.S., Verriest, E.I.: Balanced realizations near stable invariant manifolds. Automatica **42**(4), 653–659 (2006)
24. Kunisch, K., Volkwein, S.: Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition. J. Optim. Theory Appl. **102**(2), 345–371 (1999)
25. Hinze, M., Volkwein, S.: Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: error estimates and suboptimal control. In: Dimension Reduction of Large-Scale Systems. Lecture Notes in Computational and Applied Mathematics, pp. 261–306. Springer (2005)
26. Willcox, K., Peraire, J.: Balanced model reduction via the proper orthogonal decomposition. AIAA J. **40**(11), 2323–2330 (2002)
27. Kunisch, K., Volkwein, S.: Proper orthogonal decomposition for optimality systems. ESAIM Math. Modell. Numer. Anal. **42**(01), 1–23 (2008)
28. Astrid, P., Weiland, S., Willcox, K., Backx, T.: Missing point estimation in models described by proper orthogonal decomposition. IEEE Trans. Autom. Control **53**(10), 2237–2251 (2008)
29. Lall, S., Marsden, J.E., Glavaski, S.: A subspace approach to balanced truncation for model reduction of nonlinear control systems. Int. J. Robust Nonlinear Control **12**, 519–535 (2002)
30. Fujimoto, K., Tsubakino, D.: Computation of nonlinear balanced realization and model reduction based on Taylor series expansion. Syst. Control Lett. **57**(4), 283–289 (2008)
31. Blondel, V.D., Megretski, A.: Unsolved Problems in Mathematical Systems and Control Theory. Princeton University Press (2004)
32. Mäkilä, P.M., Partington, J.R.: Laguerre and Kautz shift approximations of delay systems. Int. J. Control **72**, 932–946 (1999)
33. Mäkilä, P.M., Partington, J.R.: Shift operator induced approximations of delay systems. SIAM J. Control Optim. **37**(6), 1897–1912 (1999)
34. Zhang, J., Knospe, C.R., Tsiotras, P.: Stability of linear time-delay systems: a delay-dependent criterion with a tight conservatism bound. In: Proceedings of the 2000 American Control Conference, Chicago, IL, pp. 1458–1462, June 2000
35. Al-Amer, S.H., Al-Sunni, F.M.: Approximation of time-delay systems. In: Proceedings of the 2000 American Control Conference, Chicago, IL, pp. 2491–2495, June 2000
36. Banks, H.T., Kappel, F.: Spline approximations for functional differential equations. J. Differ. Equ. **34**, 496–522 (1979)
37. Gu, G., Khargonekar, P.P., Lee, E.B.: Approximation of infinite-dimensional systems. IEEE Trans. Autom. Control **34**(6) (1992)
38. Glover, K., Lam, J., Partington, J.R.: Rational approximation of a class of infinite dimensional system i: singular value of hankel operator. Math. Control Circ. Syst. **3**, 325–344 (1990)

39. Glader, C., Hognas, G., Mäkilä, P.M., Toivonen, H.T.: Approximation of delay systems: a case study. Int. J. Control **53**(2), 369–390 (1991)
40. Ohta, Y., Kojima, A.: Formulas for Hankel singular values and vectors for a class of input delay systems. Automatica **35**, 201–215 (1999)
41. Yoon, M.G., Lee, B.H.: A new approximation method for time-delay systems. IEEE Trans. Autom. Control **42**(7), 1008–1012 (1997)
42. Michiels, W., Jarlebring, E., Meerbergen, K.: Krylov-based model order reduction of time-delay systems. SIAM J. Matrix Anal. Appl. **32**(4), 1399–1421 (2011)
43. Jarlebring, E., Damm, T., Michiels, W.: Model reduction of time-delay systems using position balancing and delay Lyapunov equations. Math. Control Signals Syst. **25**(2), 147–166 (2013)
44. Wang, Q., Wang, Y., Lam, E.Y., Wong, N.: Model order reduction for neutral systems by moment matching. Circuits Syst. Signal Process. **32**(3), 1039–1063 (2013)
45. Ionescu, T.C., Iftime, O.V.: Moment matching with prescribed poles and zeros for infinite-dimensional systems. In: American Control Conference, Montreal, Canada, pp. 1412–1417, June 2012
46. Iftime, O.V.: Block circulant and block Toeplitz approximants of a class of spatially distributed systems-An LQR perspective. Automatica **48**(12), 3098–3105 (2012)
47. Astolfi, A.: Model reduction by moment matching, steady-state response and projections. In: Proceedings of the 49th IEEE Conference on Decision and Control (2010)
48. Astolfi, A.: Model reduction by moment matching for linear and nonlinear systems. IEEE Trans. Autom. Control **55**(10), 2321–2336 (2010)
49. Scarciotti, G., Astolfi, A.: Model reduction of neutral linear and nonlinear time-invariant time-delay systems with discrete and distributed delays. IEEE Trans. Autom. Control **61**(6), 1438–1451 (2016)
50. Scarciotti, G., Astolfi, A.: Model reduction for linear systems and linear time-delay systems from input/output data. In: 2015 European Control Conference, Linz, pp. 334–339, July 2015
51. Scarciotti, G., Astolfi, A.: 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, Osaka, Japan, 15–18 Dec 2015
52. Richard, J.P.: Time-delay systems: an overview of some recent advances and open problems. Automatica **39**(10), 1667–1694 (2003)
53. Isidori, A.: Nonlinear Control Systems. Communications and Control Engineering, 3rd edn. Springer (1995)
54. Doyle, J.C., Francis, B.A., Tannenbaum, A.R.: Feedback Control Theory. Macmillan, New York (1992)
55. Antoulas, A.C.: A new result on passivity preserving model reduction. Syst. Control Lett. **54**(4), 361–374 (2005)
56. Sorensen, D.C.: Passivity preserving model reduction via interpolation of spectral zeros. Syst. Control Lett. **54**(4), 347–360 (2005)
57. Hespel, C., Jacob, G.: Approximation of nonlinear dynamic systems by rational series. Theor. Comput. Sci. **79**(1), 151–162 (1991)
58. Hespel, C.: Truncated bilinear approximants: Carleman, finite Volterra, Padé-type, geometric and structural automata. In: Jacob, G., Lamnabhi-Lagarrigue, F. (eds.) Algebraic Computing in Control. Lecture Notes in Control and Information Sciences, vol. 165, pp. 264–278. Springer (1991)
59. Gallivan, K., Vandendorpe, A., Van Dooren, P.: Sylvester equations and projection-based model reduction. J. Comput. Appl. Math. **162**(1), 213–229 (2004)
60. Dib, W., Astolfi, A., Ortega, R.: 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, pp. 6555–6560, Dec 2009
61. Ionescu, T.C., Astolfi, A., Colaneri, P.: Families of moment matching based, low order approximations for linear systems. Syst. Control Lett. **64**, 47–56 (2014)
62. Scarciotti, G., Astolfi, A.: Characterization of the moments of a linear system driven by explicit signal generators. In: Proceedings of the 2015 American Control Conference, Chicago, IL, pp. 589–594, July 2015

63. Scarciotti, G., Astolfi, A.: Model reduction by matching the steady-state response of explicit signal generators. IEEE Trans. Autom. Control **61**(7), 1995–2000 (2016)
64. Scarciotti, G., Astolfi, A.: Data-driven model reduction by moment matching for linear and nonlinear systems. Automatica **79**, 340–351 (2017)
65. Scarciotti, G.: Low computational complexity model reduction of power systems with preservation of physical characteristics. IEEE Trans. Power Syst. **32**(1), 743–752 (2017)
66. Scarciotti, G., Astolfi, A.: Moment based discontinuous phasor transform and its application to the steady-state analysis of inverters and wireless power transfer systems. IEEE Trans. Power Electron. **31**(12), 8448–8460 (2016)
67. Isidori, A., Byrnes, C.I.: Steady-state behaviors in nonlinear systems with an application to robust disturbance rejection. Annu. Rev. Control **32**(1), 1–16 (2008)
68. Scarciotti, G., Astolfi, A.: Model reduction for hybrid systems with state-dependent jumps. In: IFAC Symposium Nonlinear Control Systems, Monterey, CA, USA, pp. 862–867 (2016)
69. Scarciotti, G., Jiang, Z.P., Astolfi, A.: Constrained optimal reduced-order models from input/output data. In: Proceedings of the 55th IEEE Conference on Decision and Control, Las Vegas, NV, USA, pp. 7453–7458, 12–14 Dec 2016
70. Padoan, A., Scarciotti, G., Astolfi, A.: A geometric characterisation of persistently exciting signals generated by autonomous systems. In: IFAC Symposium Nonlinear Control Systems, Monterey, CA, USA, pp. 838–843 (2016)
71. Isidori, A., Byrnes, C.I.: Steady state response, separation principle and the output regulation of nonlinear systems. In: Proceedings of the 28th IEEE Conference on Decision and Control, Tampa, FL, USA, pp. 2247–2251 (1989)
72. Isidori, A., Byrnes, C.I.: Output regulation of nonlinear systems. IEEE Trans. Autom. Control **35**(2), 131–140 (1990)
73. Hale, J.K.: Theory of functional differential equations. Applied Mathematical Sciences Series. Springer Verlag Gmbh (1977)
74. Hale, J.K.: Behavior near constant solutions of functional differential equations. J. Differ. Equ. **15**, 278–294 (1974)
75. Byrnes, C.I., Spong, M.W., Tarn, T.J.: A several complex variables approach to feedback stabilization of linear neutral delay-differential systems. Math. Syst. Theory **17**(1), 97–133 (1984)
76. Bian, T., Jiang, Y., Jiang, Z.P.: Adaptive dynamic programming and optimal control of nonlinear nonaffine systems. Automatica **50**(10), 2624–2632 (2014)