Model reduction by moment matching for linear and nonlinear systems

Alessandro Astolfi, *Fellow, IEEE*

Abstract— The model reduction problem for (single-input, single-output) linear and nonlinear systems is addressed using the notion of moment. A re-visitation of the linear theory allows to obtain novel results for linear systems and to develop a nonlinear enhancement of the notion of moment. This, in turn, 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 model. Connections between the proposed results, projection methods, the covariance extension problem and interpolation theory are presented. Finally, the theory is illustrated by means of simple worked out examples and case studies.

I. Introduction

The model reduction problem for linear and nonlinear systems has been widely studied over the past decades. This problem has great importance in applications, because *reduced order models* are often used in analysis and design. This is the case, for example, in the study of mechanical systems, which is often based on models derived from a rigid body perspective that neglects the presence of flexible modes and elasticity; and in the study of large scale systems, such as integrated circuits or weather forecast models, which relies upon the construction of simplified models that capture the main features of the system. From a theoretical point of view, the model reduction problem generates important theoretical questions and requires advanced tools from linear algebra, functional analysis and numerical analysis. The model reduction problem can be simply, and informally, posed as follows. Given a system, described by means of linear or nonlinear differential equations together with an output map, compute a *simpler* system which *approximates* (in a sense to be specified) its behavior. To render precise this problem formulation it is necessary to define two concepts. Firstly, the meaning of the approximation. For linear systems one could introduce an approximation error given in terms of the frequency response of a suitably defined *error system*, or in terms of the response of the system for classes of input signals. For example, the methods, known as moment matching methods, which zero the transfer function of the error system for specific frequencies, belong to this class [1]. This approach does not have a direct nonlinear counterpart, despite the recent developments in this direction [2] (see also the early contributions [3], [4], [5], [6], [7], [8]). Alternatively, approximation errors expressed in terms of the H_2 or H_∞ norm of the error system have been considered both in the linear case [9], [10], [11], [12] and in the nonlinear

A. Astolfi is with the Department of Electrical and Electronic Engineering, Imperial College London, and with the Dipartimento di Informatica, Sistemi e Produzione, Università di Roma "Tor Vergata". (E-mail: a.astolfi@imperial.ac.uk).

case [13]. Finally, approximation errors based on the Hankel operator of the system have been widely considered [14], [15], [16]. This approach leads to the so-called balancing realization problem [17], [18], which has been also studied in the nonlinear framework [19], [20], [21], [22], [23], [24], [25], [26], [27]. Secondly, the concept of simplicity. For linear systems this is often understood in terms of the dimension of the system, *i.e.* an approximating system is simpler than the model to approximate if its state-space realization has fewer states. For nonlinear systems this dimensional argument may be inappropriate, as one has to take into consideration also the complexity of the functions involved in the state-space representation. In addition, the notion of dimension has, for nonlinear systems, a local nature: the dimension of a *good* reduced order model may depend upon the initial condition. Of course, there are other important issues that have to be clarified, and investigated, in establishing a model reduction theory. In particular, one may require that properties of the model (such as stability or passivity) are retained by the approximation [28], and one has to consider the computational cost associated with the construction of the approximating system. These issues have been widely investigated in the linear framework, see for example the excellent monograph [1], but are largely open for nonlinear systems. In particular, the computation of reduced order models for nonlinear systems often requires the (approximate) solution of partial differential equations, and thus it can be performed only in special circumstances. In addition, some results for time-varying systems have been proposed in [29], whereas a dedicated method for systems with the so-called ZIP property [30] has been developed in [31].

The model reduction problem for nonlinear systems have been investigated from diverse perspectives. Energy-based methods have been discussed, for example, in [21], [23], [32], [33]; model reduction of systems in special forms, such as differential-algebraic systems, bilinear systems or mechanical/Hamiltonian systems, have been discussed in [34], [35], [36], [37], [38], and reduction around a limit cycle or a manifold has been discussed in [39], [40]. Some computational issues have been addressed in [41], [42], [40], [27]. Finally, proper orthogonal decomposition has been proposed as a model reduction method applicable for linear and nonlinear systems, see, for example, [43], [44], [42], [45], [46], where theoretical issues, error bounds and applications have been discussed. We refer the reader to the references therein for further detail.

Goal of this work is to develop a theory of model reduction, based on the notion of moment, for nonlinear systems. In this process, we revisit the linear theory, providing new perspectives and results. We note that this work relies upon the

theory of the steady-state response of nonlinear systems, center manifold theory and the tools arising in the theory of output regulation for nonlinear systems [5], [6], [47], [48], [49], [50]. Preliminary versions of this work have been published in [51], [52], [53].

We complete the introduction noting that, for linear systems, the problem of model reduction by moment matching provides a different formulation of the classical (rational) interpolation problem, which has been solved by Nevanlinna-Pick theory (illustrated in several textbooks, see e.g. [54]). The interpolation problem has been also extensively studied by Byrnes, Georgiou and Lindquist (and co-workers), see *e.g.* [55], [56], [57], [58], [59], [60], that have developed a complete theory for analytic interpolation with degree constraint. Note that, the interpolation problem does not have direct counterpart for nonlinear systems. However, as a direct byproduct of the results in this paper, it is possible to pose (and solve) interpolation problems for nonlinear systems. Note that some early results on interpolation for nonlinear systems have been presented in [61], [62]. Therein, the existence of minimal rank bilinear systems approximating a general (analytic) system up to order k, for any input, has been demonstrated.

The paper is logically composed of two parts. In the first part (Section II) the problem of model reduction by moment matching for linear systems is re-visited from a point of view that allows for the nonlinear theory to be developed in the second part (Section III). Obviously, some of the *linear* results can be obtained specializing the corresponding *nonlinear* results to linear systems. Note however that in the linear case it is often possible to provide stronger statements, and in the nonlinear case there are problems or issues that are worth investigating, but that do not have special interest (or meaning) in the linear one.

Notation. Throughout the paper we use standard notation. $I\!R$, $I\!R^n$ and $I\!R^{n \times m}$ denote the set of real numbers, of n-dimensional vectors with real components, and of $n \times m$-dimensional matrices with real entries, respectively. $I\!R^+$ ($I\!R^-$) denotes the set of non-negative (non-positive) real numbers. $\mathbb{C}$ denotes the set of complex numbers, $\mathbb{C}^0$ denotes the set of complex numbers with zero real part, $\mathbb{C}^-$ denotes the set of complex numbers with negative real part. $\sigma(A)$ denotes the spectrum of the matrix $A \in I\!R^{n \times n}$. Finally, $\emptyset$ denotes the empty set.

II. MODEL REDUCTION BY MOMENT MATCHING FOR LINEAR SYSTEMS – REVISITED

A. *The notion of moment*

Consider a linear, single-input, single-output[1], continuous-time system described by equations of the form

$$\dot{x} = Ax + Bu, \qquad y = Cx, \qquad (1)$$

with $x(t) \in I\!R^n$, $u(t) \in I\!R$, $y(t) \in I\!R$, $A \in I\!R^{n \times n}$, $B \in I\!R^n$ and $C \in I\!R^{1 \times n}$ constant matrices, and the associated transfer function

$$W(s) = C(sI - A)^{-1}B. \qquad (2)$$

[1]Similar considerations can be performed for multi-input, multi-output systems.

Definition 1: [1] The 0-moment of system (1) at $s^\star \in \mathbb{C}$ is the complex number $\eta_0(s^\star) = C(s^\star I - A)^{-1}B$. The k-moment of system (1) at $s^\star \in \mathbb{C}$ is the complex number

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

Note that moments are associated with the transfer function of the system (1). Therefore, in what follows we make the following standing assumption.

(SA) The system (1) is controllable and observable.

Moments can be also characterized, for almost all $s^\star$, in a time-domain setting, as shown in the following statements (see also [1, Chapter 6]).

Lemma 1: Consider system (1) and $s^\star \in \mathbb{C}$. Suppose $s^\star \notin \sigma(A)$. Then $\eta_0(s^\star) = C\Pi$, where Π is the unique solution of the Sylvester equation

$$A\Pi + B = \Pi s^\star. \qquad (3)$$

Proof: By direct computation, equation (3) yields $\Pi = (s^\star I - A)^{-1}B$, hence $C\Pi = \eta_0(s^\star)$. ∎

Lemma 2: Consider system (1) and $s^\star \in \mathbb{C}$. Suppose $s^\star \notin \sigma(A)$. Then

$$\begin{bmatrix} \eta_0(s^\star) \\ \vdots \\ \eta_k(s^\star) \end{bmatrix} = (C\Pi\Psi_k)',$$

where

$$\Psi_k = \text{diag}(1, -1, 1, \cdots, (-1)^k) \in I\!R^{(k+1)\times(k+1)},$$

and Π is the unique solution of the Sylvester equation

$$A\Pi + BL_k = \Pi\Sigma_k, \qquad (4)$$

with $L_k = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \in I\!R^{k+1}$, and

$$\Sigma_k = \begin{bmatrix} s^\star & 1 & 0 & \cdots & 0 \\ 0 & s^\star & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & s^\star & 1 \\ 0 & \cdots & \cdots & 0 & s^\star \end{bmatrix} \in I\!R^{(k+1)\times(k+1)}.$$

Proof: Let $\Pi = \begin{bmatrix} \Pi_0 & \Pi_1 & \cdots & \Pi_k \end{bmatrix}$ and note that equation (4) can be rewritten as

$$\begin{aligned} A\Pi_0 + B &= s^\star\Pi_0, \\ A\Pi_1 &= s^\star\Pi_1 + \Pi_0, \\ &\vdots \\ A\Pi_k &= s^\star\Pi_k + \Pi_{k-1}. \end{aligned}$$

As a result,

$$\begin{aligned} \Pi_0 &= (s^\star I - A)^{-1}B, \\ \Pi_1 &= -(s^\star I - A)^{-2}B, \\ &\vdots \\ \Pi_k &= (-1)^k(s^\star I - A)^{-(k+1)}B, \end{aligned}$$

hence the claim. ∎

Remark 1: The time domain characterization of moments in Lemmas 1 and 2 is not valid for $s^\star = \infty$ (see Example 9 for a characterization of moments at $s^\star = \infty$).

Remark 2: The pair (L_k, Σ_k) is observable for any $s^\star$.

The main disadvantage of the above results is in the fact that one has to deal with complex matrices and Σ_k and L_k have a special structure. To remove this shortcoming note that moments are coordinate invariant and, by a property of real rational functions, $\eta_k(\bar{s}^\star) = \overline{\eta_k(s^\star)}$. As a result, the following statements hold.

Lemma 3: Consider system (1) and $s^\star \in \mathbb{R}$. Suppose $s^\star \notin \sigma(A)$. Then there is a one-to-one relation between the moments $\eta_0(s^\star)$, ..., $\eta_k(s^\star)$ and the matrix $C\Pi$, where Π is the unique solution of the Sylvester equation

$$A\Pi + BL = \Pi S, \tag{5}$$

with S any non-derogatory[2] real matrix such that

$$\det(sI - S) = (s - s^\star)^{k+1}, \tag{6}$$

and L such that the pair (L, S) is observable.

Proof: By observability of the pair (L, S) there is a unique invertible matrix T such that $S = T^{-1}\Sigma_k T$ and $L = L_k T$. As a result, equation (5) becomes $A\Pi + BL_k T = \Pi T^{-1}\Sigma_k T$, and this can be rewritten as $A\tilde{\Pi} + BL_k = \tilde{\Pi}\Sigma_k$, with $\tilde{\Pi} = \Pi T^{-1}$. By Lemma 2, and invariance of the moments with respect to the coordinates in the state space, the moments $\eta_0(s^\star)$, ..., $\eta_k(s^\star)$ can be univocally expressed in terms of $\tilde{\Pi}$, hence the claim. ∎

Lemma 4: Consider system (1) and $s^\star \in \mathbb{C} \setminus \mathbb{R}$. Let $s^\star = \alpha^\star + i\omega^\star$. Suppose $s^\star \notin \sigma(A)$. Then there is a one-to-one relation between the moments $\eta_0(s^\star)$, $\eta_0(\bar{s}^\star)$, ..., $\eta_k(s^\star)$, $\eta_k(\bar{s}^\star)$ and the matrix $C\Pi$, where Π is the unique solution of the Sylvester equation

$$A\Pi + BL = \Pi S, \tag{7}$$

with S any non-derogatory real matrix such that

$$\det(sI - S) = ((s - s^\star)(s - \bar{s}^\star))^{k+1}, \tag{8}$$

and L such that the pair (L, S) is observable.

Proof: The proof is similar to the one of Lemma 3 hence omitted. ∎

We complete this section with a property which is instrumental to derive a nonlinear enhancement of the notion of moment.

Theorem 1: Consider system (1), $s^\star \in \mathbb{C}$ and $k \geq 0$. Assume $\sigma(A) \subset \mathbb{C}^-$ and $s^\star \in \mathbb{C}^0$. Let

$$\dot{\omega} = S\omega, \tag{9}$$

with $\omega(t) \in \mathbb{R}^\kappa$, where

$$\kappa = \begin{cases} k+1 & \text{if } s^\star \in \mathbb{R}, \\ 2(k+1) & \text{if } s^\star \in \mathbb{C} \setminus \mathbb{R}, \end{cases}$$

and $S \in \mathbb{R}^{\kappa \times \kappa}$ any non-derogatory real matrix with characteristic polynomial as in (6), if $s^\star \in \mathbb{R}$, or as in (8), if $s^\star \in \mathbb{C} \setminus \mathbb{R}$. Let $\omega(0) \neq 0$.

Consider the interconnection of systems (1) and (9) with $u = L\omega$, and L such that the pair (L, S) is observable.

---

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

Then the moments $\eta_0(s^\star)$, ..., $\eta_k(s^\star)$ are in one-to-one relation with the (well-defined) steady-state response of the output of such interconnected system.

Proof: We provide a proof which exploits arguments with a nonlinear counterpart (an elementary, alternative, proof can be obtained using Laplace transform arguments). The considered interconnected system is described by

$$\begin{aligned} \dot{\omega} &= S\omega, \\ \dot{x} &= Ax + BL\omega, \\ y &= Cx. \end{aligned}$$

By the center manifold theorem [63], [64], which is applicable because of the assumptions on $\sigma(A)$ and $\sigma(S)$, this system has a globally well-defined invariant manifold (which is a hyperplane) given by $\mathcal{M} = \{(x, \omega) \in \mathbb{R}^{n+\kappa} \mid x = \Pi\omega\}$, with Π the unique solution of the Sylvester equation (5). Note that $\overbrace{x - \Pi\dot{\omega}} = A(x - \Pi\omega)$, hence $\mathcal{M}$ is attractive. As a result $y(t) = C\Pi\omega(t) + Ce^{At}(x(0) - \Pi\omega(0))$, where the first term on the right-hand side describes the steady-state response of the system, and the second term on the right-hand side the transient response, which proves the claim. ∎

Remark 3: The results derived so-far are direct consequences of the definition of moment. However, to the best of the author's knowledge, they have not been presented in this form, see however a similar discussion in [1, Section 6.1], the paper [65], in which Sylvester equations are used to compute the Krylov subspaces arising in the construction of the projectors which define the reduced order model, and the early results [66] (and references therein), in which the relation between least-square interpolation and moment matching (at $s = 0$ and $s = \infty$) is discussed.

Remark 4: The main advantage of the characterization of moments in terms of steady-state responses is that it allows one to define the notion of moment for systems which do not admit a representation in terms of transfer function, for example linear time-varying systems, which we do not discuss for reason of space, and nonlinear systems, which are discussed extensively in Section III.

The main disadvantage is that this characterization requires existence of the steady-state response, hence the system to be *reduced* has to possess some *strong* stability property. Note however that, for unstable systems it is still possible to define moments by means of the Sylvester equation (5). This type of definition admits a nonlinear counterpart, which is however not discussed in this paper (see also Remark 10).

B. Moment matching

We are now in a position to define, precisely, the notion of reduced order model. The system

$$\dot{\xi} = F\xi + Gu, \qquad \psi = H\xi, \tag{10}$$

with $\xi(t) \in \mathbb{R}^\nu$ and $\psi(t) \in \mathbb{R}$, is a model of system (1) at $S \in \mathbb{R}^{\tilde{\nu} \times \tilde{\nu}}$, with S such that $\sigma(S) \cap \sigma(A) = \emptyset$, if

$$\sigma(S) \cap \sigma(F) = \emptyset \tag{11}$$
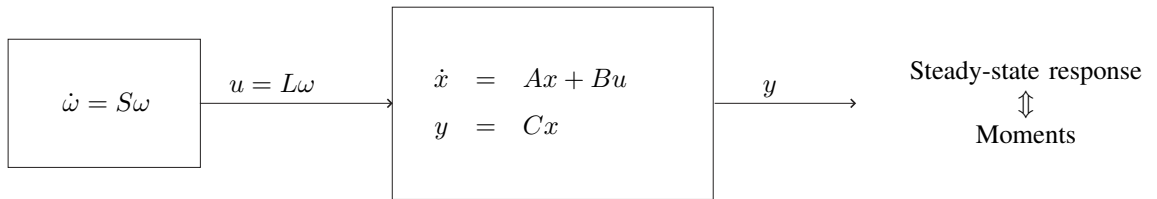
and

$$C\Pi = HP, \tag{12}$$

Fig. 1. A diagrammatic illustration of Theorem 1.

where Π is the unique solution of equation (5), with L such that the pair (L, S) is observable, and P is the unique solution of the equation

$$FP + GL = PS. \qquad (13)$$

Furthermore, system (10) is a reduced order model of system (1) if $\nu < n$. Note that, consistently with the discussion in [1], the system (10) solves the model reduction problem with moment matching at $\sigma(S)$ for system (1).

C. Model reduction by moment matching

The proposed results can, in principle, be used to solve the model reduction problem by moment matching for system (1) in two steps. In the former one has to solve the Sylvester equation (5) in the unknown Π. In the latter one has to construct matrices F, G, H and P (possibly with specific properties) such that equations (12) and (13) hold. This approach is unsatisfactory because it requires the computation of the moments, namely of the matrix $C\Pi$ and hence of Π, whereas most of the existing algorithms [1] are able to achieve moment matching without the need to compute moments. To avoid the computation of the moments, *i.e.* of the matrix Π, one could proceed as follows. Consider system (1), a matrix S, and construct a reduced order model achieving moment matching at S with any efficient algorithm that does not require the computation of the moments, see [1], [67], [68], [69], [70]. This yields a reduced order model for system (1) described by equations of the form

$$\dot{x}_M = A_M x_M + B_M u, \qquad y_M = C_M x_M, \qquad (14)$$

where $x_M(t) \in \mathbb{R}^\nu$ and $y_M(t) \in \mathbb{R}$. To find a reduced order model with desired properties it is thus sufficient to apply the model reduction procedures discussed hereafter to system (14). This approach is similar in spirit to that described in [60].

We now focus on the second step of the construction, with the goal to achieve moment matching with additional constraints on the reduced order model. To simplify this step we make the following assumptions.

(A1) The matrices F and S have the same dimensions.

(A2) S and A do not have common eigenvalues and S and F do not have common eigenvalues.

Note now that selecting $P = I$ yields a family of reduced order models for system (1) at S, hence achieving moment matching, described by equations of the form (10) with

$$F = S - \Delta L, \qquad G = \Delta, \qquad H = C\Pi,$$

namely

$$\dot{\xi} = (S - \Delta L)\xi + \Delta u, \qquad \psi = C\Pi\xi, \qquad (15)$$

with Δ any matrix such that $\sigma(S) \cap \sigma(S - \Delta L) = \emptyset$.

Remark 5: The proposed approach differs from the interpolation/projection approach developed, for example, in [1]. Therein, moment matching is achieved by means of the selection of two oblique projectors, which are computed on the basis of so-called interpolation points. Half of these interpolation points are determined by the location of the moments that have to be matched, whereas the remaining interpolation points are free parameters. These free parameters can be selected to enforce specific properties of the reduced order model, such as stability, the relative degree, or passivity. However, the relation between free interpolation points and properties of the reduced order model is nontrivial and does not possess a direct (system theoretic) interpretation (see [71], [12], [28], [72]). Finally, the projectors can also be selected to enforce a particular structure for the reduced order model, as discussed in detail in [73].

On the contrary, in this paper, the family of reduced order models achieving moment matching is parameterized directly by the matrix Δ, which has to satisfy a generic constraint. As a result, the relations between the matrix Δ and properties of the reduced order model are straightforward and easy to characterize, as discussed in the following sections.

Before ending this section we discuss two important problems associated with the parameterization given in equation (15). The former is the *completeness* of the parameterization, *i.e.* if the family (15) contains all systems, of dimension ν achieving moment matching; the latter is the connection with the so-called Georgiou-Kimura parameterization [55], [74], [75].

Proposition 1: Consider the family of systems (15). Consider a model of system (1) at S of dimension ν, and let $W(s)$ be its transfer function. Then there exists a unique Δ such that $W(s) = C\Pi(sI - (S - \Delta L))^{-1}\Delta$, *i.e.* the family of systems (15) contains all models of system (1) at S of dimension ν.

Proof: Let $W(s) = \frac{N(s)}{D(s)}$ and select Δ such that $D(s) = \det(sI - (S - \Delta L))$. Under the stated assumptions, there is a unique Δ satisfying this condition. To complete the proof we need to show that $N(s) = C\Pi \operatorname{adj}(sI - (S - \Delta L))\Delta$. The condition of moment matching implies ν independent equality conditions on the polynomials $N(s)$ and $C\Pi \operatorname{adj}(sI - (S - \Delta L))\Delta$, which are both of degree $\nu - 1$, hence these polynomials are identical, which proves the claim. ∎

We now consider the second issue, *i.e.* the connection with the Georgiou-Kimura parameterization [55], [74], [75]. The Georgiou-Kimura parameterization arises in the solution of the covariance extension problem. The covariance extension problem can be recast as the problem of matching k moments at 0. The solution of this problem, as presented in [75, Theorem 1], is given by the rational function

$$W_{GK}(s) = \frac{\psi_n(s) + \gamma_1 \psi_{n-1}(s) + \cdots + \gamma_n \psi_0(s)}{\phi_n(s) + \gamma_1 \phi_{n-1}(s) + \cdots + \gamma_n \phi_0(s)},$$

where $\gamma_i \in I\!\!R$, for $i \in [1, n]$, and $\phi(\cdot)$ and $\psi(\cdot)$ are the Szegö orthogonal polynomial of the first and second kind, respectively. It is worth comparing the above parameterization with the parameterization proposed in this work, which is given in terms of the vector Δ. To this end, note that, in the Georgiou-Kimura parameterization, the *parameters* γ_i enter linearly both in the numerator and in the denominator, whereas in the proposed representation the entries of the parameter Δ enter nonlinearly in both cases. While this may be a disadvantage, it is worth stressing that, as will become clear in what follows, it is possible to relate directly the parameter Δ with some systems theoretic properties of the reduced order model (which solves the covariance extension problem), whereas it is very hard to derive similar interpretation for the parameters γ_i. Note finally that, by Proposition 1 and [75, Theorem 1], the set of all rational transfer functions given by the Georgiou-Kimura parameterization coincides with the set characterized in this work.

Example 1: Consider a reduced order model described by the equations (15), with

$$S = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \tag{16}$$

and

$$L = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \qquad C\Pi = \begin{bmatrix} \eta_0 & \eta_1 & \eta_2 \end{bmatrix}. \tag{17}$$

This implies that the reduced order model matches the 0-, 1- and 2- moments at zero. Note that, independently from the selection of Δ, the transfer function $W(s) = C\Pi(sI - (S - \Delta L))^{-1}\Delta$ of the reduced order model is such that $W(s) = \eta_0 + \eta_1 s + \eta_2 s^2 + \ldots$.

1) The interpolation problem: As discussed in the introduction, the main difference between the model reduction problem and the classical Nevanlinna-Pick interpolation problem, is that the number of interpolation points is equal to the order of the approximating system plus one. As a result, the interpolating problem can be regarded as a model reduction problem by moment matching in which a model of order n should match $n+1$ moments. This problem can be characterized as follows. (For simplicity, we assume that one of the interpolation point is real.)

Corollary 1: Consider system (1) and an observable pair $(\tilde{L}, \tilde{S})$ with $\tilde{S} \in I\!\!R^{\nu+1 \times \nu+1}$ and $\tilde{L} \in I\!\!R^{1 \times \nu+1}$. Suppose $\sigma(\tilde{S}) \cap \sigma(A) = \emptyset$,

$$\tilde{S} = \begin{bmatrix} S & S_1 \\ 0 & s_2 \end{bmatrix}$$

and $\tilde{L} = \begin{bmatrix} L & 0 \end{bmatrix}$, with $S \in I\!\!R^{\nu \times \nu}$, $S_1 \in I\!\!R^{\nu \times 1}$, $s_2 \in I\!\!R$ and $L \in I\!\!R^{1 \times \nu}$. Then there exists a reduced order model of system (1) at $\tilde{S}$ described by equations of the form (15) with $\xi \in I\!\!R^\nu$ if and only if there is a Δ such that

$$C(s_2 I - A)^{-1}(\Pi\Delta - B)L(s_2 I - (S - \Delta L))^{-1}S_1 = 0, \tag{18}$$

with Π the unique solution of the equation $A\Pi + BL = \Pi S$.

Proof: The proof is a direct application of the definition of moment, hence it is omitted. ∎

Remark 6: Let $W(s)$ be the transfer function of the system

$$\dot{x} = Ax + (\Pi\Delta - B)u, \qquad y = Cx,$$

and $M(s)$ be the transfer function of the system

$$\dot{\xi} = (S - \Delta L)\xi + S_1 v, \qquad \psi = L\xi.$$

Condition (18) can be written as $W(s_2)M(s_2) = 0$, which expresses that Δ is such that at least one of the two transfer functions $W(s)$ and $M(s)$ have a zero at s_2. As a result, condition (18) is equivalent to

$$\det\begin{bmatrix} s_2 I - A & \Pi\Delta - B \\ C & 0 \end{bmatrix} \det\begin{bmatrix} s_2 I - (S - \Delta L) & S_1 \\ L & 0 \end{bmatrix} = 0.$$

2) Matching with prescribed eigenvalues: Consider system (15) and the problem of determining Δ such that the reduced order model has prescribed eigenvalues, *i.e.* $\sigma(F) = \{\lambda_1, \cdots, \lambda_\nu\}$, for some given λ_i's such that $\sigma(F) \cap \sigma(S) = \emptyset$. This problem has been solved, using Krylov-type projection tools, in [71], whereas model reduction by moment matching and with a stability constraint has been recently discussed in [76]. This goal is achieved selecting Δ such that

$$\sigma(S - \Delta L) = \sigma(F). \tag{19}$$

Note that, by observability of the pair (L, S), there is a unique matrix Δ such that condition (19) holds.

Example 2: Consider a reduced order model described by the equations (15), with S, L and $C\Pi$ as in equations (16) and (17). The selection $\Delta = \begin{bmatrix} 3\lambda & 3\lambda^2 & \lambda^3 \end{bmatrix}'$ yields a reduced order model with three eigenvalues at $-\lambda$, and with transfer function

$$W(s) = \frac{\lambda(3\eta_0 + 3\lambda\eta_1 + \lambda^2\eta_2)s^2 + \lambda^2(\lambda\eta_1 + 3\eta_0)s + \lambda^3\eta_0}{(s + \lambda)^3}$$

$$= \eta_0 + \eta_1 s + \eta_2 s^2 + \ldots$$

3) Matching with prescribed relative degree: Consider system (15) and the problem of determining Δ such that the system has a given relative degree $r \in [1, \nu]$. By definition of relative degree, system (15) has relative degree r if

$$C\Pi\Delta = 0,$$

$$\vdots$$

$$C\Pi(S - \Delta L)^{r-2}\Delta = 0,$$
$$C\Pi(S - \Delta L)^{r-1}\Delta \neq 0,$$

or, equivalently, if

$$\begin{bmatrix} C\Pi \\ \vdots \\ C\Pi S^{r-2} \\ C\Pi S^{r-1} \end{bmatrix} \Delta = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \gamma \end{bmatrix}, \tag{20}$$

for some nonzero γ. Therefore, to assign the relative degree of the reduced order model it is enough to solve the equation (20) in the unknown Δ. Solvability of this equation, for all $r \in [1, \nu]$, is discussed in the next statement.

Theorem 2: The following statements are equivalent.

(RD1) Equation (20) has a solution Δ for all $r \in [1, \nu]$.

(RD2) The system

$$\dot{\xi} = S\xi, \qquad \psi = C\Pi\xi, \qquad (21)$$

is observable.

(RD3) The system

$$\dot{\omega} = S\omega, \quad \dot{x} = Ax + BL\omega, \quad y = Cx, \quad (22)$$

is observable.

Proof: (RD1)$\Rightarrow$(RD2). If equation (20) has a solution Δ for all $r \in [1, \nu]$ then $C\Pi \neq 0$, and setting $r = 2$ yields that $C\Pi S$ is linearly independent from $C\Pi$, *i.e.*

$$\text{rank} \left[\begin{array}{c} C\Pi \\ C\Pi S \end{array} \right] = 2.$$

Using the same arguments we infer that, for each $r \in [1, \nu]$,

$$\text{rank} \left[\begin{array}{c} C\Pi \\ C\Pi S \\ \vdots \\ C\Pi S^{r-1} \end{array} \right] = r, \qquad (23)$$

hence for $r = \nu$ we conclude observability of system (21).

(RD2)$\Rightarrow$(RD1). Observability of system (21) implies equation (23), for all $r \in [1, \nu]$, and hence solvability of equation (20) for all $r \in [1, \nu]$.

(RD3)$\Rightarrow$(RD2). We prove this implication by contradiction. Suppose system (21) is not observable. Then there exist a (possibly complex) vector v and a scalar λ such that $Sv = \lambda v$ and $C\Pi v = 0$. Let

$$w = \left[\begin{array}{c} v \\ \Pi v \end{array} \right] = \left[\begin{array}{c} I \\ \Pi \end{array} \right] v$$

and note that

$$\left[\begin{array}{cc} S & 0 \\ BL & A \end{array} \right] w = \left[\begin{array}{c} Sv \\ (BL + A\Pi)v \end{array} \right] = \left[\begin{array}{c} I \\ \Pi \end{array} \right] Sv = \lambda w,$$

and $\left[\begin{array}{cc} 0 & C \end{array} \right] w = 0$, hence system (22) is not observable, which proves the claim.

(RD2)$\Rightarrow$(RD3). We prove this implication by contradiction. Suppose system (22) is not observable. Then there exist a (possibly complex) vector $w = [w_1'\ w_2']'$ and a scalar λ such that

$$\left[\begin{array}{cc} S & 0 \\ BL & A \end{array} \right] w = \lambda w \qquad \left[\begin{array}{cc} 0 & C \end{array} \right] w = 0.$$

This implies $Sw_1 = \lambda w_1$ and

$$\begin{aligned} BLw_1 + Aw_2 &= \Pi Sw_1 - A\Pi w_1 + Aw_2 \\ &= \lambda \Pi w_1 - A\Pi w_1 + Aw_2 \\ &= \lambda w_2, \end{aligned}$$

hence $(A - \lambda I)(w_2 - \Pi w_1) = 0$. As a result, since λ is an eigenvalue of S and S and A do not have common eigenvalues

$w_2 = \Pi w_1$, which implies $C\Pi w_1 = 0$, hence system (21) is not observable, which proves the claim. $\blacksquare$

Remark 7: If system (21), or system (22), are not observable, it may still be possible to assign the relative degree of the model for some $r \in [1, \nu]$. However, if it is possible to assign a relative degree r, then it is possible to assign any relative degree $\tilde{r} \in [1, r]$.

The matrix Δ assigning the relative degree to system (10) is not unique. This degree of freedom may be exploited to partly assign the eigenvalues or the zeros of the reduced order model. To illustrate this issue consider the following example.

Example 3: Consider a reduced order model described by the equations (15), with

$$S = \left[\begin{array}{cc} 0 & \omega^{\star} \\ -\omega^{\star} & 0 \end{array} \right],$$

$L = \left[\begin{array}{cc} 1 & 0 \end{array} \right]$ and $C\Pi = \left[\begin{array}{cc} \eta_{0s} & \eta_{0c} \end{array} \right]$, with $\omega^{\star} > 0$ and $\eta_{0s}^2 + \eta_{0c}^2 \neq 0$. This implies that the reduced order model matches the 0-moment at $\pm i\omega^{\star}$.

All matrices Δ such that the reduced order model has relative degree one are described by

$$\Delta = \gamma \left[\begin{array}{c} \eta_{0s} \\ \eta_{0c} \end{array} \right] + \delta \left[\begin{array}{c} -\eta_{0c} \\ \eta_{0s} \end{array} \right],$$

with $\gamma \neq 0$ and $\delta \in I\!\!R$. The transfer function $W(s)$ of the reduced order model is given by

$$W(s) = \frac{\gamma s + \omega^{\star}\delta}{s^2 + (\gamma\eta_{0s} - \delta\eta_{0c})s + \omega^{\star}(\omega^{\star} + \delta\eta_{0s} + \gamma\eta_{0c})}, \tag{24}$$

and the parameters δ and γ can be used to assign its poles, or the zero and the DC-gain, or to obtain a reduced order model which is asymptotically stable and minimum phase (*i.e.* it is passive).

All matrices Δ such that the reduced order model has relative degree two are described by

$$\Delta = \frac{\gamma}{\omega^{\star}} \left[\begin{array}{c} -\eta_{0c} \\ \eta_{0s} \end{array} \right],$$

with $\gamma \neq 0$. The transfer function $W(s)$ of the reduced order model is given by

$$W(s) = \frac{\gamma}{s^2 - \eta_{0c}\frac{\gamma}{\omega^{\star}}s + (\omega^{\star})^2 + \gamma\eta_{0s}}, \tag{25}$$

and the parameter γ can be selected to ensure asymptotic stability of the reduced order model provided $\eta_{0c} \neq 0$.

It is worth noting that, consistently with the theory, the value of the transfer functions in equation (24) and (25) for $s = \pm i\omega^{\star}$ depends only upon η_{0s} and η_{0c}.

Example 4: Consider a reduced order model described by the equations (15), with S, L and $C\Pi$ as in equations (16) and (17). Let $\eta_0 \neq 0$.

All matrices Δ such that the reduced order model has relative degree three are described by

$$\Delta = \gamma \left[\begin{array}{c} \dfrac{\eta_1^2 - \eta_0\eta_2}{\eta_0^3} \\ -\dfrac{\eta_1}{\eta_0^2} \\ \dfrac{1}{\eta_0} \end{array} \right]$$

with $\gamma \neq 0$. The transfer function of the reduced order model is given by

$$W(s) = \frac{\gamma \eta_0^3}{\eta_0^3 s^3 + (\eta_1^2 - \eta_0 \eta_2)\gamma s^2 - \eta_0 \eta_1 \gamma s + \eta_0^2 \gamma} \quad (26)$$

and the parameter γ can be used to assign, for example, the high-frequency gain.

4) Matching with prescribed zeros: Consider now the problem of selecting Δ such that system (15) has prescribed zeros. This problem admits the following characterization.

Theorem 3: The following statements are equivalent.

(Z1) The zeros of system (15) can be arbitrarily assigned by a proper selection of Δ.

(Z2) The zeros of the system

$$\dot{\xi} = S\xi + \Delta u, \qquad \psi = C\Pi\xi, \quad (27)$$

can be arbitrarily assigned properly selecting Δ.

(Z3) The system (21) is observable.

(Z4) The system (22) is observable.

Proof: The equivalence between (Z3) and (Z4) has been established in Theorem 2.

(Z1)$\Leftrightarrow$(Z2). To prove the claim we simply show that the zeros of the two systems coincide. For, note that

$$\begin{bmatrix} sI - (S - \Delta L) & \Delta \\ C\Pi & 0 \end{bmatrix} = \begin{bmatrix} sI - S & \Delta \\ C\Pi & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ L & 1 \end{bmatrix},$$

which proves the claim.

(Z2)$\Rightarrow$(Z3). The zeros of system (27) are the roots of the polynomial

$$\det \begin{bmatrix} sI - S & \Delta \\ C\Pi & 0 \end{bmatrix}.$$

Observability of system (21) implies that

$$\text{rank} \begin{bmatrix} sI - S \\ C\Pi \end{bmatrix} = n,$$

for all $s \in \mathbb{C}$. Let $z_i(s)$ be the determinant of the minor of order n obtained *eliminating* from the observability pencil the i-th row of the matrix $sI - S$. By observability, the n polynomials $z_1(s), \cdots, z_n(s)$ do not have common roots, *i.e.* are independent polynomials of degree $n-1$. This implies that such polynomials form a basis for the space of polynomials of degree $n-1$. Note now that (setting $\Delta = [\Delta_1 \, \Delta_2 \, \cdots \, \Delta_n]'$)

$$\begin{aligned} Z(s) &= \det \begin{bmatrix} sI - S & \Delta \\ C\Pi & 0 \end{bmatrix} \\ &= (-1)^n \left(\Delta_1 z_1(s) - \cdots + (-1)^{n-1}\Delta_n z_n(s) \right), \end{aligned} \quad (28)$$

which implies that it is possible to arbitrarily assign the polynomial which defines the zeros of the system.

(Z3)$\Rightarrow$(Z2). If the zeros of system (27) can be arbitrarily assigned, then the polynomial $Z(s)$ in equation (28) can be arbitrarily assigned, and this implies that the polynomials $z_i(s)$ are independent, hence observability of system (21). ∎

Example 5: Consider a reduced order model described by the equations (15), with

$$S = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$L = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$ and $C\Pi = \begin{bmatrix} \eta_{0s} & \eta_{0c} & \eta_0 \end{bmatrix}$. This implies that the reduced order model matches the 0-moment at $\pm i$ and at zero. Note that

$$Z(s) = \eta_0\Delta_3 + (\eta_{0s}\Delta_2 - \eta_{0c}\Delta_1)s + (\eta_{0s}\Delta_1 + \eta_{0c}\Delta_2 + \eta_0\Delta_3)s^2,$$

and this can be arbitrarily assigned provided $\eta_0(\eta_{0s}^2 + \eta_{0c}^2) \neq 0$, which implies, and is implied by, observability of the system $\dot{\xi} = S\xi$, $\psi = C\Pi\xi$.

5) Matching with a passivity constraint: Consider now the problem of selecting Δ such that system (15) is passive. This problem, addressed from a different perspective in [28], [60], admits the following characterization.

Theorem 4: The following statements hold.

(P1) The family of reduced order models (15) contains a lossless system if and only if there exists a symmetric and positive definite matrix P such that

$$S'P + PS = \Pi'C'L + L'C\Pi. \quad (29)$$

(P2) The family of reduced order models (15) contains a passive system if and only if there exists a symmetric and positive definite matrix P such that

$$S'P + PS \leq \Pi'C'L + L'C\Pi. \quad (30)$$

Proof: By definition of losslessness [77], [78], the family of systems (15) contains a lossless system if and only if there exists a symmetric and positive definite matrix P such that

$$(S - \Delta L)'P + P(S - \Delta L) = 0, \qquad P\Delta = (C\Pi)', \quad (31)$$

which is equivalent to the existence of a symmetric and positive definite matrix P such that equation (30) holds. The proof of the second statement is similar hence omitted. ∎

Remark 8: In [28] it has been shown that if the system for which a reduced order model has to be constructed is passive then it is always possible to obtain a reduced order model achieving moment matching and which is itself passive. The result in Theorem 4 is applicable even if the system to be reduced is not passive, or lossless. Note that the approximation of a general system by means of a passive system may be useful to obtain a *physical* realization of the system.

Corollary 2: Consider the family of system (15). Suppose $\sigma(S) \subset \mathbb{C}^-$. Then there is a symmetric and positive definite matrix P such that condition (30) holds.

Proof: The condition $\sigma(S) \subset \mathbb{C}^-$ implies that there exists a symmetric and positive definite matrix X such that $S'X + XS < 0$. As a result, there exists a constant $\kappa > 0$ such that condition (30) holds with $P = \kappa X$. ∎

Example 6: Consider a reduced order model described by the equations (15), with S, L and $C\Pi$ as in equations (16) and (17). Assume $\eta_0 > 0$[3]. Setting

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{12} & P_{22} & P_{23} \\ P_{13} & P_{23} & P_{33} \end{bmatrix},$$

[3]This is the case, for example, if the system to be reduced is strictly passive.

condition (30) reduces to

$$\left[\begin{array}{ccc} 2\eta_0 & \eta_1 - P_{11} & \eta_2 - P_{12} \\ \eta_1 - P_{11} & -2P_{21} & -P_{13} - P_{22} \\ \eta_2 - P_{12} & -P_{13} - P_{22} & -2P_{23} \end{array}\right] \geq 0.$$

This can be satisfied selecting $P_{13} = -P_{22}$, $P_{12} < 0$ and sufficiently large in modulus, and $P_{23} < 0$ and sufficiently large in modulus. Finally, P can be rendered positive definite selecting $P_{11} = 1$, $P_{22} > 0$ and sufficiently large, and $P_{33} > 0$ and sufficiently large.

Example 7: Consider a reduced order model described by the equations (15), with S, L and $C\Pi$ as in equations (16) and (17). Assume $\eta_0 = \eta_2 = 0$ and $\eta_1 > 0$. The selection

$$P = \left[\begin{array}{ccc} \eta_1 & 0 & -P_{22} \\ 0 & P_{22} & 0 \\ -P_{22} & 0 & P_{33} \end{array}\right],$$

with $P_{22} > 0$, and $P_{33} > 0$ and sufficiently large, is such that condition (29) holds with $P > 0$.

6) Matching with L_2-gain: In this section we show how the result developed so far can be exploited to derive an asymptotically stable reduced order model achieving moment matching and with a given L_2-gain.

Theorem 5: The following statements are equivalent.

(L1) System (15) has L_2-gain not larger than $\gamma > 0$.

(L2) There is a symmetric and positive definite matrix P such that

$$S'P + PS + \Pi'C'C\Pi \leq \gamma^2 L'L. \tag{32}$$

Proof: System (15) has L_2-gain smaller or equal to $\gamma > 0$ if and only if [79], [80] there exists a symmetric and positive definite matrix P such that

$$(S - \Delta L)'P + P(S - \Delta L) + P\frac{\Delta\Delta'}{\gamma^2}P + \Pi'C'C\Pi \leq 0. \tag{33}$$

Equation (33) can be rewritten as

$$S'P + PS + \Pi'C'C\Pi + \left[\gamma L' - P\frac{\Delta}{\gamma}\right]\left[\gamma L - \frac{\Delta'}{\gamma}P\right] \leq \gamma^2 L'L,$$

which proves the claim. ∎

Remark 9: If, for some $\gamma > 0$, there exists a symmetric and positive definite solution P of the inequality (32) then $\Delta = \frac{1}{\gamma^2}P^{-1}L'$ is such that system (15) has L_2 gain not larger than γ.

Corollary 3: Consider the family of system (15). Suppose $\sigma(S) \subset \mathbb{C}^-$. Then for any $\gamma > 0$ there is a symmetric and positive definite matrix P such that condition (32) holds.

7) Matching with a compartmental constraint: Finally, we show that under specific assumptions it is possible to select the matrix Δ such that the reduced order model is a compartmental system [81], [82], [30].

Theorem 6: Consider the family of reduced order model (15). Assume $S = S'$. The following statements are equivalent.

(C1) The system (21) is observable.

(C2) The system (22) is observable.

(C3) There is a matrix Δ such that system (27) has a diagonal positive realization.

Proof: The equivalence between (C1) and (C2) has been established in Theorem 2.

(C1)⇔(C3). Let H be the Hankel matrix associated to system (27), let H_ν be the square matrix composed of the first ν rows and the first ν columns of H, and let σH_ν be the square matrix composed of the first ν rows and the second to $\nu + 1$ columns of H. By [30, Proposition 2], system (27) admits a positive diagonal realization if and only if $H_\nu > 0$ and the number of positive (negative, zero) eigenvalues of σH_ν is equal to the number of positive (negative, zero) eigenvalues of S. Let $\Delta = (C\Pi)'$ and note that

$$H_\nu = \left[\begin{array}{c} C\Pi \\ \vdots \\ C\Pi S^{\nu-1} \end{array}\right] \left[\begin{array}{ccc} (C\Pi)' & \cdots & (C\Pi S^{\nu-1})' \end{array}\right]$$

and

$$\sigma H_\nu = \left[\begin{array}{c} C\Pi \\ \vdots \\ C\Pi S^{\nu-1} \end{array}\right] S \left[\begin{array}{ccc} (C\Pi)' & \cdots & (C\Pi S^{\nu-1})' \end{array}\right].$$

Hence, H_ν is positive definite and the number of positive (negative, zero) eigenvalues of σH_ν is equal to the number of positive (negative, zero) eigenvalues of S if and only if system (21) is observable.

∎

III. MODEL REDUCTION BY MOMENT MATCHING FOR NONLINEAR SYSTEMS

A. The notion of moment

In this section we derive a nonlinear enhancement of the notion of moment. To this end note that while most of the results in Section II do not have a direct nonlinear counterpart, we can use Lemma 1 to give a definition of moment.

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

$$\dot{x} = f(x, u), \qquad y = h(x), \tag{34}$$

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

$$\dot{\omega} = s(\omega), \qquad \theta = l(\omega), \tag{35}$$

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

$$\dot{\omega} = s(\omega), \quad \dot{x} = f(x, l(\omega)), \quad y = h(x). \tag{36}$$

Suppose, in addition, that $f(0, 0) = 0$, $s(0) = 0$, $l(0) = 0$ and $h(0) = 0$. The signal generator captures the requirement that one is interested in studying the behavior of system (34) only in *specific circumstances*. However, for this to make sense and to provide a generalization of the notion of moment, we need the following assumptions and definitions.

Assumption 1: There is a unique mapping $\pi(\omega)$, locally[4] defined in a neighborhood of $\omega = 0$, which solves the partial differential equation

$$f(\pi(\omega), l(\omega)) = \frac{\partial\pi}{\partial\omega}s(\omega). \tag{37}$$

[4]All statements are local, although global versions can be easily given.

Assumption 1 implies that the interconnected system (36) possesses an invariant manifold, described by the equation $x = \pi(\omega)$. Note that the (well-defined) dynamics of the system restricted to the invariant manifold are described by $\dot{\omega} = s(\omega)$, *i.e.* are a copy of the dynamics of the signal generator (35).

Assumption 2: The signal generator (35) 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 $l(\omega_a(t)) - l(\omega_b(t)) \not\equiv 0$.

Definition 2: Consider system (34) and the signal generator (35). Suppose Assumptions 1 and 2 hold. The function $h(\pi(\omega))$, with $\pi(\cdot)$ solution of equation (37), is *the moment of system (34) at $(s(\omega), l(\omega))$.*

Definition 3: Consider system (34) and the signal generator (35). Suppose Assumption 1 holds. Let the signal generator (35) be such that $s(\omega) = 0$ and $l(\omega) = \omega$. Then the function $h(\pi(\omega))$ is the 0-moment of system (34) at $s^\star = 0$.

The above definitions allow to derive a nonlinear counterpart of Lemma 1.

Theorem 7: Consider system (34) and the signal generator (35). Assume Assumption 2 holds. Assume the zero equilibrium of the system $\dot{x} = f(x, 0)$ is locally exponentially stable and system (35) is Poisson stable[5]. Assume $\omega(0) \neq 0$.

Then Assumption 1 holds and the moment of system (34) at $(s(\omega), l(\omega))$ coincides with the (locally well-defined) steady-state response[6] of the output of the interconnected system (36).

Proof: To begin with note that, under the stated hypotheses, Assumption 1 holds by the center manifold theory [63] and the results in [6]. Moreover, by [6], the steady-state response of the system is (locally) well-defined, and this is given by $\pi(h(\omega(t)))$, hence the claim. ∎

Remark 10: If the equilibrium $x = 0$ of system $\dot{x} = f(x, 0)$ is unstable, it is still possible to define the moment of system (34) at $(s(\omega), l(\omega))$ in terms of the function $\pi(h(\cdot))$, provided the equilibrium $x = 0$ is hyperbolic and the system (35) is Poisson stable, although it is not possible to establish a relation with the steady-state response of the interconnected system (36).

Remark 11: While for linear systems it is possible to define k-moments for every $s^\star \in \mathbb{C}$ and for any $k \geq 0$, for nonlinear systems it may be difficult, or impossible, to provide general statements if the signal $\theta(t)$, generated by system (35), is unbounded. Therefore, if the signal generator is linear we consider only 0-moments for $s^\star \in \mathbb{C}^0$, whereas if the signal generator is nonlinear we assume that it generates bounded trajectories.

Example 8: Consider a linear system described by equations of the form (1) with $x(t) \in \mathbb{R}^n$, $n > 3$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$ and the nonlinear signal generator (35) with $\omega = [\omega_1, \omega_2, \omega_3]'$,

$$s(\omega) = \begin{bmatrix} \dfrac{I_2 - I_3}{I_1}\omega_2\omega_3 \\ \dfrac{I_3 - I_1}{I_2}\omega_3\omega_1 \\ \dfrac{I_1 - I_2}{I_3}\omega_1\omega_2 \end{bmatrix}, \qquad (38)$$

with $I_1 > 0$, $I_2 > 0$, $I_3 > 0$, $I_i \neq I_j$, for $i \neq j$, and

$$l(\omega) = L\omega = \begin{bmatrix} L_1 & L_2 & L_3 \end{bmatrix} \omega, \qquad (39)$$

with $L_1 L_2 L_3 \neq 0$. This signal generator, which describes the evolution of the angular velocities of a free rigid body in space, is Poisson stable and, under the stated assumption on L, observable [83], [84].

Suppose system (1) is asymptotically stable. The moment of system (1) at $(s(\omega), l(\omega))$ can be computed as follows. Let, formally, $\pi(\omega) = \sum_{i \geq 0} \pi_i(\omega)$, with

$$\pi_i(\omega) = \begin{bmatrix} \pi_i^1(\omega) \\ \pi_i^2(\omega) \\ \vdots \\ \pi_i^n(\omega) \end{bmatrix}$$

and $\pi_i^j(\omega)$ a homogeneous polynomial of degree i in ω. Then equation (37) yields

$$\pi_1(\omega) = -A^{-1}BL\omega, \ \cdots \ \pi_i(\omega) = -A^{-i}BL\frac{d^{i-1}\omega}{dt^{i-1}}, \ \cdots .$$

Hence, the moment of system (1) at $(s(\omega), l(\omega))$ is given, formally, by

$$C\pi(\omega) = -CA^{-1}\left(BL\omega + \cdots + A^{-i+1}BL\frac{d^{i-1}\omega}{dt^{i-1}}\cdots \right),$$

which is a polynomial series in ω.

Example 9 (Markov parameters of a nonlinear system): For the linear system (1) the k-moments at infinity are defined as $\eta_k(\infty) = CA^kB$, *i.e.* the first $k+1$ moments at infinity coincide with the first $k+1$ Markov parameters [1]. To obtain a nonlinear counterpart of this notion recall that[7]

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

where $y_I(\cdot)$ denotes the impulse response of the system and $y_{F,B}(\cdot)$ denotes the free output response from $x(0) = B$.

Consider now a nonlinear affine system[8] described by equations of the form

$$\dot{x} = f(x) + g(x)u, \qquad y = h(x), \qquad (40)$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$ and $f(\cdot)$, $g(\cdot)$ and $h(\cdot)$ smooth mappings. Integrating the first of equations (40), with $x(0) = 0$ and $u(t) = \delta_0(t)$, evaluating for $t = 0$, and substituting in the second equation yields $y_I(0) = y_I^0(0) = h(g(0))$ and[9] $y_I^{(k)}(0) = L_f^k h \circ g(0)$, for $k \geq 0$. It is therefore

---

[5]See [64, Chapter 8] for the definition of Poisson stability.
[6]See [6] for the definition.

[7]To simplify notation, in this example we use $y^{(i)}(t)$ to denote the i-th order time derivative of $y(\cdot)$ at time t. Moreover, all time derivatives at $t = 0$ have to be understood as at $t = 0^+$.

[8]We focus on affine systems since for non-affine systems the impulse response, and its derivatives, may not be well-defined functions (*e.g.* they may be distributions). To illustrate this consider the system $\dot{x} = u^2$, $y = x$, with $x(t) \in \mathbb{R}$, $u(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$. Setting $u(t) = \delta_0(t)$, where $\delta_0(t)$ denotes the Dirac δ-function, and integrating, yields

$$y_I(t) = \int_0^t \delta_0^2(\tau)d\tau!$$

See [85, Chapter 10], and references therein, for an in-depth discussion on the above issue.

[9]$L_f h(\cdot)$ denotes the Lie derivative of the smooth function $h(\cdot)$ along the smooth vector field $f(\cdot)$, as defined in [64, Chapter 1].

natural to define the k-moment at infinity, for $k \geq 0$, of the nonlinear system (40) as $\eta_k(\infty) = y_I^{(k)}(0)$. Note finally that, for the considered class of nonlinear systems, $y_I^{(k)}(0) = y_{F,g(0)}^{(k)}(0)$, where $y_{F,g(0)}(t)$ denotes the free output response of the system from $x(0) = g(0)$.

These considerations allow to derive a reduced order model which matches the $0, \cdots, k-1$ moments at infinity of system (40). For, consider a linear system described by equations of the form (10), with $\xi(t) \in I\!\!R^k$, and F, G and H such that $HF^iG = y_I^{(i)}(0)$, for $i = 0, \cdots, k-1$. (Note that the matrices F, G and H can be computed using standard realization algorithms, *e.g.* Ho-Kalman realization algorithm. In addition, these matrices are not uniquely defined, hence it is possible, for example, to assign the eigenvalues of the reduced order model.) As a consequence of the discussion above, the linear system thus constructed is a model of the nonlinear system achieving moment matching at infinity. Note that, the computation of such a reduced order model does not require the solution of any partial differential equation, but simply regularity of the nonlinear system.

To illustrate the above discussion consider the model of an inverted pendulum on a cart, described by the equations

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= \frac{mlx_4^2 \sin x_3 - mg \cos x_3 \sin x_3 + u}{M + m \sin^2 x_3}, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= \frac{g(m+M) \sin x_3 - ml \cos x_3 \sin x_3 x_4^2 - \cos x_3 u}{l(M + m \sin^2 x_3)}, \\
y &= x_1
\end{aligned}
$$

where $x_1(t)$ is the position of the cart, $x_3(t)$ is the angle of the pendulum (with respect to an upward vertical), $x_2(t)$ and $x_4(t)$ are the corresponding, linear and angular, velocities, $u(t)$ is an external input force, and m, M, l and g are positive constants.

The first five moments of this system at infinity can be directly calculated, using the expression above, yielding $\eta_0(\infty) = 0$, $\eta_1(\infty) = \frac{1}{M}$, $\eta_2(\infty) = 0$, $\eta_3(\infty) = \frac{m(glM^2-1)}{l^2M^4}$, $\eta_4(\infty) = 0$. From these expressions it is possible to derive a reduced order (linear) model.

For example, a reduced order model, matching the first two moments, is described by equations of the form (10), with $\xi(t) \in I\!\!R^2$, and

$$
F = \begin{bmatrix} 0 & 1 \\ -f_{21} & -f_{22} \end{bmatrix} \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad H = \begin{bmatrix} \frac{1}{M} & 0 \end{bmatrix}.
$$

The coefficients f_{21} and f_{22} can be selected to assign the eigenvalues of the matrix F, or to achieve other *interpolation* conditions. For example, selecting $f_{21} = -m\frac{glM^2-1}{M^3l^2}$ and $f_{22} = 0$ ensures matching of the first four moments at infinity.

Finally, a four-dimensional reduced order (linear) model achieving matching of the first four moments at infinity, and with the same eigenvalues of the linearization of the nonlinear system around $x = 0$ is described by equations of the form (10), with $\xi(t) \in I\!\!R^4$, and

$$
F = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & g\frac{M+m}{Ml} & 0 \end{bmatrix} \quad G = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}
$$

$$
H = \begin{bmatrix} -\frac{m+glM^3}{l^2M^4} & 0 & \frac{1}{M} & 0 \end{bmatrix}.
$$

Note that this last reduced order model does not coincide with the linearized model around the zero equilibrium.

1) The frequency response of a nonlinear system: The discussion in the previous sections allows to derive a nonlinear enhancement of the notion of frequency response of a linear system. This relies upon the notion of steady-state response of a nonlinear system, as developed in [5], [6].

Consider system (34) and the signal generator (35). Let the signal generator (35) be such that

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

with $\omega(0) \neq 0$, $\omega^\star \neq 0$ and $L_1^2 + L_2^2 \neq 0$. Then, under the hypotheses of Theorem 7, for all $\omega^\star \in I\!\!R$, Assumption 1 holds and the output of the interconnected system (36) converges towards a locally well-defined steady-state response, which, by definition, does not depend upon the initial condition $x(0)$. 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 $h(\pi(\omega(t))) = \sum_{k=-\infty}^{\infty} c_k e^{k\omega^\star t}$. Consider now the operator $\mathcal{P}_+$ which acts on a Fourier series as follows

$$
\mathcal{P}_+\left(\sum_{k=-\infty}^{\infty} \alpha_k e^{k\omega^\star t} \right) = \sum_{k=0}^{\infty} \alpha_k e^{k\omega^\star t}.
$$

With this operator we can define the frequency response of the nonlinear system (34) as

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

This function depends upon the frequency $\omega^\star$, just as in the linear case, and, unlike the linear case, upon the initial condition $\omega(0)$ of the signal generator. Note finally that if the system (34) were linear, hence described by the equations (1), then $F(\omega(0), \omega^\star) = |W(i\omega^\star)| e^{\angle W(i\omega^\star)}$, where $W(s) = C(sI - A)^{-1}B$.

B. Moment matching

Analogously to the linear case, we now introduce the notion of reduced order model and characterize the solution of the model reduction problem by moment matching.

Definition 4: The system

$$
\dot{\xi} = \phi(\xi, u), \qquad \psi = \kappa(\xi), \qquad (41)
$$

with $\xi(t) \in I\!\!R^\nu$, is a *model at* $(s(\omega), l(\omega))$ *of system (34)* if system (41) has the same moment at $(s(\omega), l(\omega))$ as (34). In this case, system (41) is said to *match* the moment of system (34) at $(s(\omega), l(\omega))$. Furthermore, system (41) is a reduced order model of system (34) if $\nu < n$.

Lemma 5: Consider the system (34), the system (41) and the signal generator (35). Suppose Assumptions 1 and 2 hold. System (41) *matches* the moments of (34) at $(s(\omega), l(\omega))$ if the equation

$$\phi(p(\omega), l(\omega)) = \frac{\partial p}{\partial \omega} s(\omega) \qquad (42)$$

has a unique solution $p(\cdot)$ such that

$$h(\pi(\omega)) = \kappa(p(\omega)), \qquad (43)$$

where $\pi(\cdot)$ is the (unique) solution of equation (37).

Proof: The claim is a direct consequence of the definition of moment. ∎

C. Model reduction by moment matching

In this section we provide a nonlinear counterpart of the construction in Section II-C. For, note that to construct a reduced order model it is necessary to determine mappings $\phi(\cdot, \cdot)$, $\kappa(\cdot)$ and $p(\cdot)$ such that equations (42) and (43) hold, where $\pi(\cdot)$ is the solution of equation (37).

To solve this problem we make the following assumption.

Assumption 3: There exists mappings $\kappa(\cdot)$ and $p(\cdot)$ such that $k(0) = 0$, $p(0) = 0$, $p(\cdot)$ is locally C^1, equation (43) holds and $\det \frac{\partial p(\omega)}{\partial \omega}(0) \neq 0$, *i.e.* the mapping $p(\cdot)$ possesses a local inverse $p^{-1}(\cdot)$.

Remark 12: Similar to the linear case Assumption 3 holds selecting $p(\omega) = \omega$ and $k(\omega) = h(\pi(\omega))$.

A direct computation shows that a family of reduced order models, all achieving moment matching, provided equation (42) has a unique solution $p(\cdot)$, is described by

$$\dot{\xi} = \phi(\xi) + \frac{\partial p(\omega)}{\partial \omega} \delta(\xi) u, \qquad \psi = \kappa(\xi),$$

where $\kappa(\cdot)$ and $p(\cdot)$ are such that Assumption 3 holds, $\delta(\xi) = \tilde{\delta}(p^{-1}(\xi))$, where $\tilde{\delta}(\cdot)$ is a free mapping, and

$$\phi(\xi) = \left[\frac{\partial p(\omega)}{\partial \omega} \left(s(\omega) - \delta(p(\omega)) l(\omega) \right) \right]_{\omega = p^{-1}(\xi)}.$$

In particular, selecting $p(\omega) = \omega$ yields a family of reduced order models described by[10]

$$\dot{\xi} = s(\xi) - \delta(\xi) l(\xi) + \delta(\xi) u, \qquad \psi = h(\pi(\xi)), \qquad (44)$$

where $\delta(\cdot)$ is any mapping such that the equation

$$s(p(\omega)) - \delta(p(\omega)) l(p(\omega)) + \delta(p(\omega)) l(\omega) = \frac{\partial p}{\partial \omega} s(\omega) \quad (45)$$

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

Similarly to what discussed for the linear case in Section II-C, it is possible to use the *parameter* $\delta(\cdot)$ to achieve specific properties of the reduced order model. In what follows, we implicitly assume that the $\delta(\cdot)$ achieving a specific property is such that equation (45) has a unique solution.

As already noted, in the nonlinear case it is not possible to obtain the simple characterizations given in Section II-C, and to address and solve the same problems. On the contrary, there are problems that are interesting only in the nonlinear framework.

[10]Note that this family of models has the same structure as the family given in equation (15).

1) Matching with asymptotic stability: Consider the problem of achieving model reduction by moment matching with a reduced order model, described by equations of the form (44), such that the model has an asymptotically stable equilibrium. Such a reduced order model can be constructed selecting, if possible, the free mapping $\delta(\cdot)$ such that the zero equilibrium of the system (recall that $s(0) = 0$ and $l(0) = 0$) $\dot{\xi} = s(\xi) - \delta(\xi) l(\xi)$ is locally asymptotically stable. This is possible, for example, if the pair $\left(\frac{\partial l(\xi)}{\partial \xi}(0), \frac{\partial s(\xi)}{\partial \xi}(0) \right)$ is observable, or detectable. Note, however, that this is not necessary.

2) Matching with prescribed relative degree: Consider the problem of selecting the mapping $\delta(\cdot)$ in system (44) such that the reduced model has a given relative degree $r \in [1, \nu]$ at some point ξ_0. For such a problem the following fact holds.

Theorem 8: Consider the following statements.

(RD1)* For all $r \in [1, \nu]$ there exists a $\delta(\cdot)$ such that system (44) has relative degree r at ξ_0.

(RD2)* The codistribution

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

has dimension ν at ξ_0.

(RD3)* The system (44) is locally observable at ξ_0.

(RD4)* The system (36) is locally observable at ξ_0.

Then (RD1)*⇔(RD2)*⇒(RD3)* ⇐(RD4)*.

Proof: (RD1)*⇒(RD2)*. To begin with note that system (44) has relative degree r at ξ_0 if and only if system

$$\dot{\xi} = s(\xi) + \delta(\xi) u, \qquad \psi = h(\pi(\xi)), \qquad (46)$$

has relative degree r at ξ_0. By definition of relative degree [64, Chapter 4] system (46) has relative degree r if

$$L_\delta h(\pi(\xi)) = L_\delta L_s h(\pi(\xi)) = \cdots = L_\delta L_s^{r-2} h(\pi(\xi)) = 0, \qquad (47)$$

for all ξ in a neighborhood of ξ_0, and

$$L_\delta L_s^{r-1} h(\pi(\xi_0)) \neq 0. \qquad (48)$$

Equations (47) and (48) can be rewritten as

$$\begin{bmatrix} dh(\pi(\xi)) \\ \vdots \\ dL_s^{r-2} h(\pi(\xi)) \\ dL_s^{r-1} h(\pi(\xi)) \end{bmatrix} \delta(\xi) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \gamma(\xi) \end{bmatrix},$$

for some function $\gamma(\cdot)$ such that $\gamma(\xi_0) \neq 0$. This equation has a solution for all $r \in [1, n]$ if and only if the matrix in the left hand side has rank r at ξ_0. Setting $r = \nu$ yields that the codistribution $d\mathcal{O}_\nu$ has dimension ν.

(RD2)*⇒(RD1)*. Since $d\mathcal{O}_\nu$ has rank ν and it is spanned by ν differentials, for all $r \in [1, \nu]$ the distribution

$$d\mathcal{O}_i = \text{span}\{dh(\pi(\xi)), \cdots, dL_s^{i-1} h(\pi(\xi))\}$$

has dimension i, which implies that the equations (47) and (48) have a solution $\delta(\cdot)$ locally well-defined around ξ_0, hence the claim.

(RD2)*⇒(RD3)*. This implication follows from [83, Theorem 3.32].

(RD4)*⇒(RD3)*. We prove this statement by contradiction. Suppose system (44) is not observable at ξ_0 then there is a

point ξ_a (in a neighborhood of ξ_0) which is indistinguishable from ξ_0 (recall [83, Definitions 3.27 and 3.28]). Consider now system (36) and the points $[\xi_0', \pi(\xi_0)']'$ and $[\xi_a', \pi(\xi_a)']'$. Simple computations, recalling that the manifold $x = \pi(\xi)$ is invariant, show that these points are not distinguishable, hence system (36) is not locally observable at ξ_0, which proves the claim. ∎

Remark 13: Although the implication $(RD3)^* \Rightarrow (RD2)^*$ does not hold in general, $(RD3)^*$ implies that the codistribution $d\mathcal{O}_\nu$ has dimension ν for all ξ in an open and dense set around ξ_0 (see [83, Corollary 3.35]).

3) Matching with prescribed zero dynamics: Consider the problem of selecting the mapping $\delta(\cdot)$ in system (44) such that the reduced model has zero dynamics with specific properties. To simplify the study of this problem we assume that condition $(RD2)^*$ holds, which allows to obtain a special form for system (44).

Lemma 6: Consider system (44). Assume condition $(RD2)^*$ holds. Then there exists a coordinates transformation $\chi = \Xi(\xi)$, locally defined around ξ_0, such that, in the new coordinates, system (44) is described by equations of the form

$$
\begin{aligned}
\dot{\chi}_1 &= \chi_2 + \tilde{\delta}_1(\chi)(v - \tilde{l}(\chi)), \\
\dot{\chi}_2 &= \chi_3 + \tilde{\delta}_2(\chi)(v - \tilde{l}(\chi)), \\
&\vdots \\
\dot{\chi}_\nu &= \tilde{f}(\chi) + \tilde{\delta}_\nu(\chi)(v - \tilde{l}(\chi)), \\
\psi &= \chi_1,
\end{aligned} \tag{49}
$$

where $[\tilde{\delta}_1(\chi), \cdots, \tilde{\delta}_\nu(\chi)]' = \delta(\Xi^{-1}(\chi))$, $\tilde{l}(\chi) = l(\Xi^{-1}(\chi))$, and $\tilde{f}(\chi) = L_s^\nu h(\pi(\Xi^{-1}(\chi)))$.

Proof: By condition $(RD2)^*$ the functions $h(\pi(\cdot))$, ..., $L_s^{\nu-1} h(\pi(\cdot))$ qualify as local coordinates around ξ_0. In these coordinates the system (44) is trivially described by equations of the form (49). ∎

As a consequence of the result established in Lemma 6 we have the following statement.

Proposition 2: Consider system (44). Assume condition $(RD2)^*$ holds and ξ_0 is an equilibrium of system (44). Then, for all[11] $r \in [1, \nu-1]$, there is a $\delta(\cdot)$ such that system (44) has relative degree r and its zero dynamics have a locally exponentially stable equilibrium. In addition, there is a coordinates transformation, locally defined around ξ_0, such that, in the new coordinates, the zero dynamics of system (44) are described by equations of the form

$$
\begin{aligned}
\dot{z}_1 &= z_2 + \hat{\delta}_1(z)z_1, \\
\dot{z}_2 &= z_3 + \hat{\delta}_2(z)z_1, \\
&\vdots \\
\dot{z}_{\nu-r} &= \hat{f}(z) + \hat{\delta}_{\nu-r}(z)z_1,
\end{aligned} \tag{50}
$$

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

$$
\hat{f}(z) = \tilde{f}(\chi) \Big|_{\chi = [0, \cdots, 0, z_1, \cdots, z_{\nu-r}]'}.
$$

Proof: By Lemma 6, condition $(RD2)^*$, and invariance of the zero dynamics with respect to coordinates transformation, we can consider the system described by equations of the form

---
[11]For $r = \nu$ the system does not have zero dynamics.

(49). Fix now $r \in [1, \nu-1]$. Select $\tilde{\delta}_i(\cdot)$ such that $\tilde{\delta}_r(\chi_0) \neq 0$ and $\tilde{\delta}_j(\chi) = 0$, for all[12] $j \in [1, r-1]$. The resulting system has relative degree r and the zero dynamics are described by the last $\nu - r$ equations of system (49) with $v = \tilde{l}(\chi)$ and $\chi_1 = \ldots = \chi_r = 0$, which have the form given in equation (50). To complete the proof, recall that ξ_0 is (by assumption) an equilibrium, hence, the system (50) has an equilibrium which can be rendered locally exponentially stable by a proper selection of the functions $\hat{\delta}_1(\cdot), \cdots, \hat{\delta}_{\nu-r}(\cdot)$. ∎

4) Matching with a passivity constraint: Consider now the problem of selecting the mapping $\delta(\cdot)$ such that system (44) is lossless or passive. For such a problem the following fact holds.

Theorem 9: Consider the following statements.

(P1)* The family of reduced order models (44) contains, locally around ξ_0, a lossless (passive, respectively) system with a differentiable storage function.

(P2)* There exists a differentiable function $V(\cdot)$, locally positive definite around ξ_0, such that[13]

$$
dV_\xi s(\xi) = h(\pi(\xi))l(\xi), \quad (V_\xi s(\xi) \leq h(\pi(\xi))l(\xi) \text{ resp.}), \tag{51}
$$

locally around ξ_0.

(P3)* There exists a differentiable function $V(\cdot)$, locally positive definite around ξ_0, such that equation (51) holds and

$$
V_{\xi\xi}(\xi_0) > 0. \tag{52}
$$

Then $(P1)^* \Rightarrow (P2)^*$, $(P3)^* \Rightarrow (P2)^*$, and $(P3)^* \Rightarrow (P1)^*$.

Proof: We provide the proof for the lossless case, since the claims on passivity can be proved with identical arguments. $(P1)^* \Rightarrow (P2)^*$. As detailed in [48], if the family of systems (44) contains, locally around ξ_0, a lossless system with a differentiable storage function $V(\cdot)$ then there exists $\delta(\cdot)$ such that $V_\xi(s(\xi) - \delta(\xi)l(\xi)) = 0$ and $V_\xi \delta = h(\pi(\xi))$, locally around ξ_0. Replacing the second equality in the first yields the claim.

$(P3)^* \Rightarrow (P2)^*$. This implication is trivial.

$(P3)^* \Rightarrow (P1)^*$. Equation (52) implies that the equation $V_\xi \delta(\xi) = h(\pi(\xi))$, in the unknown $\delta(\cdot)$, has a (unique) solution $\delta^\star(\cdot)$, which is locally well-defined, and continuous, around ξ_0. Consider now system (44) with $\delta(\cdot) = \delta^\star(\cdot)$. Then $V_\xi(s(\xi) - \delta^\star(\xi)l(\xi)) = 0$, and this, together with the definition of $\delta^\star(\cdot)$, implies $(P1)^*$. ∎

5) Matching with L_2-gain: We now consider the problem of selecting the mapping $\delta(\cdot)$ such that system (44) has a given L_2-gain.

Theorem 10: Consider the following statements.

(L1)* The family of reduced order models (44) contains, locally around ξ_0, a system with L_2-gain not larger than $\gamma > 0$ and with a differentiable storage function.

(L2)* There exists a differentiable function $V(\cdot)$, locally positive definite around ξ_0, such that equation (52) holds and such that

$$
V_\xi s(\xi) + (h(\pi(\xi)))^2 \leq \gamma^2 l^2(\xi), \tag{53}
$$

---
[12]If $r = 1$ this set is empty.

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

locally around ξ_0.

Then $(L2)^* \Rightarrow (L1)^*$.

The proof of this statement is similar to the proof of Theorem 9, hence it is omitted.

6) Matching for linear systems at $(s(\omega), l(\omega))$: In this section we consider the model reduction problem for linear systems at $(s(\omega), l(\omega))$, *i.e.* we consider the case in which the system to be reduced is linear and the signal generator is a nonlinear system. For such a problem, under suitable assumptions, it is possible to obtain in an explicit way a formal description of reduced order models, as detailed in the following statement.

Proposition 3: Consider the linear system (1), with $x(t) \in \mathbb{R}^n$ and $\sigma(A) \subset \mathbb{C}^-$. Consider the signal generator (35), with $\omega(t) \in \mathbb{R}^\nu$, $n > \nu$ and $l(\omega) = L\omega$. Assume that the signal generator is Poisson stable and that $s(\cdot)$ can be expressed, locally around $\omega = 0$, as a formal power series, *i.e.* $s(\omega) = \sum_{i \geq 1} s^{[i]}(\omega)$, where $s^{[i]}(\cdot)$ denotes a polynomial vector field which is homogeneous of degree i. Suppose in addition that $s^{[1]}(\omega) = 0$. Then a family of reduced order models achieving moment matching at $(s(\omega), l(\omega))$ is formally described by the equations

$$\dot{\xi} = s(\xi) - \delta(\xi)L\xi + \delta(\xi)u \qquad \psi = C\pi(\xi)$$

with $\delta(\cdot)$ a free mapping and

$$\pi(\xi) = \sum_{i \geq 1} \pi^{[i]}(\xi), \tag{54}$$

where

$$\pi^{[1]}(\xi) = -A^{-1}BL\xi \tag{55}$$

and

$$\pi^{[k]}(\xi) = A^{-1} \sum_{i=1}^{k-1} \frac{\partial \pi^{[i]}(\xi)}{\partial \xi} s^{[k-i+1]}(\xi), \tag{56}$$

for $k \geq 2$.

Proof: Note that equation (37) can be written as $A\pi(\omega) + BL\omega = \frac{\partial \pi}{\partial \omega}s(\omega)$. In addition, by the stated assumptions and Theorem 7, this equation has a unique solution locally around $\omega = 0$. Then, by a direct computation, we conclude that the solution $\pi(\cdot)$ of this equation admits the formal power series description given by equations (54), (55) and (56). The result then follows from the discussion in Section III-C. ∎

Example 10 (Example 8 continued): Exploiting the results in Proposition 3 and the discussion in Example 8, we infer that a reduced order model for a linear asymptotically stable system at the $(s(\omega), l(\omega))$ given in equations (38) and (39) is described by

$$\begin{aligned} \dot{\xi} &= s(\xi) - \delta(\xi)L\xi + \delta(\xi)u, \\ \psi &= -CA^{-1}\left[BL\omega + \cdots + A^{-i+1}BL\frac{d^{i-1}\omega}{dt^{i-1}}\cdots\right]_{\omega=\xi}. \end{aligned}$$

Simulations have been run selecting $I_1 = 1$, $I_2 = 2$, $I_3 = 3$, $L_1 = 1$, $L_2 = 1/2$, $L_3 = 1/3$ and $\delta(\cdot) = \text{diag}(1/I_1, 1/I_2, 1/I_3)L'$, which yields (see [86] for a proof) a reduced order model with a globally asymptotically stable equilibrium at $\xi = 0$.

The linear system, that has to be reduced, is a randomly selected asymptotically stable system of dimension 15. The initial conditions of the signal generator have been selected as $\omega(0) = \frac{1}{5}[1\ 1\ 1]'$.

The linear system and the reduced order model, both driven by the signal generator, have been numerically integrated from zero initial conditions. Figure 2 (left) displays the output $y(t)$ of the linear system when driven by the signal generator, and the signals $\psi^{[1]}(t)$, $\psi^{[2]}(t)$ and $\psi^{[3]}(t)$, obtained by truncating the formal power series defining $\psi(t)$ to the first, second and third order terms, respectively. Figure 2 (right) displays the approximation errors $y(t) - \psi^{[1]}(t)$, $y(t) - \psi^{[2]}(t)$, $y(t) - \psi^{[3]}(t)$. Note that, in steady-state,

$$\begin{aligned} \max(|y(t) - \psi^{[I]}(t)|) &\approx 0.276 > \\ \max(|y(t) - \psi^{[II]}(t)|) &\approx 0.164 > \\ \max(|y(t) - \psi^{[III]}(t)|) &\approx 0.076, \end{aligned}$$

which shows that the approximation error decreases by adding terms in the formal power series defining the output of the reduced order model.

7) Matching for nonlinear systems at $S\omega$: In this section we consider the model reduction problem for nonlinear systems at $s(\omega) = S\omega$, *i.e.* we consider the case in which the signal generator is a linear system.

This problem is of particular interest since, exploiting the discussion in Section III-C, we infer that the reduced order models have a very simple description, *i.e.* a family of reduced order models is given by the equations $\dot{\xi} = (S - \delta(\xi)L)\xi + \delta(\xi)u, \psi = h(\pi(\omega))$, where $\delta(\xi)$ is a free mapping. In particular, selecting $\delta(\xi) = \Delta$, for some constant matrix Δ, we have that the family of reduced order models is described by a linear differential equation with a nonlinear output map. This structure has two main advantages. The former is that the matrix Δ can be selected to achieve additional goals, such as to assign the eigenvalues or the relative degree of the reduced order model (provided additional assumptions on the output map holds). The latter is that the computation of (an approximation of) the reduced order model boils down to the computation of (an approximation of) the output map $h(\pi(\omega))$. This computation can be carried out in the spirit of the results in [50, Section 4.2 and 4.3]. We complete this section discussing the model reduction problem with 0-moment matching at $s^\star = 0$, *i.e.* the model reduction problem at $s(\omega) = 0$. This problem can be solved, under specific assumptions, without solving any partial differential equation, as detailed in the following statement.

Proposition 4 (0-moment matching at $s^\star = 0$): Consider system (34) and the signal generator $\dot{\omega} = 0$, $\theta = \omega$. Assume the zero equilibrium of the system $\dot{x} = f(x, 0)$ is locally exponentially stable. Then the zero moment at $s^\star = 0$ of system (34) is (locally) well-defined and given by $h(\pi(\cdot))$, with $\pi(\cdot)$ the unique solution of the algebraic equation $f(\pi(\omega), \omega) = 0$. Finally, a reduced order model, for which the zero equilibrium is locally asymptotically stable is given by $\dot{\xi} = -\delta(\xi)(\xi - u), \psi = h(\pi(\xi))$, with $\delta(\cdot)$ such that $\delta(0) > 0$.

Proof: We simply need to show that equation (42) has a unique solution. For, note that in this case, equation (42) rewrites as $-\delta(p(\omega))(p(\omega) - \omega)) = 0$, which, by positivity of $\delta(0)$, has indeed a unique solution (locally). ∎
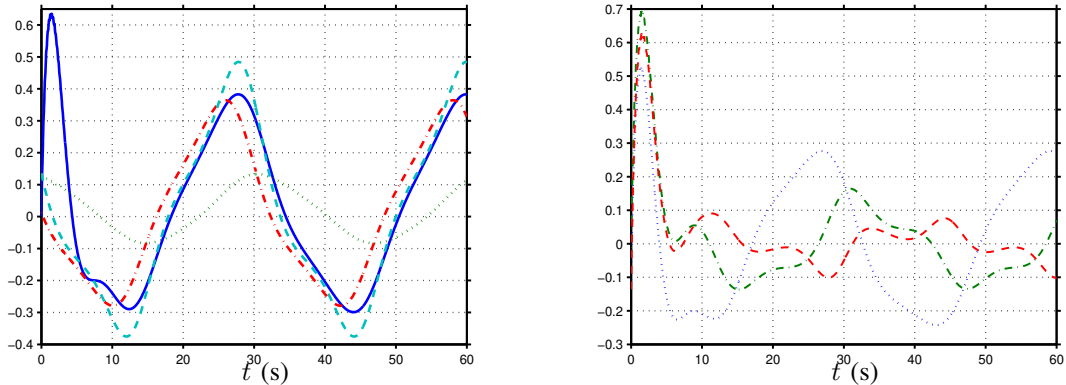
Fig. 2. Time histories of the output of the driven linear system and of the *approximating* outputs of the driven reduced order model (left): $y(t)$ (solid), $\psi^{[I]}(t)$ (dotted), $\psi^{[II]}(t)$ (dash-dotted) and $\psi^{[III]}(t)$ (dashed). Time histories of the approximation errors (right): $y(t) - \psi^{[I]}(t)$ (dotted), $y(t) - \psi^{[II]}(t)$ (dash-dotted), $y(t) - \psi^{[III]}(t)$ (dashed).

Example 11: The averaged model of the DC–to–DC Ćuk converter is given by the equations [87]

$$L_1 \frac{d}{dt} i_1 = -(1-u) v_2 + E,$$
$$C_2 \frac{d}{dt} v_2 = (1-u) i_1 + u i_3,$$
$$L_3 \frac{d}{dt} i_3 = -u v_2 - v_4, \qquad (57)$$
$$C_4 \frac{d}{dt} v_4 = i_3 - G v_4,$$
$$y = v_4,$$

where $i_1(t) \in \mathbb{R}^+$ and $i_3(t) \in \mathbb{R}^-$ describe currents, $v_2(t) \in \mathbb{R}^+$ and $v_4(t) \in \mathbb{R}^-$ voltages, L_1, C_2, L_3, C_4, E and G positive parameters and $u(t) \in (0,1)$ a continuous control signal which represents the slew rate of a PWM circuit used to control the switch position in the converter. The 0-moment of the system at $s^\star = 0$ is $h(\pi(\omega)) = \frac{\omega}{\omega - 1} E$, and a locally asymptotically stable reduced order model achieving moment matching at $s^\star = 0$ is

$$\dot{\xi} = -\delta(\xi)(\xi - u), \qquad \psi = E \frac{\xi}{\xi - 1}, \qquad (58)$$

with $\delta(0) > 0$, which is well-defined if $\xi \neq 1$. This is consistent with the fact that the 0-moment at $s^\star = 0$ is defined for $\omega \neq 1$. Simulations have been run to assess the properties of the reduced order model. The parameters have been selected as in [87], the input signal is piecewise constant, with jumps every 0.05 seconds. The reduced order model is described by equations (58), where the function $\delta(\cdot)$ depends upon the input signal u and it is equal to the real part of the slowest eigenvalue of the system (57) (which is a linear system for constant u). Figure 3 (top) displays the output $y(t)$ of the averaged model of the Ćuk converter and of the output of the reduced order model $\psi(t)$. The figure shows that the reduced order model provides a good *static* approximation of the behavior of the system but does not capture its dynamic (under-damped) behavior. The dynamic behavior can be captured constructing a second order model, which is (in the spirit of the model (58)) a linear system with a nonlinear output map. Since such a model is required to match only one moment, it is possible to assign

its eigenvalues at the location of the dominant modes of system (57) with u fixed. The output $\psi_2(t)$ of this two dimensional reduced order model is also displayed in Figure 3. Note that this signal may provide a better approximation of $y(t)$, as shown in Figure 3 (bottom), in which the errors $y(t) - \psi(t)$ and $y(t) - \psi_2(t)$ are plotted.
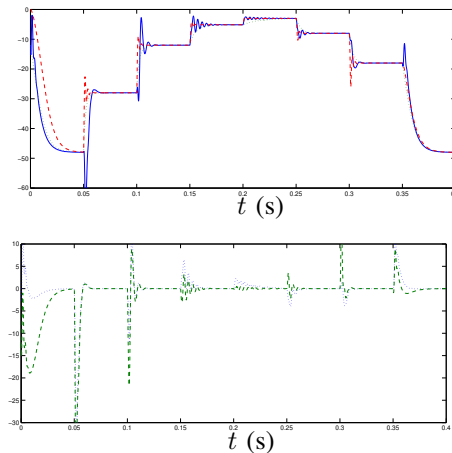

Fig. 3. Time histories of the output of the averaged model of the Ćuk converter and of the *approximating* outputs of the reduced order models (top): $y(t)$ (solid), $\psi(t)$ (dotted), $\psi_2(t)$ (dashed). Time histories of the approximation errors (bottom): $y(t) - \psi(t)$ (dotted), $y(t) - \psi_2(t)$ (dashed).

IV. CONCLUSIONS

The problem of model reduction for linear and nonlinear systems has been posed and solved exploiting the notion of moment. To this end, a nonlinear enhancement of the notion of moment has been developed. The linear and nonlinear problems have been dealt with from a unique perspective, which shades new light on the linear problem, and allows for a simple solution of the nonlinear one. Special emphasis has been devoted to the problem of parameterizing a family of reduced order models, and to the problem of selecting a member of the family with specific properties. This type of problems, in the linear framework, has been solved with projection methods and in the framework of interpolation theory, which however

do not provide a unifying perspective, *i.e.* the selection of projectors and of interpolation parameters does not have a simple interpretation. Several issues or research directions are left open in the present paper: it may be possible to exploit the connection between moment matching and interpolation theory, available for linear systems, to develop a nonlinear interpolation theory; it may be possible to extend the notion of moment to nonlinear system with unstable, but hyperbolic, equilibria (as briefly mentioned in Remark 10); the model reduction problem for specific classes of systems, for example Hamiltonian systems or periodic systems, may be studied; the multivariable version of the proposed results can be developed; the use of reduced order models in control design has to be investigated, especially in the nonlinear case. Finally, the engineering relevance of the theory should be evaluated with more realistic case studies.

REFERENCES

[1] A. Antoulas, *Approximation of Large-Scale Dynamical Systems*. SIAM Advances in Design and Control, 2005.

[2] T. Hu, A. Teel, and Z. Lin, "Lyapunov characterization of forced oscillations," *Automatica*, vol. 41, pp. 1723–1735, 2005.

[3] L. Chua and C. Ng, "Frequency-domain analysis of nonlinear systems: General theory. formulation of transfer functions," *IEE Journal on Electronic Circuits and Systems*, pp. 165–185, 1979.

[4] C. Byrnes and A. Isidori, "A frequency domain philosophy for nonlinear systems, with application to stabilization and to adaptive control," in *23rd Conference on Decision and Control*, 1985, pp. 1031–1037.

[5] A. Isidori and C. I. Byrnes, "Steady state response, separation principle and the output regulation of nonlinear systems," in *28th Conf. Decision and Control, Tampa, Fl*, 1989, pp. 2247–2251.

[6] ——, "Output regulation of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 35, no. 2, pp. 131–140, 1990.

[7] A. Isidori and A. Astolfi, "Disturbance attenuation and H$_\infty$-control via measurement feedback in nonlinear systems," *IEEE Trans. Autom. Control*, vol. 37, pp. 1283–1293, Sept. 1992.

[8] J. Huang, "Frequency domain approach to the robust regulator theory for volterra systems," in *34th Conference on Decision and Control, New Orleans, LA*, 1995, pp. 1690–1695.

[9] D. Kavranoglu and M. Bettayeb, "Characterization of the solution to the optimal H$_\infty$ model reduction problem," *Systems and Control Letters*, vol. 20, pp. 99–108, 1993.

[10] X.-X. Huang, W.-Y. Yan, and K. Teo, "H$_2$ near optimal model reduction," *IEEE Trans. Autom. Control*, vol. 46, pp. 1279–1285, 2001.

[11] A. Antoulas and A. Astolfi, "H$_\infty$ norm approximation," in *Unsolved problems in Mathematical Systems and Control Theory, V. Blondel and A. Megretski Editors*. Princeton University Press, 2004, pp. 267–270.

[12] S. Gugercin, A. Antoulas, and C. Beattie, "H$_2$ model reduction for large-scale linear dynamical systems," *SIAM Journal of Matrix Analysis*, vol. 30, pp. 609–638, 2008.

[13] J. Scherpen, "H$_\infty$ balancing for nonlinear systems," *Int. J. Rob. Nonl. Contr.*, vol. 6, pp. 645–668, 1996.

[14] V. Adamjan, A. D.Z., and M. 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.

[15] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L$^\infty$-error bounds," *International Journal of Control*, vol. 39, pp. 1115–1193, 1984.

[16] M. Safonov, R. Chiang, and D. Limebeer, "Optimal Hankel model reduction for nonminimal systems," *IEEE Trans. Autom. Control*, vol. 35, no. 4, pp. 496–502, 1990.

[17] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Autom. Control*, vol. 26, no. 1, pp. 17–32, 1981.

[18] D. Meyer, "Fractional balanced reduction: model reduction via a fractional representation," *IEEE Trans. Autom. Control*, vol. 35, no. 12, pp. 1341–1345, 1990.

[19] J. Scherpen, "Balancing for nonlinear systems," *Systems and Control Letters*, vol. 21, pp. 143–153, 1993.

[20] J. Scherpen and A. van der Schaft, "Normalized coprime factorizations and balancing for unstable nonlinear systems," *Int. Journal of Control*, vol. 60, pp. 1193–1222, 1994.

[21] J. Scherpen and W. Gray, "Minimality and local state decompositions of a nonlinear state space realization using energy functions," *IEEE Trans. Autom. Control*, vol. 45, no. 11, pp. 2079–2086, 2000.

[22] W. Gray and J. Scherpen, "On the nonuniqueness of singular value functions and balanced nonlinear realizations," *Systems and Control Letters*, vol. 44, pp. 219–232, 2001.

[23] J. Scherpen and W. Gray, "Nonlinear Hilbert adjoints: Properties and applications to Hankel singular value analysis," *Nonlinear Analysis, Theory, Methods and Applications*, vol. 51, no. 5, pp. 883–901, 2002.

[24] K. Fujimoto and J. Scherpen, "Nonlinear input-normal realizations based on the differential eigenstructure of Hankel operators," *IEEE Trans. Autom. Control*, vol. 50, pp. 2–18, 2005.

[25] W. Gray and J. Scherpen, "Hankel singular value functions from Schmidt pairs for nonlinear input-output systems," *Systems and Control Letters*, vol. 54, pp. 135–144, 2005.

[26] A. Krener, "Model reduction for linear and nonlinear control systems," in *45th Conference on Decision and Control, San Diego, CA*, 2006, bode Lecture.

[27] K. Fujimoto and D. Tsubakino, "Computation of nonlinear balanced realization and model reduction based on Taylor series expansion," *Systems and Control Letters*, vol. 57, pp. 283–289, 2008.

[28] A. Antoulas, "A new result on passivity preserving model reduction," *Systems and Control Letters*, vol. 54, pp. 361–374, 2005.

[29] S. Lall and C. Beck, "Error bounds for balanced model reduction of linear time-varying systems," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 946–956, 2003.

[30] A. Astolfi and P. Colaneri, "A note on the existence of positive realizations," *Linear algebra and its applications*, vol. 390, pp. 329–343, Oct. 2004.

[31] B. Srinivasan and P. Myszkorowski, "Model reduction of systems with zeros interlacing the poles," *Systems and Control Letters*, vol. 30, pp. 19–24, 1997.

[32] E. Verriest and W. Gray, "Geometry and topology of the state space via balancing," in *Mathematical Theory of Networks and Systems, Kyoto, Japan*, 2006.

[33] W. Gray and J. Mesko, "General input balancing and model reduction for linear and nonlinear systems," in *European Control Conference, Brussels, Belgium*, 1997.

[34] S. Al-Baiyat, M. Bettayeb, and U. Al-Saggaf, "New model reduction scheme for bilinear systems," *International Journal of Systems Science*, vol. 25, pp. 1631–1642, 1994.

[35] S. Lall, P. Krysl, and J. Marsden, "Structure-preserving model reduction for mechanical systems," *Physica D*, vol. 184, no. 1–4, pp. 304–318, 2003.

[36] J. Soberg, K. Fujimoto, and T. Glad, "Model reduction of nonlinear differential-algebraic equations," in *7th IFAC Symposium on Nonlinear Control Systems, Pretoria, South Africa*, 2007, pp. 712–717.

[37] K. Fujimoto and H. Kajiura, "Balanced realization and model reduction of port-Hamiltonian systems," in *American Control Conference, New York, NY*, 2007, pp. 930–934.

[38] K. Fujimoto, "Balanced realization and model order reduction for port-Hamiltonian systems," *Journal of System Design and Dynamics*, (to appear) 2008.

[39] E. Verriest and W. Gray, "Dynamics near limit cycles: Model reduction and sensitivity," in *Mathematical Theory of Networks and Systems, Padova, Italy*, 1998.

[40] W. Gray and E. Verriest, "Balanced realizations near stable invariant manifolds," *Automatica*, vol. 42, no. 4, pp. 653–659, 2006.

[41] S. Lall, J. Marsden, and S. Glavaski, "A subspace approach to balanced truncation for model reduction of nonlinear control systems," *Int. J. Robust Nonlinear Control*, vol. 12, pp. 519–535, 2002.

[42] K. Willcox and J. Peraire, "Balanced model reduction via the proper orthogonal decomposition," *AIAA Journal*, vol. 40, no. 11, pp. 2323–2330, 2002.

[43] K. Kunisch and S. Volkwein, "Control of Burgers' equation by reduced order approach using proper orthogonal decomposition," *Journal of Optimization Theory and Applications*, vol. 102, no. 2, pp. 345–371, 1999.

[44] M. Hinze and S. Volkwei, "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, V. M. P. Benner and D. Sorensen, Eds. Springer Verlag, Heidelberg, 2005, pp. 261–306.

[45] K. Kunisch and S. Volkwein, "Proper orthogonal decomposition for optimality systems," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 42, no. 1, pp. 1–23, 2008.

[46] P. Astrid, S. Weiland, K. Willcox, and T. Backx, "Missing point estimation in models described by proper orthogonal decomposition," *IEEE Trans. Autom. Control*, (to appear) 2008.

[47] C. I. Byrnes and A. Isidori, "Asymptotic stabilization of minimum phase nonlinear systems," *IEEE Trans. Autom. Control*, vol. 36, pp. 1122–1137, 1991.

[48] C. I. Byrnes, A. Isidori, and J. C. Willems, "Passivity, feedback equivalence, and the global stabilization of minimum phase nonlinear systems," *IEEE Trans. Autom. Control*, vol. 36, pp. 1228–1240, Nov. 1991.

[49] A. Isidori, "A tool for semiglobal stabilization of uncertain non-minimum-phase nonlinear systems via output feedback," *IEEE Trans. Autom. Control*, vol. 45, pp. 1817–1827, Oct. 2000.

[50] J. Huang, *Nonlinear Output Regulation: Theory and Applications*. SIAM Advances in Design and Control, 2004.

[51] A. Astolfi, "Model reduction by moment matching (semi-plenary presentation)," in *IFAC Symposium on Nonlinear Control System Design, Pretoria, S. Africa*, 2007, pp. 95–102.

[52] ——, "A new look at model reduction by moment matching for linear systems," in *46th Conference on Decision and Control, New Orleans, LS*, 2007, pp. 4361–4366.

[53] ——, "Model reduction by moment matching for nonlinear systems," in *47th Conference on Decision and Control, Cancun, Mexico*, 2008.

[54] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback control theory*. Macmillan Publishing Comp., New York, 1992.

[55] T. Georgiou, "Partial realization of covariance sequences," Ph.D. dissertation, University of Florida, 1983.

[56] C. Byrnes, A. Lindquist, S. Gusev, and A. Matveev, "A complete parameterization of all positive rational extensions of a covariance sequence," *IEEE Trans. Autom. Control*, vol. 40, pp. 1841–1857, 1995.

[57] T. Georgiou, "The interpolation problem with a degree constraint," *IEEE Trans. Autom. Control*, vol. 44, pp. 631–635, 1999.

[58] C. Byrnes, A. Lindquist, and T. Georgiou, "A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint," *IEEE Trans. Autom. Control*, vol. 46, pp. 822–839, 2001.

[59] C. Byrnes, T. Georgiou, A. Lindquist, and A. Megretski, "Generalized interpolation in H^∞ with a complexity constraint," *Trans. Amer. Math. Soc.*, vol. 358, pp. 965–987, 2006.

[60] G. Fanizza, J. Karlsson, A. Lindquist, and R. Nagamune, "Passivity-preserving model reduction by analytic interpolation," *Linear Algebra and Its Applications*, vol. 425, pp. 608–633, 2007.

[61] C. Hespel and G. Jacob, "Approximation of nonlinear dynamic systems by rational series," *Theoret. Comput. Science*, vol. 79, pp. 151–162, 1991.

[62] C. Hespel, "Truncated bilinear approximants: Carleman, finite Volterra, Padé-type, geometric and structural automata," in *Algebraic Computing in Control, G. Jacob and F. Lamnabhi-Lagarrigue (editors)*. Springer-Verlag, Berlin, 1991, pp. 264–278.

[63] J. Carr, *Applications of Center Manifold Theory*. Springer Verlag, 1981.

[64] A. Isidori, *Nonlinear Control Systems, Third Edition*. Springer Verlag, 1995.

[65] K. Gallivan, A. Vandendorpe, and P. Van Dooren, "Model reduction and the solution of Sylvester equations," in *MTNS, Kyoto*, 2006.

[66] I. Smith and T. Lucas, "Least-squares moment matching reduction methods," *Electronics Letters*, vol. 31, no. 11, pp. 929–930, 1995.

[67] A. Antoulas, J. Ball, J. Kang, and J. 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.

[68] P. Feldman and R. Freund, "Efficient linear circuit analysis by Pade approximation via a Lanczos method," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 639–649, 1995.

[69] E. Grimme, D. Sorensen, and P. van Dooren, "Model reduction of state space systems via an implicitly restarted Lanczos method," *Numerical algorithms*, vol. 12, pp. 1–31, 1995.

[70] I. Jaimoukha and E. Kasenally, "Implicitly restarted Krylov subspace methods for stable partial realizations," *SIAM J. Matrix Anal. Appl.*, vol. 18, pp. 123–132, 1997.

[71] A. Antoulas, "On pole placement in model reduction," *preprint*, 2007.

[72] C. Beattie and S. Gugercin, "Interpolation theory for structure-preserving model reduction," in *47th Conference on Decision and Control, Cancun, Mexico*, 2008.

[73] R.-C. Li and Z. Bai, "Structure-preserving model reduction using a Krylov subspace projection formulation," *Comm. Math. Sci.*, vol. 3, no. 2, pp. 179–199, 2005.

[74] H. Kimura, "A canonical form for partial realization of covariance sequences," Osaka University," Tech. Report 83-01, 1983.

[75] ——, "Positive partial realization of covariance sequences," in *Modeling, Identification and Control, C.I. Byrnes and A. Lindquist (editors)*. North Holland, 1986, pp. 499–513.

[76] S. Gugercin and K. Willcox, "Krylov projection framework for Fourier model reduction," *Automatica*, vol. 44, no. 1, pp. 209–215, 2008.

[77] J. C. Willems, "Dissipative dynamical systems," *Arch. Rational Mechanics and Analysis*, vol. 45, pp. 321–393, 1972.

[78] A. van der Schaft, $\mathcal{L}_2$-*Gain and Passivity Techniques in Nonlinear Control*. Springer Verlag, 1996.

[79] B. A. Francis and J. C. Doyle, "Linear control theory with an H_∞-optimal criterion," *Siam J. Contr. amd Optimiz.*, vol. 25, pp. 815–844, 1987.

[80] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-space solutions to standard H_2 and H_∞ control problems," *IEEE Trans. Autom. Control*, vol. 34, pp. 831–847, Aug. 1989.

[81] L. Farina and S. Rinaldi, *Positive linear systems theory and applications*. John Wiley and Sons, 2000.

[82] L. Farina, "Is a system representable as a compartmental system?" in *4rd European Control Conference, Brussels, Belgium*, 1997.

[83] H. Nijmeijer and A. J. Van der Schaft, *Nonlinear Dynamical Control Systems*. Springer Verlag, 1989.

[84] A. Astolfi, "Output feedback control of the angular velocity of a rigid body," *Systems and Control Letters*, vol. 36, no. 3, pp. 181–192, 1999.

[85] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*. AIMS Series on Applied Mathematics, 2007, vol. 2.

[86] D. Aeyels and M. Szafranski, "Comments on the stabilizability of the angular velocity of a rigid body," *Systems and Control Letters*, vol. 10, pp. 35–39, 1988.

[87] H. Rodriguez, R. Ortega, and A. Astolfi, "Adaptive partial state feedback control of the DC-to-DC Ćuk converter," in *American Control Conference, Portland, OR*, 2005, pp. 5121–5126.

Alessandro Astolfi was born in Rome, Italy, in 1967. He graduated in electrical engineering from the University of Rome in 1991. In 1992 he joined ETH-Zurich where he obtained a M.Sc. in Information Theory in 1995 and the Ph.D. degree with Medal of Honour in 1995 with a thesis on discontinuous stabilization of nonholonomic systems. In 1996 he was awarded a Ph.D. from the University of Rome "La Sapienza" for his work on nonlinear robust control. Since 1996 he is with the Electrical and Electronic Engineering Department of Imperial College, London (UK), where he is currently Professor in Non-linear Control Theory. From 1998 to 2003 he was also an Associate Professor at the Dept. of Electronics and Information of the Politecnico di Milano. Since 2005 he is also Professor at Dipartimento di Informatica, Sistemi e Produzione, University of Rome Tor Vergata. His research interests are focused on mathematical control theory and control applications, with special emphasis for the problems of discontinuous stabilization, robust stabilization, robust control and adaptive control. He is author of more than 90 journal papers, of 25 book chapters and of over 200 papers in refereed conference proceedings. He is Associate Editor of Systems and Control Letters, Automatica, the International Journal of Control, the European Journal of Control, the Journal of the Franklin Institute, and the International Journal of Adaptive Control and Signal Processing. He is the Chair of the IEEE CSS conference Editorial Board. He was awarded the IEEE Control Systems Society Antonio Ruberti Young Researcher Prize (2007) in recognition of distinguished cutting-edge contributions by a young researcher to the theory or applications of systems and control.