

Signal Generator Agnostic Moment Matching

Debraj Bhattacharjee¹, Graduate Student Member, IEEE, Alessio Moreschini², Member, IEEE, and Alessandro Astolfi³, Fellow, IEEE

Abstract—We study the model-reduction problem by moment matching for linear and nonlinear systems in a data-driven setting. We show that reduced-order models can be directly computed from input–output data without requiring knowledge of the structure of the signal generator or its internal state. The reduced-order models thus obtained match the moments of the unknown underlying system asymptotically. Our formulation provides a simple way to enforce additional constraints on the structure of the reduced-order model, which could be used to incorporate prior knowledge about the underlying system. In addition, we show that our method can be directly applied to a large class of linear and nonlinear time-delay systems with minimal modifications. Finally, we provide a simple algorithmic formulation that can be used directly with data, and demonstrate its effectiveness on a benchmark example—a nonlinear RC ladder circuit.

Index Terms—Data-driven algorithms, moment matching, nonlinear model reduction, time-delay systems.

I. INTRODUCTION

A. Overview

MATHEMATICAL models play a pivotal part in ensuring the efficient operation of large-scale dynamical systems. These models are extensively used in design, analysis, and control. In recent years, however, the increased complexity of these systems has led to difficulties in modeling the underlying process accurately, while also demanding large amounts of computational power once aggregated models are available. This has necessitated the construction of simpler models that can replicate the behavior of the underlying system under specific

operating conditions. The procedure for identifying such simpler descriptions is called *model reduction* [1], [2].

Model reduction algorithms find extensive usage in several disciplines owing to their rich mathematical structure and versatility across application domains. These include large-scale integrated circuit design, where the effects of billions of transistors need to be captured [3]; the modeling of complex mechanical systems, where the effect of flexible modes is often neglected [4]; weather forecast models, which deal with large amounts of atmospheric data [5]; and more recently in disease modeling, to better understand the spread of global epidemics [6]. For linear systems, two classes of methods have predominantly found popularity in the context of model reduction: methods based on the singular value decomposition (SVD) and methods based on Krylov subspace operators. The Hankel operator [7], [8], [9] and the celebrated balanced realization [10], [11], [12], [13] belong to the first group, whereas methods based on interpolation theory belong to the latter group [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24]. Moment matching methods, which belong to the second group and are of primary interest to us, allow the construction of reduced-order models such that the error between the underlying system and the reduced-order model is zero at some points of the complex plane. In the nonlinear setting, significant attention has been devoted to developing innovative approaches, as well as extending techniques borrowed from the linear framework. Model reduction methods for special classes of systems, such as Hamiltonian systems and bilinear systems, have been developed in [25], [26], [27], [28], and [29], whereas enhancements of the balancing framework to the nonlinear setting have been proposed in [30], [31], [32], [33], and [34]. Moment matching methods, which are known for their reduced computational complexity, have historically been associated with linear systems. However, the characterization of moments through the solution of a Sylvester equation [35], [36] allows the extension of the moment matching approach to the nonlinear setting by exploiting the steady-state response of the underlying system [18]. This has led to several developments in this area, including methods for time-delay systems, bilinear systems, and the Loewner framework [27], [28], [29], [37], [38], [39], [40], [41], [42], [43], [44], [45].

B. Motivation

Most model reduction approaches rely upon the construction of a high-dimensional mathematical model of the system to be reduced. However, with the increased complexity of interconnected systems, it is often difficult to derive such a model. Instead, due to the emergence of sophisticated sensors and

Received 12 February 2025; accepted 21 May 2025. Date of publication 2 June 2025; date of current version 27 October 2025. This work was supported in part by the European Union's Horizon 2020 Research and Innovation Programme through KIOS CoE under Grant 739551, in part by the Italian Ministry for Research in the framework of the 2020 Program for Research Projects of National Interest (PRIN) under Grant 2020RTWES4, and in part by the EPSRC Grant "Model reduction from data" under Grant EP/W005557 and Grant EP/X033546. A preliminary version of this paper was presented in part at the 2024 European Control Conference (ECC) [DOI: 10.23919/ECC64448.2024.10590933]. Recommended by Senior Editor D. Dochain. (Corresponding author: Debraj Bhattacharjee.)

Debraj Bhattacharjee and Alessio Moreschini are with the Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ London, U.K. (e-mail: d.bhattacharjee22@imperial.ac.uk; a.moreschini@imperial.ac.uk).

Alessandro Astolfi is with the Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ London, U.K., and also with the Dipartimento di Ingegneria Civile e Ingegneria Informatica, Università di Roma "Tor Vergata", 00133 Roma, Italy (e-mail: a.astolfi@imperial.ac.uk).

Digital Object Identifier 10.1109/TAC.2025.3576063

high-performance computing platforms, large-scale monitoring and collection of data from such systems are relatively easy [46]. This has led to the inception of data-driven model reduction approaches that circumvent the need for obtaining models of the underlying system. Data-driven model reduction algorithms, such as proper orthogonal decomposition and dynamic mode decomposition, which are based on the SVD, have found extensive usage in the community due to their versatility [31], [47], [48], [49]. In the moment matching framework for linear systems, data-driven versions of the Loewner framework have been presented for model reduction using frequency-domain measurements or the notion of transfer function [16], [38], [39]. Note, however, that this limits the extension of these techniques to nonlinear systems. As an alternative, a data-driven algorithm for model reduction by moment matching based on time-domain measurements has been presented in [50]. In this method, the moment of the system is first calculated based on these time-domain observables, following which reduced-order models are derived. This method has since been extended to the so-called swapped interconnection that allows two-sided moment matching [51]. A data-driven moment matching framework for nonlinear systems based on noisy output measurements has been discussed in [52] exploiting kernel-based methodologies. These methods require knowledge of the structure of the signal generator that drives the system as well as the knowledge of its internal states. The Loewner framework has also seen several developments in the area of model reduction for special classes of nonlinear systems, such as quadratic and/or bilinear systems [28], [29], [40], [41]. For more general classes of nonlinear systems, these methods typically first convert the nonlinear system under consideration into a quadratic and/or bilinear form, following which standard machinery from the Loewner framework is applied for model reduction. Such techniques typically require knowledge of the interpolation points, an augmentation of the number of states [28], and in several cases, the knowledge of the dynamics of the underlying system for the intermediate conversion.

However, in several applications and notably in the case in which the underlying system is part of a closed-loop architecture, only qualitative information may be known about the signal generator. The absence of knowledge of the signal generator (or correspondingly, the set of interpolation points) introduces difficulties in the application of existing moment-matching-based methods. In the Loewner framework, it obstructs the construction of the appropriate Loewner matrices required for model reduction, whereas in the interconnection-based moment matching framework, it does not allow the computation of the moment of the system as an intermediate step. In addition, since obtaining reduced-order models is the ultimate objective, it may be unnecessary to compute the moment of the system as an intermediate step.

C. Contributions

To deal with the data-driven model reduction problem when only qualitative information is available about the signal generator, e.g., that the signal generator has an unknown *implicit*

representation, we exploit the ideas presented in [18] and [50] to derive reduced-order models directly from input–output data. By leveraging the interconnection-based interpretation of the notion of moment, we transform the moment matching problem from a problem of interpolation of points in the complex plane to a problem of interpolation of signals in the time domain [53]. We list our contributions below.

- 1) We derive a class of reduced-order models, using input–output data, which achieve moment matching when both the underlying system and the signal generator are unknown, for both linear and nonlinear systems, thus not requiring knowledge of the interpolation points. In addition, we show that our method is readily applicable to a large class of linear time-delay and nonlinear time-delay systems.
- 2) We show that the construction of reduced-order models does not require the computation of moment of the underlying system as an intermediate step.
- 3) We highlight how our method offers greater flexibility in model structure selection of the reduced-order model when compared to existing data-driven moment matching approaches, and discuss the robustness of our method to variations in the model of the signal generator.
- 4) We provide two simple and easy-to-use algorithms that can be used to construct reduced-order models from data and highlight their effectiveness on a standard benchmark example—a nonlinear *RC* ladder.

In this article, we extend the results presented in [54] to nonlinear and time-delay systems. Furthermore, we provide simple algorithmic formulations that will help the reader apply the methods described herein.

D. Organization

The rest of this article is organized as follows. Section II is devoted to the development of the procedure for linear systems, and Section III mirrors the procedure for nonlinear systems. Inside each of these major sections, we focus on four aspects. First, we recall the preliminaries required for deriving the main results of this article. Second, we show how to obtain reduced-order models from input–output data without the knowledge of the underlying system and the signal generator. Third, we describe how this procedure can be extended to time-delay systems. Fourth, we highlight key observations and insights in the form of a discussion. To illustrate and validate the theoretical findings derived in the preceding sections, we consider a nonlinear *RC* ladder circuit as a benchmark example in Section IV. Finally, we present concluding remarks and future perspectives in Section V.

E. Notations

We use standard notation. $\mathbb{R}_{\geq 0}$ denotes the set of nonnegative real numbers, $\mathbb{R}_{> 0}$ denotes the set $\mathbb{R}_{\geq 0} \setminus \{0\}$, $\mathbb{C}_{< 0}$ denotes the set of complex numbers with negative real part, $\mathbb{C}_{\geq 0}$ denotes the set $\mathbb{C} \setminus \mathbb{C}_{< 0}$, and $\mathbb{C}_0$ denotes the set of complex numbers with zero real part. The symbol I denotes the identity matrix of appropriate dimensions, $\sigma(A)$ denotes the spectrum of the square matrix A , and $\|A\|$ denotes the induced Euclidean matrix

norm of the matrix A . The Kronecker product is denoted by $\otimes$. The vectorization operator of a matrix A is denoted by $\text{vec}(A)$ and is obtained by stacking A 's columns. Unless otherwise specified, all functions and mappings are assumed to be sufficiently smooth.

II. LINEAR SYSTEMS

A. Preliminaries on Linear Moment Matching

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

$$\begin{aligned}\dot{x} &= Ax + Bu \\ y &= Cx\end{aligned}\quad (1)$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^n$, and $C \in \mathbb{R}^{1 \times n}$. Let the associated transfer function $W(s) = C(sI - A)^{-1}B$ be minimal, i.e., the pair (A, B) is controllable and the pair (C, A) observable. We first recall the notion of moment associated with the transfer function of (1) [1, Sect. 11.1].

Definition L1: The 0-moment of the system (1) at $s_i \in \mathbb{C} \setminus \sigma(A)$ is the complex number $\eta_0(s_i) = C(s_i I - A)^{-1}B$. The k -moment of the system (1) at $s_i \in \mathbb{C} \setminus \sigma(A)$ is the complex number

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

The notion of moment for systems of the form (1) has been described in terms of a Sylvester equation in [18] and [35], as outlined below.

Lemma L1 (See [18]): Consider the system (1), let $s_i \in \mathbb{C}$ be such that $s_i \notin \sigma(A)$, for all $i = 1, \dots, \eta$. Then, there exists a one-to-one relation between the moments $\eta_0(s_1), \dots, \eta_{k_1}(s_1), \dots, \eta_0(s_\eta), \dots, \eta_{k_\eta}(s_\eta)$ and the matrix $C\Pi$, where Π is the unique solution of the Sylvester equation

$$A\Pi + BL = \Pi S \quad (2)$$

where $S \in \mathbb{R}^{\nu \times \nu}$ is any nonderogatory matrix with characteristic polynomial

$$p(s) = \prod_{i=1}^{\eta} (s - s_i)^{k_i} \quad (3)$$

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

This algebraic formulation of moments allows establishing a relationship, through the Sylvester equation, between the moments of a system and its steady-state output response (provided it exists) [18]. To present this notion of moment in terms of the steady-state response of the system, we need, however, to define the signal generator that drives the system. We consider an *implicit* signal generator given by

$$\begin{aligned}\dot{\omega} &= S\omega \\ u &= L\omega.\end{aligned}\quad (4)$$

Furthermore, we need the following assumptions to define the notion of moment for the underlying system.

Assumption L1: The system (1) is asymptotically stable, i.e., $\sigma(A) \subset \mathbb{C}_{<0}$.

Assumption L2: The signal generator defined by (4) is such that $\sigma(S) \subset \mathbb{C}_0$ and the matrix S has simple eigenvalues. Furthermore, the triple $(L, S, \omega(0))$ is minimal.

We are now ready to recall the result that connects the moments of a system to its steady-state output response when it is driven by a signal generator of the form (4).

Theorem L1 (See [18]): Consider the system (1) and the signal generator (4). Suppose that Assumptions L1 and L2 hold. Let $S \in \mathbb{R}^{\nu \times \nu}$ be any nonderogatory matrix with characteristic polynomial as defined in (3). Then, there exists a one-to-one relation between the moments $\eta_0(s_1), \dots, \eta_{k_1}(s_1), \dots, \eta_0(s_\eta), \dots, \eta_{k_\eta}(s_\eta)$ and the steady-state output of the interconnected system (1) and (4).

Exploiting Lemma L1 and Theorem L1, the notion of reduced-order model can be formalized as follows.

Definition L2: The system described by the equations

$$\begin{aligned}\dot{\xi} &= F\xi + Gu \\ \psi &= H\xi\end{aligned}\quad (5)$$

with $\xi(t) \in \mathbb{R}^\rho$, $u(t) \in \mathbb{R}$, $\psi(t) \in \mathbb{R}$, $F \in \mathbb{R}^{\rho \times \rho}$, $G \in \mathbb{R}^\rho$, and $H \in \mathbb{R}^{1 \times \rho}$, is a model of the system (1) at (S, L) if the system (5) has the same moment at (S, L) as the system (1). In this case, the system (5) is said to *match* the moment of the system (1). Furthermore, the system (5) is said to be a reduced-order model of the system (1) if $\rho < n$.

The condition under which this reduced-order model achieves moment matching with the underlying system in the linear setting is described next.

Lemma L2 (See [18]): Consider the system (1), the system (5), and the signal generator (4). The system (5) is a model of the system (1) if $\sigma(S) \cap \sigma(A) = \emptyset$ and $\sigma(S) \cap \sigma(F) = \emptyset$, such that

$$C\Pi = HP \quad (6)$$

where Π is the unique solution of the Sylvester equation (2) and P is the unique solution of the Sylvester equation

$$FP + GL = PS. \quad (7)$$

While deriving reduced-order models, one usually starts with a high-dimensional model of the underlying system. This, of course, assumes knowledge of the full-order realization. In practice, however, it may be difficult to obtain state-space realizations of complex systems. Instead, it is easier to obtain input–output data from such systems and then use these data to obtain reduced-order models. It was shown in [50] that reduced-order models achieving moment matching can be constructed from input–output data. However, this procedure still requires full information about the structure of the signal generator via the use of S and L . In a completely data-driven setting, one may only have access to input–output data and some qualitative knowledge about the signal generator itself. In such situations, the aforementioned procedure cannot be utilized for model reduction. To address these issues, we conclude this section with the following assumptions and properties.

Assumption L3: The elements of the window of sample times T_k^w are given by

$$T_k^w := \{t_{k-w+1}, \dots, t_{k-1}, t_k\}$$

where $0 \leq t_0 < \dots < t_{k-w} < \dots < t_k$ and $k \geq w - 1 \geq 0$, are selected such that $\text{rank}([\omega(t_{k-w+1}) \dots \omega(t_k)]) = \nu$, for all $k \geq w - 1$.

Assumptions L2 and L3 guarantee that the signals generated by (4) are persistently exciting over the time window T_k^w [55], [56], [57]. Similar to Assumption L1, we impose a stability condition on the reduced-order model (5).

Assumption L4: The system (5) is asymptotically stable, i.e., $\sigma(F) \subset \mathbb{C}_{<0}$, and the pair (F, G) is controllable.

B. Direct Computation of Linear Reduced-Order Models From Data

In this section, we briefly recall a recently introduced enhancement [54] of Scarcitti and Astolfi [50]. The results of Bhat-tacharjee and Astolfi [54] allow for the construction of reduced-order models from input–output data when the system (1) is driven by an *unknown* implicit signal generator described by equations of the form (4), with u and y as the only available measurements. More precisely, the reduced-order models obtained through this technique do not require knowledge of the structure of the signal generator or of its internal state. We note again that this formulation is based on the premise that if the structure of the signal generator is unknown, the Sylvester equation (7) cannot be solved, and thus, reduced-order models cannot be constructed using the standard procedure. We emphasize that we need to recall some of the results presented in [54] in order to naturally extend the same approach to time-delay and nonlinear systems, as shown in later sections. We first recall the construction of reduced-order models from a batch of input–output data.

Theorem L2 (See [54]): Consider the system (1), the system (5), and the signal generator (4). Suppose that Assumptions L1–L4 hold. Let the time snapshots $\tilde{R}_k \in \mathbb{R}^{w \times \rho}$ and $\tilde{\Gamma}_k \in \mathbb{R}^w$, with $w \geq \rho > 0$, be defined as

$$\tilde{R}_k = \begin{bmatrix} (\xi(t_{k-w+1}))^\top \\ \vdots \\ (\xi(t_{k-1}))^\top \\ (\xi(t_k))^\top \end{bmatrix}, \quad \tilde{\Gamma}_k = \begin{bmatrix} y(t_{k-w+1}) \\ \vdots \\ y(t_{k-1}) \\ y(t_k) \end{bmatrix}$$

respectively, such that $k \geq w - 1$. Assume that the matrix $\tilde{R}_k$ has full column rank. Then

$$\text{vec}(\tilde{H}_k) = (\tilde{R}_k^\top \tilde{R}_k)^{-1} \tilde{R}_k^\top \tilde{\Gamma}_k. \quad (8)$$

The relation between the matrix $\tilde{H}$ in Theorem L2 and the output matrix of the reduced-order model (5) can be seen in the result below.

Theorem L3 (See [54]): Consider the system (1), the system (5), and the signal generator (4). Let H be the solution of (6). Suppose that Assumptions L1–L4 hold. Then, there exist sequences $\{t_k\}$ such that $\lim_{k \rightarrow \infty} \tilde{H}_k = H$.

The computation of $\text{vec}(\tilde{H}_k)$ at each time instant (8) involves the inversion of a matrix of dimension $\rho \times \rho$. While ρ and ν are unrelated quantities, it is natural to consider that in practice, the

dimension of the reduced-order model exceeds the number of interpolation points, $\rho \geq \nu$. If the number of interpolation points (and correspondingly, the size of the reduced-order model) is large, this procedure may not be computationally efficient. Therefore, we now recall a recursive version of (8) that alleviates this issue.

Theorem L4 (See [54]): Consider the system (1), the system (5), and the signal generator (4). Let $\alpha_k = P\omega(t_k)$ and assume that $\Phi_k = (\tilde{R}_k^\top \tilde{R}_k)^{-1}$ and $\Upsilon_k = ((\tilde{R}_{k-1}^\top \tilde{R}_{k-1})^{-1} + \alpha_k \alpha_k^\top)^{-1}$ are full rank for all $t \geq t_r$, where $t_r > t_w$. Given $\text{vec}(\tilde{H}_r)$, Φ_r , α_r , and Υ_r , the recursive least-square formulation for $\text{vec}(\tilde{H}_k)$ for all $t \geq t_r$ and $k \geq w - 1$ is given by

$$\begin{aligned} \text{vec}(\tilde{H}_k) &= \text{vec}(\tilde{H}_{k-1}) + \Phi_k \alpha_k (y(t_k) - \alpha_k^\top \text{vec}(\tilde{H}_{k-1})) \\ &\quad - \Phi_k \alpha_{k-w} (y(t_{k-w}) - \alpha_{k-w}^\top \text{vec}(\tilde{H}_{k-1})) \end{aligned} \quad (9)$$

where

$$\Phi_k = \Upsilon_k - \Upsilon_k \alpha_{k-w} (\alpha_{k-w}^\top \Upsilon_k \alpha_{k-w} - I)^{-1} \alpha_{k-w}^\top \Upsilon_k \quad (10)$$

and

$$\Upsilon_k = \Phi_{k-1} - \Phi_{k-1} \alpha_k (I + \alpha_k^\top \Phi_{k-1} \alpha_k)^{-1} \alpha_k^\top \Phi_{k-1}. \quad (11)$$

The recursive formulation is appealing due to the absence of any matrix inversions since only scalar divisions are involved. Thus, any increase in the size of the reduced-order model scales the computational cost linearly. This estimation procedure is run until $\|\tilde{H}_k - \tilde{H}_{k-1}\| < \delta$, where $\delta > 0$ is the convergence tolerance chosen by the user.

As described in this section, the objective of the designer is to construct a reduced-order model from input–output data without the knowledge of the underlying system or the signal generator. As such, the designer is tasked with the construction of the state dynamics of the reduced-order model, and the proposed technique computes the output mapping for the same reduced-order model such that it achieves moment matching with the underlying system. We provide a brief discussion on the selection of the state dynamics of the reduced-order model in Section II-D. We conclude this section by providing an algorithmic formulation of Theorem L4 in Algorithm 1.

C. Matching Linear Systems With Delays

In this section, we highlight how the previously introduced results from [54] can be extended to a large class of linear time-delay systems. In particular, we focus on linear systems where the delay affects only the state dynamics of the system but not its output. To this end, consider the linear single-input, single-output, continuous-time, time-delay system described by the set of equations

$$\begin{aligned} \dot{x} &= \sum_{j=0}^{\mu} A_j x_{\tau_j} + B u_{\tau_u} \\ y &= C x \\ x(\theta) &= \phi(\theta), \quad -T \leq \theta \leq 0 \end{aligned} \quad (12)$$

where $x(\theta)$ denotes the initial condition described by the continuous function $\phi : [-T, 0] \rightarrow \mathbb{R}^n$, and $x_{\tau_j} \in \mathbb{R}^n$ denotes the translation operator applied to $x(t) \in \mathbb{R}^n$, i.e., $x_{\tau_j} = x(t - \tau_j)$,

Algorithm 1: Model Reduction for Linear Systems.**Input:** F, G **Output:** $\tilde{H}$ **Data:** u, y **Initialization:** $\tilde{H}_0, \Phi_0, w, \delta$ $k \leftarrow 1;$ **do**

```

    // Compute  $\xi$  using  $F, G$ , and  $u$ 
     $\alpha_k \leftarrow \xi(t_k);$ 
     $\alpha_{k-w} \leftarrow \xi(t_{k-w});$ 
    // Solve for  $\Upsilon_k$  using equation (11)
     $\Upsilon_k \leftarrow \text{compute}\Upsilon(\Phi_{k-1}, \alpha_k);$ 
    // Solve for  $\Phi_k$  using equation (10)
     $\Phi_k \leftarrow \text{compute}\Phi(\Upsilon_k, \alpha_{k-w});$ 
    // Solve for  $\tilde{H}_k$  using equation (9)
     $\tilde{H}_k \leftarrow \text{compute}H(\tilde{H}_{k-1}, \Phi_k, y, \alpha_k, \alpha_{k-w});$ 
     $k \leftarrow k + 1;$ 

```

while $\|\tilde{H}_k - \tilde{H}_{k-1}\| > \delta;$

for $\tau_j \in \mathbb{R}_{\geq 0}$, with $\tau_0 = 0$ and $j = 1, \dots, \mu$. Moreover, $u_{\tau_u} = u(t - \tau_u) \in \mathbb{R}$ for $\tau_u \in \mathbb{R}_{\geq 0}$, $y(t) \in \mathbb{R}$, $B \in \mathbb{R}^n$, $C \in \mathbb{R}^{1 \times n}$, and $A_j \in \mathbb{R}^{n \times n}$, with $j = 0, \dots, \mu$. Let

$$\bar{A}(s) = A_0 + \sum_{j=1}^{\mu} e^{-s\tau_j} A_j.$$

Then, if $s_i \notin \sigma(\bar{A}(s))$, for all $s_i \in \sigma(S)$, the moments of the system (12), namely

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

are in one-to-one relation with the matrix $C\Pi$ [37], where $G_d(s)$ is given by

$$G_d(s) = C \left(sI - A_0 - \sum_{j=1}^{\mu} e^{-s\tau_j} A_j \right)^{-1} e^{-s\tau_u} B$$

and Π is the unique solution of the Sylvester-like equation

$$A_0\Pi + \sum_{j=1}^{\mu} A_j\Pi e^{-s\tau_j} - \Pi S = -BLe^{-s\tau_u}. \quad (13)$$

The time response of this linear time-delay system can be described as in the finite-dimensional linear setting, as shown next.

Lemma LTD1 (See [37]): Consider the interconnection of the system (12) and the signal generator (4). Suppose that Assumptions L2 and L3 hold and $\sigma(\bar{A}(s)) \subset \mathbb{C}_{<0}$. Then, the output, y , of the system (12) is described by

$$y(t) = C\Pi\omega(t) + \epsilon_1(t)$$

with $\epsilon_1(t) = C\mathcal{L}^{-1}\{(sI - \bar{A}(s))^{-1}(x(0) - \Pi\omega(0))\}$ being an exponentially decaying function.

We now recall the notion of reduced-order model for linear time-delay systems.

Definition LTD1: Consider the interconnection of the system (12) and the signal generator (4), and suppose that Assumptions L2 and L3 hold. Furthermore, assume that the system (12) is minimal and $\sigma(\bar{A}(s)) \subset \mathbb{C}_{<0}$. Then, the system

$$\dot{\xi} = \sum_{j=0}^{\vartheta} F_j \xi_{\gamma_j} + Gu_{\gamma_u}$$

$$\psi = H\xi$$

$$\xi(\theta) = \chi(\theta), \quad -T \leq \theta \leq 0 \quad (14)$$

with $\xi(t) \in \mathbb{R}^{\rho}$ is a *model* at (S, L) of the system (12) if the system (14) has the same moment at (S, L) as the system (12). Here, $F_j \in \mathbb{R}^{\rho \times \rho}$, for $j = 0, \dots, \vartheta$, $G \in \mathbb{R}^{\rho}$, $H \in \mathbb{R}^{1 \times \rho}$, $\xi(\theta)$ denotes the initial condition described by the continuous function $\chi: [-T, 0] \rightarrow \mathbb{R}^{\rho}$, and delays $\gamma_u \geq 0$ and $\gamma_j \geq 0$ for $j = 0, \dots, \vartheta$. In this case, the system (14) is said to *match* the moment of the system (12) at (S, L) . Furthermore, the system (14) is a reduced-order model of the system (12) if $\rho < n$.

The condition under which this reduced-order model achieves moment matching with the underlying system (12) is given next.

Lemma LTD2 (See [37]): Consider the system (12), the system (14), and the signal generator (4). Let $\bar{F}(s) = F_0 + \sum_{j=1}^{\vartheta} e^{-s\gamma_j} F_j$. Assume that $\sigma(S) \cap \sigma(\bar{A}(s)) = \emptyset$ and $\sigma(S) \cap \sigma(\bar{F}(s)) = \emptyset$. The system (14) is a model of the system (12) if

$$C\Pi = HP$$

where Π is the unique solution of the Sylvester-like equation (13) and P is the unique solution of the Sylvester-like equation

$$F_0 P + \sum_{j=1}^{\rho} F_j P e^{-s\gamma_j} - \Pi S = -GL e^{-s\gamma_u}.$$

Analogous to the linear finite-dimensional setting, the following property holds.

Lemma LTD3 (See [37]): Consider the interconnection of the system (14) and the signal generator (4). Suppose that Assumptions L2 and L3 hold, and that $\sigma(\bar{F}(s)) \subset \mathbb{C}_{<0}$. Then, the output, y , of the system (14) is described by

$$\psi(t) = HP\omega(t) + \epsilon_2(t)$$

with $\epsilon_2(t) = H\mathcal{L}^{-1}\{(sI - \bar{F}(s))^{-1}(\xi(0) - P\omega(0))\}$ being an exponentially decaying signal.

We now present a result that outlines how reduced-order models can be computed from input–output data for linear-time delay systems.

Theorem LTD1: Consider the system (12), the system (14), and the signal generator (4). Suppose that Assumptions L2 and L3 hold. Furthermore, assume that $\sigma(\bar{A}(s)) \subset \mathbb{C}_{<0}$ and $\sigma(\bar{F}(s)) \subset \mathbb{C}_{<0}$. Let the time snapshots $R_k \in \mathbb{R}^{w \times \rho}$ and $\Gamma_k \in \mathbb{R}^w$, with $w \geq \rho > 0$, be defined as

$$R_k = \begin{bmatrix} (\xi(t_{k-w+1}))^\top \\ \vdots \\ (\xi(t_{k-1}))^\top \\ (\xi(t_k))^\top \end{bmatrix}$$

and

$$\Gamma_k = \begin{bmatrix} y(t_{k-w+1}) - (\epsilon_1(t_{k-w+1}) - \epsilon_2(t_{k-w+1})) \\ \vdots \\ y(t_{k-1}) - (\epsilon_1(t_{k-1}) - \epsilon_2(t_{k-1})) \\ y(t_k) - (\epsilon_1(t_k) - \epsilon_2(t_k)) \end{bmatrix}$$

respectively, such that $k \geq w - 1$. Assume that the matrix R_k has full column rank. Then

$$\text{vec}(H_k) = (R_k^\top R_k)^{-1} R_k^\top \Gamma_k. \quad (15)$$

Proof: From Lemma LTD2, for all $\omega(t) \neq 0$, we have that

$$C\Pi\omega(t) = H P\omega(t).$$

From Lemmas LTD1 and LTD3, we have that

$$\psi(t) = y(t) - (\epsilon_1(t) - \epsilon_2(t))$$

which leads to

$$H\xi(t) = y(t) - (\epsilon_1(t) - \epsilon_2(t)).$$

Using the Kronecker product and the vectorization operator, we obtain

$$\text{vec}(H\xi(t)) = \text{vec}(y(t) - (\epsilon_1(t) - \epsilon_2(t)))$$

which leads to

$$(\xi(t))^\top \text{vec}(H) = \text{vec}(y(t) - (\epsilon_1(t) - \epsilon_2(t))).$$

□

Similar to Section II-B, we solve (15) over a set of sample times given by T_k^w . Since $\sigma(\bar{A}(s)) \subset \mathbb{C}_{<0}$ and $\sigma(\bar{F}(s)) \subset \mathbb{C}_{<0}$, ϵ_1 and ϵ_2 exponentially decay to zero. This allows giving an approximate version of the results that does not depend on the initial conditions of the underlying system.

To this end, suppose that Assumptions L2 and L3 hold. Furthermore, assume that $\sigma(\bar{A}(s)) \subset \mathbb{C}_{<0}$ and $\sigma(\bar{F}(s)) \subset \mathbb{C}_{<0}$. Let the time snapshots $R_k \in \mathbb{R}^{w \times \rho}$ and $\Gamma_k \in \mathbb{R}^w$, with $w \geq \rho > 0$, be defined as

$$R_k = \begin{bmatrix} (\xi(t_{k-w+1}))^\top \\ \vdots \\ (\xi(t_{k-1}))^\top \\ (\xi(t_k))^\top \end{bmatrix}, \quad \Gamma_k = \begin{bmatrix} y(t_{k-w+1}) \\ \vdots \\ y(t_{k-1}) \\ y(t_k) \end{bmatrix}$$

respectively, such that $k \geq w - 1$. Assume that the matrix R_k has full column rank. Then

$$\text{vec}(H_k) \approx (R_k^\top R_k)^{-1} R_k^\top \Gamma_k. \quad (16)$$

Algorithm 1 can now be readily used to obtain reduced-order models from input-output data for linear time-delay systems.

D. Discussion

When compared to existing results in the literature, our results are advantageous on several fronts.

- 1) First, using existing methods, one cannot construct reduced-order models if the signal generator or the set of interpolation points is not known, while, in our work, we do not need to know the structure of the signal generator or its internal state [16], [38], [39], [50].

- 2) Second, in our method, there is no need to compute the moment of the full-order system, $C\Pi$, as an intermediate step while obtaining reduced-order models, as is seen in Algorithm 1.
- 3) Third, in our formulation, both F and G in the reduced-order model (5) [system (14) in the time-delay setting] are free parameters that can be chosen freely to preserve properties of the underlying system, such as matching with prescribed eigenvalues or matching with prescribed relative degree.
- 4) Fourth, in the linear setting, any choice of the matrices F and G satisfying Assumption L4 yields a reduced-order model that achieves moment matching [54]. This is advantageous in a setting where the designer does not have any knowledge about the underlying system. However, this would also mean that for a choice of F and G that is not representative of the state dynamics of the underlying system, the designer would need to wait longer to collect steady-state data.
- 5) Finally, reduced-order models obtained from this procedure are robust to variations in the signal generator (4). To see this, note that in the standard case in which the signal generator (4) is known [50], reduced-order models are dependent on the estimate of $C\Pi$. Therefore, if the signal generator changes (4), $C\Pi$ needs to be recomputed, leading to a recomputation of the reduced-order model. However, since we do not need to compute $C\Pi$, any variation in the signal generator is automatically captured by the output matrix H of the reduced-order model.

The time taken by $\tilde{H}$ to converge in Algorithm 1 depends on two primary factors—the initial estimate $\tilde{H}_0$ and the availability of steady-state data. A poor initial estimate and a longer waiting period for steady-state input-output data would lead to a situation in which the user would have to wait longer for $\tilde{H}$ to converge. On the other hand, the batch version of Algorithm 1 as presented in Theorem L2 does not depend on the initial estimate $\tilde{H}_0$. Therefore, if steady-state data were to be available, the batch algorithm would yield the correct estimate in one step. This concludes the results and discussion for linear systems. We now shift our attention to nonlinear systems.

III. NONLINEAR SYSTEMS

A. Preliminaries on Nonlinear Moment Matching

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

$$\begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x) \end{aligned} \quad (17)$$

with state $x(t) \in \mathbb{R}^n$, input $u(t) \in \mathbb{R}$, output $y(t) \in \mathbb{R}$, and (sufficiently) smooth mappings $f(\cdot, \cdot)$ and $h(\cdot)$ such that $f(0, 0) = 0$ and $h(0) = 0$. We assume also that the system (17) is minimal, i.e., (locally) observable¹ and (locally) accessible.² Consider also a nonlinear signal generator described by equations of the

¹ See [58, Def. 2.10] for the notion of observability.

² See [58, Def. 2.11] for the notion of accessibility.

form

$$\begin{aligned}\dot{\omega} &= s(\omega) \\ u &= l(\omega)\end{aligned}\quad (18)$$

with state $\omega(t) \in \mathbb{R}^s$, output $u(t) \in \mathbb{R}$, and (sufficiently) smooth mappings $s(\cdot)$ and $l(\cdot)$ such that $s(0) = 0$ and $l(0) = 0$. Then, the system (17) and the signal generator (18) form an interconnected system given by

$$\begin{aligned}\dot{\omega} &= s(\omega) \\ \dot{x} &= f(x, l(\omega)) \\ y &= h(x).\end{aligned}\quad (19)$$

Assumption NL1: The system (17) is such that the zero equilibrium of $\dot{x} = f(x, 0)$ is locally exponentially stable.

Assumption NL2: The signal generator (18) is (locally) observable and $\omega = 0$ is a Poisson stable³ equilibrium.

Under Assumptions NL1 and NL2, there exists a mapping $\pi(\cdot)$, locally defined in a neighborhood of $\omega = 0$, with $\pi(0) = 0$, which satisfies the partial differential equation

$$f(\pi(\omega), l(\omega)) = \frac{\partial \pi}{\partial \omega} s(\omega) \quad (20)$$

see [59, Prop. 8.1.1]. This implies that the interconnected system (19) possesses an invariant manifold, described by the equation $x = \pi(\omega)$. These assumptions allow defining the notion of moment of the nonlinear system under consideration.

Definition NLI: Consider the system (17) and the signal generator (18). The moment of the system (17) at (s, ℓ) is given by the mapping $h(\pi(\cdot))$, where $\pi(\cdot)$ is the unique solution of (20).

We now recall the connection between the notion of moment of the interconnected system (19) with its steady-state output response.

Theorem NLI (See [18]): Consider the system (17) and the signal generator (18). Suppose that Assumptions NL1 and NL2 hold. Then, the moment of the system (17) at (s, ℓ) coincides with the (locally well-defined) steady-state response of the output of the interconnected system (19).

We now introduce the notion of reduced-order model in the nonlinear setting; see [18].

Definition NL2: The system described by the equations

$$\begin{aligned}\dot{\xi} &= \phi(\xi, u) \\ \psi &= \kappa(\xi)\end{aligned}\quad (21)$$

with $\xi(t) \in \mathbb{R}^\rho$ is a *model* at (s, ℓ) of the system (17) if the system (21) has the same moment at (s, ℓ) as the system (17). In this case, the system (21) is said to *match* the moment of the system (17) at (s, ℓ) . Furthermore, the system (21) is a reduced-order model of the system (17) if $\rho < n$.

The condition under which this reduced-order model achieves moment matching with the underlying system in the nonlinear setting is described next.

Lemma NLI (See [18]): Consider the system (17), the system (21), and the signal generator (18). Suppose that Assumptions

NL1 and NL2 hold. Then, the system (21) *matches* the moment of the system (17) at (s, ℓ) if the equation

$$\phi(p(\omega), l(\omega)) = \frac{\partial p}{\partial \omega} s(\omega) \quad (22)$$

has a unique solution $p(\cdot)$ with boundary condition $p(0) = 0$ such that

$$h(\pi(\cdot)) = \kappa(p(\cdot)) \quad (23)$$

where $\pi(\cdot)$ is the unique solution of (20).

Similar to the linear setting, the construction of this class of reduced-order models requires the knowledge of the structure of the signal generator. Therefore, we focus our attention on a formulation that circumvents this requirement.

B. Direct Computation of Nonlinear Reduced-Order Models From Data

As described in the previous section, the construction of a reduced-order model that achieves moment matching with the underlying system requires the computation of π . Even when there is perfect knowledge of the dynamics of the system (17), solving for π in (20) may be difficult. In the case in which the mapping f is unknown, π could be obtained numerically but would require knowledge of the states of the system (17) [50]. In a data-driven setting, however, one may only have access to input–output data. Therefore, to consider the moment-matching problem in a purely data-driven setting, we introduce the following hypotheses on the reduced-order model, analogous to the linear case.

Assumption NL3: The system (21) is minimal. The zero equilibrium of the system $\dot{\xi} = \phi(\xi, 0)$ is locally exponentially stable. Furthermore, the system is (locally) accessible.

Combining Assumption NL3 with Assumption NL2 implies that there exists a unique mapping $p(\cdot)$, locally defined in a neighborhood of $\omega = 0$ and $p(0) = 0$, which solves the partial differential equation

$$\phi(p(\omega), l(\omega)) = \frac{\partial p}{\partial \omega} s(\omega). \quad (24)$$

We now state a result that allows the expression of the moment of the underlying system and the reduced-order model in terms of their corresponding output responses.

Lemma NL2 (See [50]): Consider the system (17) and the signal generator (18). Suppose that Assumptions NL1 and NL2 hold. Then, the output response of the system (17) is given by the equation

$$y(t) = h(\pi(\omega(t))) + \epsilon(t) \quad (25)$$

where $\epsilon(t)$ is an exponentially decaying signal. Similarly, consider the system (21) and the signal generator (18). Suppose that Assumptions NL2 and NL3 hold. Then, the output response of the system (21) is given by the equation

$$\psi(t) = \kappa(\xi(t)) + \varepsilon(t) \quad (26)$$

where $\varepsilon(t)$ is an exponentially decaying signal.

With this in mind, we are now ready to state a result that allows for the definition of the moment of the underlying system in terms of the state of the reduced-order model.

³See [59, Sect. 8.1] for the notion of Poisson stability.

Lemma NL3: Consider the system (17), the system (21), and the signal generator (18). Suppose that Assumptions NL1–NL3 hold. Then, the system (21) matches the moment of the system (17) at (s, l) and

$$\lim_{t \rightarrow \infty} (h(\pi(\omega(t))) - \kappa(\xi(t))) = 0. \quad (27)$$

Proof: Assumptions NL2 and NL3 guarantee that (24) has a solution $p(\cdot)$ and that the interconnection of the system (21) and the signal generator (18) has an invariant manifold described by the equation

$$p(\omega) = \xi. \quad (28)$$

By Lemma NL1, system (21) matches the moments of the system (17) at (s, l) . Finally, using (28) in the moment matching condition (23), we conclude (27). $\square$

Using the result in Lemma NL3, (25) can be transformed into a form in which the output of the underlying system is related to the state of the reduced order model, as shown next. From an identification viewpoint, the mapping $\kappa(\cdot)$ must be identified. However, an exact identification may not be feasible and an approximation should be considered. Therefore, we consider a standard assumption for approximating a mapping by a family of basis functions [60]. Typically, the family of basis functions representing the mapping can be implemented by a trial-and-error procedure by using a polynomial expansion or an expansion based on the class of input signals. It is worthwhile noting here that there are results of “universal” approximation for some families of basis functions [61], [62]. In light of this, we now introduce an assumption on the mapping $\kappa(p(\cdot))$.

Assumption NL4: The mapping $\kappa(p(\cdot))$ belongs to the function space described by the family of continuous basis functions $\varphi: \mathbb{R}^p \rightarrow \mathbb{R}$, with $i = 1, \dots, M$ (M could be ∞), i.e., there exist constants $\Gamma_i \in \mathbb{R}$ with $i = 1, \dots, M$ such that $\kappa(\xi) = \sum_{i=1}^M \Gamma_i \varphi_i(\xi)$. If $M = \infty$, then we assume that the series is uniformly convergent.

While it may be true that a specific family of basis functions may be able to exactly represent a certain mapping $\kappa(p(\cdot))$ with a finite basis, this is not the case generally. Therefore, consider the following approximation

$$\bar{\Gamma} = \begin{bmatrix} \bar{\Gamma}_1 & \bar{\Gamma}_2 & \dots & \bar{\Gamma}_N \end{bmatrix}$$

$$\Omega(\xi) = \begin{bmatrix} \varphi_1(\xi) & \varphi_2(\xi) & \dots & \varphi_N(\xi) \end{bmatrix}$$

such that $N \leq M$. Using a weighted sum of these basis functions, (26) can be rewritten as

$$y(t) = \sum_{i=1}^N \bar{\Gamma}_i \varphi_i(\xi(t)) + e(t) + \epsilon(t)$$

$$= \bar{\Gamma} \Omega(\xi(t))^\top + e(t) + \epsilon(t)$$

where $e(t) = \sum_{i=N+1}^M \bar{\Gamma}_i \varphi_i(\xi(t))$ is the error due to the termination of the summation at N . Finally, since ϵ is an exponentially decaying signal, consider the approximation

$$\tilde{y} = \sum_{i=1}^N \bar{\Gamma}_i \varphi_i(\xi) = \bar{\Gamma} \Omega(\xi)^\top \quad (29)$$

which neglects the truncation error e and the transient error ϵ . Analogous to Section II-B, let $\bar{\Gamma}_k$ be the estimate of the matrix $\bar{\Gamma}$ at T_k^w . Since ϵ is not known, $\bar{\Gamma}_k$ [described by (29)] needs to be approximated. This approximation can be determined from data as follows.

Theorem NL2: Consider the system (17), the system (21), and the signal generator (18). Suppose that Assumptions NL1–NL3 hold. Let the time snapshots $U_k \in \mathbb{R}^{w \times N}$ and $\Xi_k \in \mathbb{R}^w$, with $w \geq N > 0$, be defined as

$$\tilde{U}_k = \begin{bmatrix} \Omega(\xi(t_{k-w+1})) \\ \vdots \\ \Omega(\xi(t_{k-1})) \\ \Omega(\xi(t_k)) \end{bmatrix}, \quad \tilde{\Xi}_k = \begin{bmatrix} y(t_{k-w+1}) - \tilde{e}(t_{k-w+1}) \\ \vdots \\ y(t_{k-1}) - \tilde{e}(t_{k-1}) \\ y(t_k) - \tilde{e}(t_k) \end{bmatrix}$$

respectively, where $\tilde{e}(t_k) = \epsilon(t_k) + e(t_k)$ and $k \geq w - 1$. If $\tilde{U}_k$ is full column rank, then

$$\text{vec}(\hat{\Gamma}_k) = (\tilde{U}_k^\top \tilde{U}_k)^{-1} \tilde{U}_k^\top \tilde{\Xi}_k \quad (30)$$

is an approximation of the estimate $\bar{\Gamma}_k$.

Proof: Using (29), we can express the output response of the underlying system as

$$\bar{\Gamma} \Omega(\xi(t))^\top = y(t) - (\epsilon(t) + e(t)).$$

By applying the vectorization operator on both sides of the equation above, we obtain

$$\Omega(\xi(t)) \text{vec}(\bar{\Gamma}) = \text{vec}(y(t) - \epsilon(t) - e(t)). \quad (31)$$

By computing (31) at all elements of T_k^w , we obtain

$$\tilde{U}_k \text{vec}(\bar{\Gamma}_k) = \tilde{\Xi}_k. \quad (32)$$

Since the matrix $\tilde{U}_k$ has full column rank by assumption, we can compute $\bar{\Gamma}_k$ from (32), thus yielding (30). $\square$

For the approximation $\bar{\Gamma}_k$, defined in (30) to be well defined for all k , the set of sample times T_k^w needs to be selected in a way that the matrix $\tilde{U}_k^\top \tilde{U}_k$ is full column rank. This condition on T_k^w highlights a property of persistence of excitation of the signal generator (18), which is guaranteed by the following assumption; see [57].

Assumption NL5: The initial condition $\omega(0)$ of the signal generator (18) is almost periodic, and all the solutions of the system are analytic. Furthermore, the signal generator (18) satisfies the excitation rank condition [57] at $\omega(0)$.

We now define the approximation $\bar{\Gamma}_k$ in terms of the mapping $\kappa(p(\cdot))$ to simplify notation, following which we will be ready to prove the main result of this section.

Definition NL3: Consider the system (21) and the signal generator (18). The estimated moment of the system (21) is defined as

$$\tilde{\kappa}_{N,k}(\cdot) = \bar{\Gamma}_k \Omega(\cdot)^\top$$

where $\bar{\Gamma}_k$ is computed from (30).

Theorem NL3: Consider the system (17), the system (21), and the signal generator (18). Suppose that Assumptions NL1–NL5

Algorithm 2: Model Reduction for Nonlinear Systems.

Input: ϕ, Ω
Output: $\bar{\Gamma}$
 // Compute ξ using ϕ and u
Data: y, ξ
Initialization: $\bar{\Gamma}_0, \Phi_0, w, \delta$
 $k \leftarrow 1$;
do
 // Compute α using Ω
 $\alpha_k \leftarrow (\Omega(\xi(t_k)))^\top$;
 $\alpha_{k-w} \leftarrow (\Omega(\xi(t_{k-w})))^\top$;
 // Solve for Υ_k using equation (36)
 $\Upsilon_k \leftarrow \text{compute}\Upsilon(\Phi_{k-1}, \alpha_k)$;
 // Solve for Φ_k using equation (35)
 $\Phi_k \leftarrow \text{compute}\Phi(\Upsilon_k, \alpha_{k-w})$;
 // Solve for $\bar{\Gamma}_k$ using equation (34)
 $\bar{\Gamma}_k \leftarrow \text{compute}\bar{\Gamma}(\bar{\Gamma}_{k-1}, \Phi_k, y, \alpha_k, \alpha_{k-w})$;
 $k \leftarrow k + 1$;
while $\|\bar{\Gamma}_k - \bar{\Gamma}_{k-1}\| > \delta$;

hold. Then, there exist sequences $\{t_k\}$ such that

$$\lim_{k \rightarrow \infty} \left(y(t_k) - \lim_{N \rightarrow M} \tilde{\kappa}_{N,k}(\xi(t_k)) \right) = 0. \quad (33)$$

Proof: Assumptions NL1 and NL2 guarantee that $h(\pi(\cdot))$ is well defined and $h(\pi(\omega))$ matches the steady-state output y of the interconnected system (19). Similarly, Assumptions NL2 and NL3 guarantee that $\kappa(p(\cdot))$ is well defined. Furthermore, Assumptions NL1 and NL3 imply that there exist sequences $\{t_k\}$, with $\lim_{k \rightarrow \infty} t_k = \infty$, such that $\|\epsilon(t_k)\|$ vanishes exponentially. Finally, Assumption NL5 guarantees that the approximation $\bar{\Gamma}_k$ is well defined for all k . Therefore, by Assumption NL4, (33) holds. $\square$

For computational efficiency and robustness to variations in the signal generator, we now present an online version of (30) as we did in the linear setting with an accompanying algorithmic formulation of Theorem NL2.

Theorem NL4: Consider the system (17), the system (21), and the signal generator (18). Let $\alpha_k = (\Omega(\xi(t_k)))^\top$ and assume that $\Phi_k = (\tilde{U}_k^\top \tilde{U}_k)^{-1}$ and $\Upsilon_k = ((\tilde{U}_{k-1}^\top \tilde{U}_{k-1})^{-1} + \alpha_k \alpha_k^\top)^{-1}$ are full rank for all $t \geq t_r$, where $t_r > t_w$. Given $\text{vec}(\bar{\Gamma}_r)$, Φ_r , α_r , and Υ_r , the recursive least-square formulation for $\text{vec}(\bar{\Gamma}_k)$ for all $t \geq t_r$ and $k \geq w - 1$ is given by

$$\begin{aligned} \text{vec}(\bar{\Gamma}_k) &= \text{vec}(\bar{\Gamma}_{k-1}) + \Phi_k \alpha_k (y(t_k) - \alpha_k^\top \text{vec}(\bar{\Gamma}_{k-1})) \\ &\quad - \Phi_k \alpha_{k-w} (y(t_{k-w}) - \alpha_{k-w}^\top \text{vec}(\bar{\Gamma}_{k-1})) \end{aligned} \quad (34)$$

where

$$\Phi_k = \Upsilon_k - \Upsilon_k \alpha_{k-w} (\alpha_{k-w}^\top \Upsilon_k \alpha_{k-w} - I)^{-1} \alpha_{k-w}^\top \Upsilon_k \quad (35)$$

and

$$\Upsilon_k = \Phi_{k-1} - \Phi_{k-1} \alpha_k (I + \alpha_k^\top \Phi_{k-1} \alpha_k)^{-1} \alpha_k^\top \Phi_{k-1}. \quad (36)$$

We conclude this section by providing an algorithmic formulation of Theorem NL4 in Algorithm 2.

C. Matching Nonlinear Systems With Delays

In this section, we highlight how the results presented thus far can be readily extended to a large class of nonlinear systems with delays. In particular, we focus on nonlinear systems where the delay affects only the state dynamics of the system but not its output. To this end, consider a nonlinear single-input, single-output, continuous-time, time-delay system described by the equations

$$\begin{aligned} \dot{x} &= f(x_{\tau_0}, \dots, x_{\tau_\zeta}, u_{\tau_{\zeta+1}}, \dots, u_{\tau_\mu}) \\ y &= h(x) \\ x(\theta) &= \phi(\theta), \quad -T \leq \theta \leq 0 \end{aligned} \quad (37)$$

where $x(\theta)$ denotes the initial condition described by the continuous function $\phi: [-T, 0] \rightarrow \mathbb{R}^n$, and $x_{\tau_j} \in \mathbb{R}^n$ denotes the translation operator applied to $x(t) \in \mathbb{R}^n$, i.e., $x_{\tau_j} = x(t - \tau_j)$ for $\tau_j \in \mathbb{R}_{\geq 0}$ with $\tau_0 = 0$ and $j = 1, \dots, \zeta$. Furthermore, the translation operator applied on the input $u(t) \in \mathbb{R}$ gives $u_{\tau_j} = u(t - \tau_j)$ for $\tau_j \in \mathbb{R}_{\geq 0}$ with $j = \zeta + 1, \dots, \mu$, and the output $y(t) \in \mathbb{R}$. In addition, $f(0, \dots, 0, 0, \dots, 0) = 0$ and $h(0) = 0$. Consider the signal generator (18) and the interconnected system described by the equations

$$\begin{aligned} \dot{\omega} &= s(\omega) \\ \dot{x} &= f(x_{\tau_0}, \dots, x_{\tau_\zeta}, l(\omega_{\tau_{\zeta+1}}), \dots, l(\omega_{\tau_\mu})) \\ y &= h(x). \end{aligned} \quad (38)$$

In order to provide a generalization of the notion of moment, consider the following assumptions.

Assumption NLTD1: There is a unique mapping $\pi(\cdot)$, locally defined in a neighborhood of $\omega = 0$ and $\pi(0) = 0$, which solves the partial differential equation

$$\frac{\partial \pi}{\partial \omega} s(\omega) = f(\pi(\bar{\omega}_{\tau_0}), \dots, \pi(\bar{\omega}_{\tau_\zeta}), l(\bar{\omega}_{\tau_{\zeta+1}}), \dots, l(\bar{\omega}_{\tau_\mu})) \quad (39)$$

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

This allows the definition of moment of the system in terms of its steady-state response.

Definition NLTD1: Consider the system (37) and the signal generator (18). Suppose Assumption NL2 and Assumption NLTD1 hold. The function $h(\pi(\cdot))$, with $\pi(\cdot)$ being the solution of (39), is the *moment of the system* (37) at (s, l) .

This allows one to define the moment of the system (37) in terms of its steady-state response. We now recall the connection between this notion of moment and the steady-state output response of the interconnected system (38).

Lemma NLTD1 (See [37]): Consider the system (37) and the signal generator (18). Suppose that Assumption NL2 and NLTD1 hold. Assume that the zero equilibrium of the system $\dot{x} = f(x_{\tau_0}, \dots, x_{\tau_\zeta}, 0, \dots, 0)$ is locally exponentially stable. Then, the moment of the system (37) at (s, l) coincides with the (locally well-defined) steady-state response of the output of the interconnected system (38).

Finally, we define a class of reduced-order models that achieve moment matching at (s, l) with this underlying nonlinear time-delay system (37).

Definition NLTD2: The system described by the equations

$$\begin{aligned}\dot{\xi} &= \varphi(\xi_{\gamma_0}, \dots, \xi_{\gamma_\theta}, u_{\gamma_{\theta+1}}, \dots, u_{\gamma_\eta}) \\ \psi &= \kappa(\xi) \\ \xi(\theta) &= \chi(\theta), \quad -T \leq \theta \leq 0\end{aligned}\quad (40)$$

with $\xi(t) \in \mathbb{R}^\rho$ is a *model at (s, l)* of the system (37) if the system (40) has the same moment at (s, l) as the system (37). Here, $\xi(\theta)$ denotes the initial condition described by the continuous function $\chi: [-T, 0] \rightarrow \mathbb{R}^\rho$ and delays $\gamma_j \geq 0$ for $j = 0, \dots, \eta$. In this case, the system (40) is said to *match* the moment of the system (37) at (s, l) . Furthermore, the system (40) is a reduced-order model of the system (37) if $\rho < n$.

Analogous to the underlying system, we introduce an assumption on the reduced-order model.

Assumption NLTD2: There is a unique mapping $p(\cdot)$, locally defined in a neighborhood of $\omega = 0$ and $p(0) = 0$, which solves the partial differential equation

$$\frac{\partial p}{\partial \omega} s(\omega) = \varphi(p(\bar{\omega}_{\gamma_0}), \dots, p(\bar{\omega}_{\gamma_\theta}), l(\bar{\omega}_{\gamma_{\theta+1}}), \dots, l(\bar{\omega}_{\gamma_\eta})) \quad (41)$$

where $\bar{\omega}_{\gamma_i} = \Phi_{\gamma_i}^s(\omega)$, with $i = 0, \dots, \mu$, is the flow of the vector field s at γ_i .

Assumptions NLTD1 and NLTD2 ensure the existence of a well-defined steady state of the underlying system (37) and the reduced-order model (40) [37]. In the model-based setting, Assumption NLTD1 is critical since the construction of the reduced-order model depends on the mapping π . In our data-driven setting, however, this assumption is automatically fulfilled as long as system (37) has a well-defined steady-state output response. For the reduced-order model, similar to the design procedure in Section III-B, the state dynamics of the reduced-order model can be easily adjusted by the designer to obtain a well-defined steady state, thus satisfying Assumption NLTD2. We can now express the moment of this reduced-order nonlinear time-delay system in terms of its steady-state output response.

Lemma NLTD2 (See [37]): Consider the system (40) and the signal generator (18). Suppose that Assumptions NL2 and NLTD2 hold. Assume that the zero equilibrium of the system $\dot{\xi} = \varphi(\xi_{\gamma_0}, \dots, \xi_{\gamma_\theta}, 0, \dots, 0)$ is locally exponentially stable. Then, the moment of the system (40) at (s, l) coincides with the (locally well-defined) steady-state response of the output of the interconnected system (18) and (40). In addition, the function $\kappa(p(\cdot))$, with $p(\cdot)$ being the solution of (41), is the moment of the system (40) at (s, l) . The system (40) matches the moment of the system (37) when

$$h(\pi(\cdot)) = \kappa(p(\cdot)).$$

Analogous to the linear time-delay case in Section II-C, the output of the systems (37) and (40) can be expressed as the moment of the system and an exponentially decaying signal and

are omitted here for brevity. We are now ready to state the final result of this section.

Theorem NLTD1: Consider the system (37), the system (40), and the signal generator (18), and assume that the respective zero equilibrium of systems (37) and (40) are locally exponentially stable. Suppose that Assumptions NL2, NL4, NL5, NLTD1, and NLTD2 hold. Then, there exist sequences $\{t_k\}$ such that

$$\lim_{k \rightarrow \infty} \left(y(t_k) - \lim_{N \rightarrow M} \tilde{\kappa}_{N,k}(\xi(t_k)) \right) = 0. \quad (42)$$

Proof: The proof for this result follows the lines of the proof of Theorem NL3 and is omitted for brevity. $\square$

The preceding result allows the direct application of Algorithm 2 to any interconnected nonlinear time-delay system of the form (38).

D. Discussion

We see benefits similar to the linear setting in the nonlinear framework. We highlight these points again for the reader's convenience.

- 1) Construction of the reduced-order models (21) [system (40) in the time-delay setting] does not require the knowledge of the structure of the signal generator or its internal state.
- 2) This procedure does not require the computation of the moment of the underlying system $h(\pi(\cdot))$ as an intermediate step during the construction of reduced-order models.
- 3) The function ϕ in the reduced-order model (21) is a free mapping that can be suitably chosen to incorporate prior knowledge about the underlying system. Herein, the designer has the freedom to enforce additional constraints on the structure of the reduced-order model, such as stability, transient behavior, delay structure, etc. The designer also has the ability to incorporate prior knowledge in the output mapping by appropriately selecting a set of basis functions that are representative of the nonlinearities affecting the system. Distinct from the model-based setting, the output mapping κ relies on a learning algorithm. In this respect, the complete structure of this output mapping is not known a priori. As such, the learning algorithm would need to be suitably constrained to account for more general input–output behaviors, such as passivity, $\mathcal{L}_2$ stability, etc.
- 4) Our method is applicable to a more general class of nonlinear systems compared to [28], [29], and [41], which are specialized for quadratic and/or bilinear systems. For more general nonlinear systems, these methods require an intermediate conversion step to a quadratic and/or bilinear form, often requiring knowledge of the underlying system and an augmentation of the number of states.
- 5) The reduced-order models obtained through this procedure are robust to variations in the signal generator (18).

As alluded to in Section III-C, the assumptions related to the existence of a well-defined steady state of the full-order model (Assumptions L1, NL1, and NLTD1 in conjunction with

Assumptions L2 and NL2, respectively) are critical in the model-based setting since the construction of the reduced-order model depends on the computation of the moment of the underlying system. In a data-driven setting, such as ours, however, these assumptions are automatically fulfilled as long as the full-order system has a well-defined steady-state output response. Furthermore, the state dynamics of the reduced-order model can be easily adjusted by the designer to obtain a well-defined steady state, thus satisfying Assumptions L4, NL3, and NLTD2, respectively. While in the linear setting, local properties implicitly extend to the global case, this is generally not true for nonlinear systems. Local properties of nonlinear systems are attached to specific trajectories of the signal generator (18). To tackle this, Algorithm 2 can be appropriately modified to account for multiple trajectories. For a detailed discussion on this issue, see [50]. This concludes the results and discussion on nonlinear systems. We now apply the results obtained thus far to a benchmark example.

IV. EXAMPLE—NONLINEAR RC LADDER

In this section, we apply Algorithm 2 developed in Section III-B on a benchmark example—the “Nonlinear RC ladder circuit.” This benchmark is part of the MORwiki examples for model reduction [63]. We consider two scenarios in this example. The first scenario considers the delay-free setting, whereas the second scenario highlights how delays can be incorporated in the construction of the reduced-order model.

A. Delay-Free Setting

1) Model Description: The nonlinear RC ladder benchmark is an electronic test circuit describing a first-order nonlinear system modeled by a resistor–capacitor network. This network exhibits a distinct nonlinear behavior due to nonlinear resistors. These nonlinear resistors, each of which consists of a resistor connected in parallel with a diode, are connected in parallel to a capacitor [64]. More details about this benchmark can be found in [65].

2) Experimental Setup: For data collection from the full-order system, we consider the model of the nonlinear RC ladder circuit as described in [66]. Herein, the RC ladder circuit, where the nonlinear resistor is composed of a resistor connected in parallel with a diode, models a resistor–capacitor network. The underlying model of this system is given by a single-input, single-output gradient system [67] of the form

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_k \\ \vdots \\ \dot{x}_N \end{bmatrix} = \begin{bmatrix} -g(x_1) - g(x_1 - x_2) + u \\ g(x_1 - x_2) - g(x_2 - x_3) \\ \vdots \\ g(x_{k-1} - x_k) - g(x_k - x_{k+1}) \\ \vdots \\ -g(x_{N-1} - x_N) \end{bmatrix} \quad (43)$$

$y = x_1$

where $g : \mathbb{R} \rightarrow \mathbb{R}$ is a mapping of the form

$$g(x_i) = g_D(x_i) + x_i$$

which combines the effect of a diode and a resistor. The diode is modeled as a nonlinear resistor based on the Shockley model [68] and is described by the relation

$$g_D(x_i) = i_S(e^{u_P x_i} - 1) \quad (44)$$

such that the material parameters $i_S > 0$ and $u_P > 0$. In this benchmark, the material parameters are chosen such that $i_S = 1$ and $u_P = 40$ [66]. We call this model FOM. For this example, we choose $N = 1000$. The input to the underlying system is of the form

$$u(t) = \sin(200\pi t) + 10 \sin(100\pi t)(1 + \sin(200\pi t)) \quad (45)$$

through which we collect input–output data from this procedure. We only describe the output equation (45) of the unknown signal generator since we are interested in constructing a reduced-order model that is agnostic to the structure of the signal generator, and therefore do not know the evolution of ω , as described in (18). Note, however, that the procedure does not require the description of the input as a closed-form expression. We show (45) only for the reader’s perusal. We simulate the underlying system FOM for a duration of 5 s and collect input–output data sampled at 10^{-4} s.

3) Model Selection: As described in Algorithm 2, for obtaining the reduced-order model, we need to provide as input the state of the reduced-order system. This means that we are free to choose the model structure describing the state-dynamics of the reduced-order model. Therefore, we describe below two model selections that we choose as the reduced-order model.

- 1) *Reduced-Order Model 1 (ROM1):* We choose the state dynamics to have the same structure as the underlying system (43) with $N = 10$.
- 2) *Reduced-Order Model 2 (ROM2):* We model the state-dynamics of the reduced-order system as a circuit of chained inverter gates, which is sometimes also known as an inverter chain [69]. This system is described by the differential equation

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_N \end{bmatrix} = - \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} + \begin{bmatrix} u \\ v \tanh(ax_1) \\ \vdots \\ v \tanh(ax_{N-1}) \end{bmatrix}$$

where $N = 10$, supply voltage $v = 100$, and a physical parameter $a = 0.5$.

4) Model Validation: Our objective is to estimate the output mapping for these reduced-order models such that they achieve moment matching with the underlying system. To accomplish this, we need to construct a family of basis functions, as discussed in Algorithm 2. Here, as in [70, eq. (2)], we consider a choice of basis functions, which is a mixture of constant, polynomial, and trigonometric terms as follows

$$\Omega(\xi) = \begin{bmatrix} 1 & \xi & \xi^2 & \xi^3 & \xi^4 & \sin(\xi) & \cos(\xi) \end{bmatrix}. \quad (46)$$

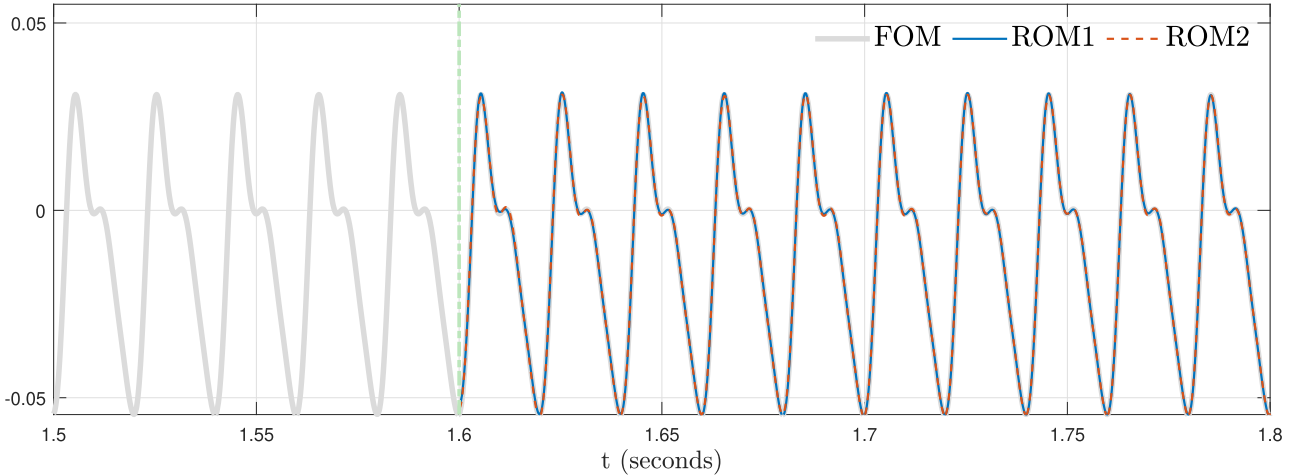


Fig. 1. Comparison of the output responses of the full-order model (FOM), with the reduced-order models (ROM1) and (ROM2), in the delay-free setting (see Section IV-A). The output response of FOM is displayed between 1.5 and 1.8 s. The estimation of the moments associated with the reduced-order models ROM1 and ROM2 uses training data to the right of the vertical dashed green line. This is done to eliminate the effect of the transient behavior of the FOM, which occurs on the left-hand side of the vertical dashed green line.

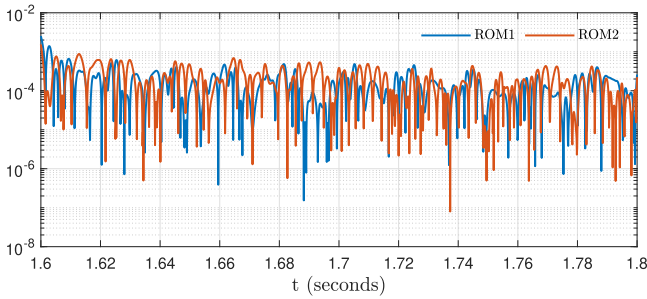


Fig. 2. Semilog plot of the absolute value of the steady-state error between the outputs of the full-order model (FOM) and the corresponding reduced-order systems ROM1 and ROM2 in a simulation of length 5 s in the delay-free setting (see Section IV-A).

We are now ready to identify $\bar{\Gamma}$ for our reduced-order model by applying Algorithm 2. We note again that the only observables that our algorithm uses are the measurables u and y . We do not require knowledge of the signal generator or its internal state ω for this procedure. Once $\bar{\Gamma}$ has been estimated, we use it to compute the output of both the reduced-order systems (ROM1 and ROM2) using (29) when driven by the signal generator (45). In Fig. 1, we compare the output responses of the underlying system and both the reduced-order models, when driven by the signal generator (45). Theorem NL3 mandates that for the estimation of the output mapping $\kappa(p(\cdot))$, the transient data from the simulation must be discarded. Therefore, we do this by considering only steady-state data for constructing the reduced-order model. The separation between the “transient” and the “steady state” in the data is represented by the vertical dashed green line in Fig. 1. As seen there, the steady-state output responses of the reduced-order models match almost exactly that of the full-order model. To analyze the choice of structure for the reduced-order models, we show in Fig. 2 the absolute value of the steady-state error between the steady-state output response

of the full-order model and the corresponding reduced-order models, ROM1 and ROM2, for a simulation of length 5 s. The error metrics associated with ROM1 appear to be better than that of ROM2. Although the errors are relatively small for ROM2, this error plot indicates that matching can be improved further by refining the choice of basis functions since, in steady state, the output of the reduced-order model should be the same as that of the full-order system. As a final note, compared to the approach described in [28, Sect. 4.2] for the same underlying system nor do we need to know the structure of the underlying system nor do we need to augment the number of states as an intermediate step.

B. Setting With Delays

To illustrate that our proposed method can be readily applied to systems with delay, we consider a modified version of the system presented in Section IV-A with a smaller number of states. Furthermore, we consider a nontrivial signal generator that is nonlinear, which, however, is unknown to the designer.

1) Model Description and Experimental Setup: We consider the nonlinear RC ladder circuit described in Section IV-A with the distinction that a delay is embedded in the state of the system. The modified underlying system has the same structure as system (43) barring the mapping $g : \mathbb{R} \rightarrow \mathbb{R}$, which is modified to incorporate the delay, that is

$$g(x_i(t)) = g_D(x_i(t)) + x_i(t - \tau_D) \quad (47)$$

with g_D as described in (44). In this example, we keep $i_S = 1$ and $u_P = 40$ but choose $N = 10$ in the full-order model for simplicity and we call this model FOMD. In addition, we choose $\tau_D = 0.05$ and drive the underlying system with a Van der Pol oscillator, which can be represented by a signal generator of the

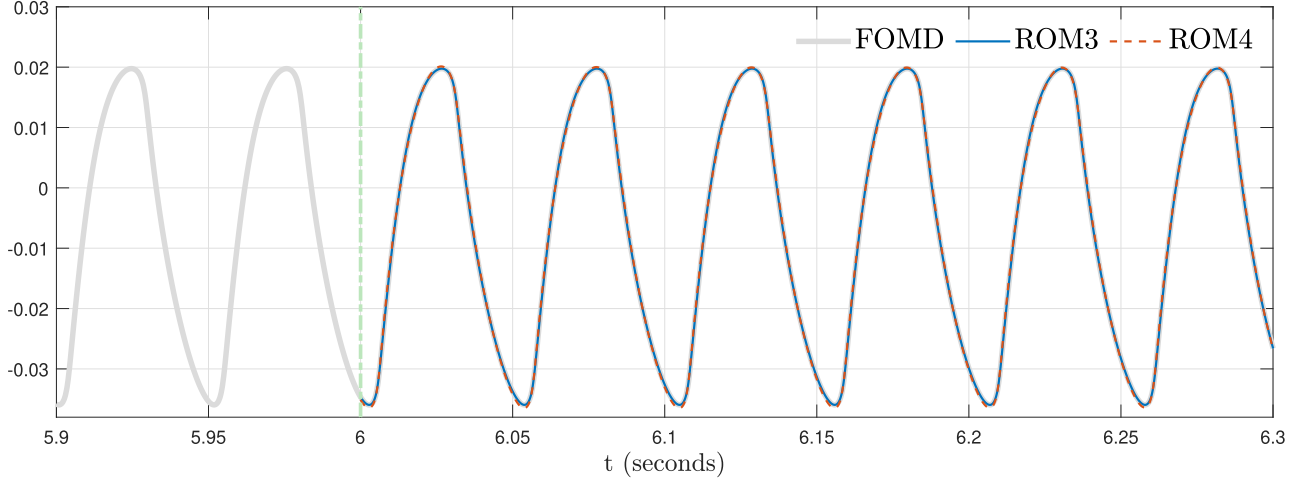


Fig. 3. Comparison of the output responses of the full-order model, (FOMD), with the reduced-order models, (ROM3) and (ROM4), in the setting with delays (see Section IV-B). The output response of FOMD is displayed between 5.9 and 6.3 s. The estimation of the moments associated with the reduced-order models ROM3 and ROM4 uses training data to the right of the vertical dashed green line. This is done to eliminate the effect of the transient behavior of the FOMD, which occurs on the left-hand side of the vertical dashed green line.

form (18), where

$$s(\omega) = \lambda \begin{bmatrix} \omega_2 \\ \mu(1 - \omega_1^2)\omega_2 - \omega_1 \end{bmatrix}, \quad l(\omega) = \varrho\omega_1 \quad (48)$$

with $\lambda = 200$, $\mu = 4$, and $\varrho = 2$. We emphasize again that we provide a description of the signal generator only for the reader's perusal, as the technique that we have presented in this work does not require the structure of the signal generator or its internal state, as seen in Algorithms 1 and 2. We simulate the underlying system FOMD for a duration of 10 s and collect input–output data sampled at 10^{-4} s.

2) Model Selection: Similar to Section IV-A, we have two model selections for the reduced-order model.

- 1) *Reduced-Order Model 3 (ROM3):* We choose the state dynamics to have the same structure as the underlying system (43) with $N = 4$. Similar to the full-order model, we use (47) to describe the effect of the delay but choose $\tau_D = 0.02$, which is distinct from the delay present in the underlying system. We make this choice to highlight that the asymptotic behavior of the algorithm allows the user to appropriately embed knowledge of the underlying system into the state dynamics of the reduced-order model. While the delays are numerically different, our procedure is able to account for this difference while constructing the output map of the reduced-order system.
- 2) *Reduced-Order Model 4 (ROM4):* We model the state-dynamics of the reduced-order system as a circuit of chained inverter gates but also incorporate a delay in the state-dynamics. This system is described by the differential equation

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_N(t) \end{bmatrix} = - \begin{bmatrix} x_1(t - \tau_D) \\ x_2(t - \tau_D) \\ \vdots \\ x_N(t - \tau_D) \end{bmatrix} + \begin{bmatrix} u \\ v \tanh(ax_1(t)) \\ \vdots \\ v \tanh(ax_{N-1}(t)) \end{bmatrix}$$

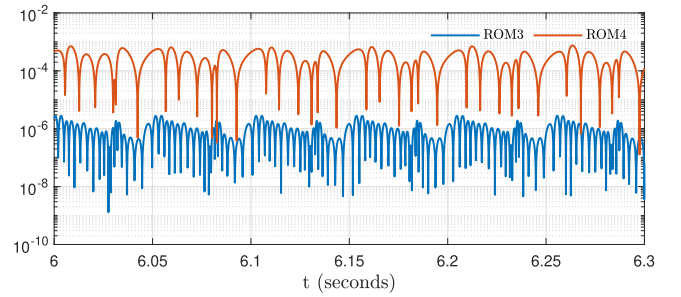


Fig. 4. Semilog plot of the absolute value of the steady-state error between the outputs of the full-order model (FOMD) and the corresponding reduced-order systems ROM3 and ROM4 in a simulation of length 10 s in the setting with delays (see Section IV-B).

where $N = 4$, $\tau_D = 0.04$, supply voltage $v = 100$, and a physical parameter $a = 0.5$.

3) Model Validation: As before, our objective is to find the output mapping for reduced-order models ROM3 and ROM4 such that they achieve moment matching with the underlying system FOMD. We use Algorithm 2 to compute $\bar{\Gamma}$, which is then used to compute the output of the reduced-order models ROM3 and ROM4 with the basis functions (46). In Fig. 3, we compare the output responses of the underlying system and both the reduced-order models when driven by the Van der Pol oscillator (48). Similar to Section IV-A, we discard transient data during the identification procedure. The separation between the “transient” and the “steady state” in the data is represented by the vertical dashed green line in Fig. 3. We plot the absolute value of the steady-state error between the output responses of the underlying system FOMD and the corresponding reduced-order models, ROM3 and ROM4, in Fig. 4. A few observations stand out in this set of results. First, we notice that the steady-state value of the state dynamics of the reduced-order model ROM4 is achieved at a much larger time compared to ROM2 in Section IV-A, due to which we use data from 6 s onward

for the computation of the output map of the reduced-order models. Very interestingly, we also notice that even with an imprecise knowledge of the delay in the underlying system, the reduced-order models are able to match the output response of the underlying system in steady state. The tradeoff is that the designer needs to wait longer to collect steady-state data, which in this case is achieved at a longer time (about 5 s for ROM4). Indeed, if one were to start with steady-state data, the batch version of Algorithm 2 would converge in one step. Finally, we observe that the reduced-order model ROM3 is more accurate than ROM4, compared to the delay-free setting involving the models ROM1 and ROM2. This, however, can be attributed to the fact that in the delay-free setting, the underlying system had a much higher dimension ($N = 1000$) as compared to the setting with delays ($N = 10$), because of which both ROM1 and ROM2 have comparable performances.

V. CONCLUSION AND PERSPECTIVES

We have presented a procedure to construct reduced-order models from input–output data, which is agnostic to the knowledge of the signal generator or its internal state for both linear and nonlinear systems. This procedure allows the construction of reduced-order models without needing to compute the moment of the underlying system as an intermediate step. Leveraging this intuition, we have shown that our procedure can be readily applied to linear and nonlinear time-delay systems without any further modification. We have provided a batch formulation of our results as well as a recursive one. The recursive formulation offers two major benefits. First, it effectively addresses the computational challenges that are typically associated with the batch formulation. Second, it inherently incorporates robustness to fluctuations in the signal generator. In particular, when operating online over a moving window of training data, the proposed framework can adaptively capture the change in moment due to a variation in the input signal. This is contingent on the change in the signal occurring once the underlying system has achieved the associated steady-state behavior. These advantages underscore the robustness and adaptability of our technique, making it a valuable tool for real-time applications, thus demonstrating the efficacy and potential of our proposed methodology. In addition, this formulation provides an easy way of enforcing additional constraints on the reduced-order model that can be used to incorporate prior knowledge about the underlying system. This could mean matching specific properties, such as matching with prescribed eigenvalues and prescribed zeros in the linear setting, whereas it could mean matching more general properties applicable to both linear and nonlinear systems. Extension of this technique to preserve physical property, such as passivity or port-Hamiltonian systems [43], [71], is under consideration.

There are several promising avenues for future research. These include extending the proposed procedure to the multivariable setting [72], [73] and adapting it to scenarios in which sensor-obtained data may be subject to noise. Indeed, while the methodology presented in this study is ideally suited for scenarios in which data are collected without the presence of noise, the findings can be seamlessly extended through the use

of Kernel-based methods as in [52]. This extension allows for the accommodation and mitigation of noise effects, thereby preventing overfitting when dealing with noisy input–output data. Furthermore, an exploration of the utility of the reduced-order models derived in this study for improving the data-driven output regulation problem—possibly for nonlinear systems—presents a promising opportunity, especially in connection with the work in [74], [75], [76], and [77].

REFERENCES

- [1] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA, USA: SIAM, 2005.
- [2] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox, *Model Reduction and Approximation: Theory and Algorithms*. Philadelphia, PA, USA: SIAM, 2017.
- [3] P. Benner, M. Hinze, and E. J. W. Ter Maten, *Model Reduction for Circuit Simulation*, vol. 74. Berlin, Germany: Springer, 2011.
- [4] P. Koutsovasilis and M. Beiteltschmidt, “Comparison of model reduction techniques for large mechanical systems: A study on an elastic rod,” *Multibody System Dyn.*, vol. 20, pp. 111–128, 2008.
- [5] I. Horenko, R. Klein, S. Dolaptchiev, and C. Schütte, “Automated generation of reduced stochastic weather models i: Simultaneous dimension and model reduction for time series analysis,” *Multiscale Model. Simul.*, vol. 6, no. 4, pp. 1125–1145, 2008.
- [6] M. J. Keeling and K. T. Eames, “Networks and epidemic models,” *J. Roy. Soc. Interface*, vol. 2, no. 4, pp. 295–307, 2005.
- [7] K. Glover, “All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ -error bounds,” *Int. J. Control*, vol. 39, no. 6, pp. 1115–1193, 1984.
- [8] V. M. Adamyan, D. Z. Arov, and M. G. Krein, “Analytic properties of schmidt pairs for a Hankel operator and the generalized Schur–Takagi problem,” *Matematicheskii Sbornik*, vol. 128, no. 1, pp. 34–75, 1971.
- [9] 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, Apr. 1990.
- [10] B. Moore, “Principal component analysis in linear systems: Controllability, observability, and model reduction,” *IEEE Trans. Autom. Control*, vol. 26, no. 1, pp. 17–32, Feb. 1981.
- [11] D. G. Meyer, “Fractional balanced reduction: Model reduction via fractional representation,” *IEEE Trans. Autom. Control*, vol. 35, no. 12, pp. 1341–1345, Dec. 1990.
- [12] W. S. Gray and J. Mesko, “General input balancing and model reduction for linear and nonlinear systems,” in *Proc. 1997 Eur. Control Conf.*, 1997, pp. 2862–2867.
- [13] 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, Jun. 2003.
- [14] A. C. Antoulas, J. A. Ball, J. Kang, and J. C. Willems, “On the solution of the minimal rational interpolation problem,” *Linear Algebra its Appl.*, vol. 137, pp. 511–573, 1990.
- [15] C. I. Byrnes, T. T. Georgiou, and A. Lindquist, “A generalized entropy criterion for nevanlinna–pick interpolation with degree constraint,” *IEEE Trans. Autom. Control*, vol. 46, no. 6, pp. 822–839, Jun. 2001.
- [16] A. Mayo and A. Antoulas, “A framework for the solution of the generalized realization problem,” *Linear Algebra its Appl.*, vol. 425, no. 2–3, pp. 634–662, 2007.
- [17] C. A. Beattie and S. Gugercin, “Interpolation theory for structure-preserving model reduction,” in *Proc. 47th IEEE Conf. Decis. Control*, 2008, pp. 4204–4208.
- [18] A. Astolfi, “Model reduction by moment matching for linear and nonlinear systems,” *IEEE Trans. Autom. Control*, vol. 55, no. 10, pp. 2321–2336, Oct. 2010.
- [19] G. Scarcioiti and A. Astolfi, “Model reduction by matching the steady-state response of explicit signal generators,” *IEEE Trans. Autom. Control*, vol. 61, no. 7, pp. 1995–2000, Nov. 2015.
- [20] Y. Nakatsukasa, O. Sète, and L. N. Trefethen, “The AAA algorithm for rational approximation,” *SIAM J. Sci. Comput.*, vol. 40, no. 3, pp. A1494–A1522, 2018.
- [21] J. D. Simard, A. Moreschini, and A. Astolfi, “Parameterization of all moment matching interpolants,” in *Proc. 2023 Eur. Control Conf.*, 2023, pp. 1–6.

- [22] A. Moreschini, J. D. Simard, and A. Astolfi, "Model reduction in the Loewner framework for second-order network systems on graphs," in *Proc. 62nd IEEE Conf. Decis. Control*, 2023, pp. 6713–6718.
- [23] D. Bhattacharjee and A. Astolfi, "Closed-loop model reduction by moment matching for linear systems," in *Proc. 62nd IEEE Conf. Decis. Control*, 2023, pp. 4954–4959.
- [24] G. Scarcioiti and A. Astolfi, "Interconnection-based model order reduction—a survey," *Eur. J. Control*, vol. 75, 2024, Art. no. 100929.
- [25] S. Lall, P. Krysl, and J. E. Marsden, "Structure-preserving model reduction for mechanical systems," *Physica D: Nonlinear Phenomena*, vol. 184, no. 1–4, pp. 304–318, 2003.
- [26] J. Sjöberg, K. Fujimoto, and T. Glad, "Model reduction of nonlinear differential-algebraic equations," *IFAC Proc. Vol.*, vol. 40, no. 12, pp. 176–181, 2007.
- [27] A. C. Antoulas, I. V. Gosea, and A. C. Ionita, "Model reduction of bilinear systems in the Loewner framework," *SIAM J. Sci. Comput.*, vol. 38, no. 5, pp. B889–B916, 2016.
- [28] I. V. Gosea and A. C. Antoulas, "Data-driven model order reduction of quadratic-bilinear systems," *Numer. Linear Algebra with Appl.*, vol. 25, no. 6, 2018, Art. no. e2200.
- [29] P. Benner, S. Gugercin, and S. W. Werner, "Structure-preserving interpolation of bilinear control systems," *Adv. Computat. Math.*, vol. 47, no. 3, 2021, Art. no. 43.
- [30] J. M. A. Scherpen and A. J. van der Schaft, "Normalized coprime factorizations and balancing for unstable nonlinear systems," *Int. J. Control*, vol. 60, no. 6, pp. 1193–1222, 1994.
- [31] K. Willcox and J. Peraire, "Balanced model reduction via the proper orthogonal decomposition," *AIAA J.*, vol. 40, no. 11, pp. 2323–2330, 2002.
- [32] W. S. Gray and J. M. Scherpen, "Hankel singular value functions from Schmidt pairs for nonlinear input–output systems," *Syst. Control Lett.*, vol. 54, no. 2, pp. 135–144, 2005.
- [33] K. Fujimoto and D. Tsubakino, "Computation of nonlinear balanced realization and model reduction based on Taylor series expansion," *Syst. Control Lett.*, vol. 57, no. 4, pp. 283–289, 2008.
- [34] B. Besselink, N. van de Wouw, J. M. Scherpen, and H. Nijmeijer, "Model reduction for nonlinear systems by incremental balanced truncation," *IEEE Trans. Autom. Control*, vol. 59, no. 10, pp. 2739–2753, Oct. 2014.
- [35] K. Gallivan, A. Vandendorpe, and P. Van Dooren, "Sylvester equations and projection-based model reduction," *J. Comput. Appl. Math.*, vol. 162, no. 1, pp. 213–229, 2004.
- [36] K. Gallivan, A. Vandendorpe, and P. Van Dooren, "Model reduction and the solution of Sylvester equations," in *Proc. 17th Int. Symp. Math. Theory Netw. Syst.*, 2006, .
- [37] G. Scarcioiti and A. Astolfi, "Model reduction of neutral linear and nonlinear time-invariant time-delay systems with discrete and distributed delays," *IEEE Trans. Autom. Control*, vol. 61, no. 6, pp. 1438–1451, Jun. 2016.
- [38] P. Schulze and B. Unger, "Data-driven interpolation of dynamical systems with delay," *Syst. Control Lett.*, vol. 97, pp. 125–131, 2016.
- [39] P. Schulze, B. Unger, C. Beattie, and S. Gugercin, "Data-driven structured realization," *Linear Algebra its Appl.*, vol. 537, pp. 250–286, 2018.
- [40] D. Karachalios, I. V. Gosea, and A. C. Antoulas, "The Loewner framework for system identification and reduction," in *Model Order Reduction: Volume I: System- and Data-Driven Methods and Algorithms*, Berlin, Germany: De Gruyter, 2021, pp. 181–228.
- [41] I. V. Gosea, C. Poussot-Vassal, and A. C. Antoulas, "Data-driven modeling and control of large-scale dynamical systems in the Loewner framework: Methodology and applications," in *Handbook of Numerical Analysis*, vol. 23, Amsterdam, The Netherlands: Elsevier, 2022, pp. 499–530.
- [42] H. Bai, J. Mao, and G. Scarcioiti, "Model reduction for quadratic-bilinear time-delay systems using nonlinear moments," *IFAC-PapersOnLine*, vol. 56, no. 1, pp. 91–95, 2023.
- [43] A. Moreschini, J. D. Simard, and A. Astolfi, "Data-driven model reduction for port-Hamiltonian and network systems in the Loewner framework," *Automatica*, vol. 169, 2024, Art. no. 111836.
- [44] M. F. Shakib, G. Scarcioiti, A. Y. Pogromsky, A. Pavlov, and N. van de Wouw, "Model reduction by moment matching with preservation of global stability for a class of nonlinear models," *Automatica*, vol. 157, 2023, Art. no. 111227.
- [45] J. D. Simard and A. Astolfi, "On the construction and parameterization of interpolants in the Loewner framework," *Automatica*, vol. 159, 2024, Art. no. 111329.
- [46] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*. Philadelphia, PA, USA: SIAM, 2016.
- [47] C. W. Rowley, T. Colonius, and R. M. Murray, "Model reduction for compressible flows using POD and Galerkin projection," *Physica D: Nonlinear Phenomena*, vol. 189, no. 1–2, pp. 115–129, 2004.
- [48] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *J. Fluid Mech.*, vol. 656, pp. 5–28, 2010.
- [49] P. J. Schmid, "Dynamic mode decomposition and its variants," *Annu. Rev. Fluid Mech.*, vol. 54, pp. 225–254, 2022.
- [50] G. Scarcioiti and A. Astolfi, "Data-driven model reduction by moment matching for linear and nonlinear systems," *Automatica*, vol. 79, pp. 340–351, 2017.
- [51] J. Mao and G. Scarcioiti, "Data-driven model reduction by moment matching for linear systems through a swapped interconnection," in *Proc. 2022 Eur. Control Conf.*, 2022, pp. 1690–1695.
- [52] A. Moreschini, M. Scandella, and T. Parisini, "Nonlinear data-driven moment matching in reproducing kernel Hilbert spaces," in *Proc. 2024 Eur. Control Conf.*, 2024, pp. 3440–3445.
- [53] A. Moreschini and A. Astolfi, "Closed-loop interpolation by moment matching for linear and nonlinear systems," *IEEE Trans. Autom. Control*, vol. 70, no. 5, pp. 2918–2933, May 2025.
- [54] D. Bhattacharjee and A. Astolfi, "Data-driven model reduction by moment matching for linear systems driven by an unknown implicit signal generator," in *Proc. 2024 Eur. Control Conf.*, 2024, pp. 3434–3439.
- [55] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems: Theory—Implementation—Applications*. Berlin, Germany: Springer, 2012.
- [56] K. J. Åström and B. Wittenmark, *Adaptive Control*. North Chelmsford, MA, USA: Courier Corporation, 2013.
- [57] A. Padoan, G. Scarcioiti, and A. Astolfi, "A geometric characterization of the persistence of excitation condition for the solutions of autonomous systems," *IEEE Trans. Autom. Control*, vol. 62, no. 11, pp. 5666–5677, Nov. 2017.
- [58] G. Scarcioiti and A. Astolfi, "Nonlinear model reduction by moment matching," *Foundations Trends Syst. Control*, vol. 4, no. 3–4, pp. 224–409, 2017.
- [59] A. Isidori, *Nonlinear Control Systems*, (3rd ed.). Berlin, Germany: Springer-Verlag London, 1995.
- [60] R. Tóth, *Modeling and Identification of Linear Parameter-Varying Systems*, vol. 403. Berlin, Germany: Springer, 2010.
- [61] J. Park and I. W. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Computation*, vol. 3, no. 2, pp. 246–257, 1991.
- [62] H. Rocha, "On the selection of the most adequate radial basis function," *Appl. Math. Modelling*, vol. 33, no. 3, pp. 1573–1583, 2009.
- [63] The MORwiki Community, "MORwiki - model order reduction wiki," n.d., [Online]. Available: <http://modelreduction.org>
- [64] M. Rewienski and J. White, "A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 22, no. 2, pp. 155–170, Feb. 2003.
- [65] The MORwiki Community, "Nonlinear RC ladder," MORwiki—Model Order Reduction Wiki, 2018.
- [66] Y. Chen, "Model order reduction for nonlinear systems," Ph.D. thesis, Massachusetts Inst. Technol., Cambridge, MA, USA, 1999.
- [67] M. Condon and R. Ivanov, "Empirical balanced truncation of nonlinear systems," *J. Nonlinear Sci.*, vol. 14, pp. 405–414, 2004.
- [68] T. Reis, *Mathematical Modeling and Analysis of Nonlinear Time-Invariant RLC Circuits*. Berlin, Germany: Springer, 2014.
- [69] C. Gu, "Model order reduction of nonlinear dynamical systems," Univ. California at Berkeley, Berkeley, CA, USA, 2011.
- [70] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proc. Nat. Acad. Sci.*, vol. 113, no. 15, pp. 3932–3937, 2016.
- [71] A. Moreschini, M. Bin, A. Astolfi, and T. Parisini, "A generalized passivity theory over abstract time domains," *IEEE Trans. Autom. Control*, vol. 70, no. 1, pp. 2–17, Jan. 2025.
- [72] M. F. Shakib, G. Scarcioiti, A. Y. Pogromsky, A. Pavlov, and N. van de Wouw, "Time-domain moment matching for multiple-input multiple-output linear time-invariant models," *Automatica*, vol. 152, 2023, Art. no. 110935.
- [73] J. D. Simard, A. Moreschini, and A. Astolfi, "Parameterization of all differential-algebraic moment matching interpolants," *IEEE Trans. Autom. Control*, vol. 70, no. 3, pp. 1875–1882, Mar. 2025.
- [74] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 909–924, Mar. 2019.

- [75] C. De Persis and P. Tesi, "Learning controllers for nonlinear systems from data," *Annu. Rev. Control*, vol. 56, 2023, Art. no. 100915.
- [76] D. Carnevale, S. Galeani, M. Sassano, and A. Serrani, "External models for output regulation based on moment estimation from input-output data," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 7777–7782, 2017.
- [77] L. Gentilini, M. Bin, and L. Marconi, "Data-driven output regulation via Gaussian processes and Luenberger internal models," *IFAC-PapersOnLine*, vol. 56, no. 1, pp. 367–372, 2023.



Debraj Bhattacharjee (Graduate Student Member, IEEE) received the M.S. degree in Electrical Engineering from the University of Minnesota—Twin Cities, Minneapolis, MN, USA, in 2018.

Between 2018 and 2022, he was a developer in the Controls and Identification group at MathWorks, Natick, MA, USA. Since 2022, he has been a part of the Control and Power research group, Department of Electrical and Electronic Engineering, Imperial College

London, U.K., where he is currently working towards a Ph.D. degree in Electrical and Electronic Engineering. His research focuses on the area of control theory with an emphasis on model reduction.



Alessio Moreschini (Member, IEEE) received the Ph.D. (Hons.) degree in Automatica from the University of Rome "La Sapienza," Rome, Italy, in 2021, and the Ph.D. degree in Systems and Control from Université Paris-Saclay, Paris, France, in 2021.

From 2021 to 2022, he was a Postdoctoral Fellow with the Department of Computer, Control, and Management Engineering, University of Rome "La Sapienza." Since 2022, he has been a Research Associate with the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K.

His research interests include the field of nonlinear systems and control theory, with a particular focus on passivity-based control, sampled-data systems, and model order reduction.

Dr. Moreschini is currently an Associate Editor for *Automatica* and the EUCA Conference Editorial Board. He is also a Member of the IEEE CSS Technical Committee on Nonlinear Systems and Control and a Corresponding Member of the IFAC Technical Committee on Nonlinear Control Systems.



Alessandro Astolfi (Fellow, IEEE) received the Laurea in electronic engineering from the University of Rome La Sapienza, Rome, Italy, in 1991, the M.Sc. degree in information theory and a Ph.D. degree (with Medal of Honor) with a thesis on discontinuous stabilization of non-holonomic systems from ETH-Zurich, Zurich, Switzerland, in 1995, and a Ph.D. degree for his work on nonlinear robust control from the University of Rome "La Sapienza," in 1996.

In 1992, he joined ETH-Zurich as a Research Associate. Since 1996, he has been with the Electrical and Electronic Engineering Department, Imperial College London, London, U.K., where he is currently a Professor of Nonlinear Control Theory and College Consul for the Faculty of Engineering and Business School. From 2010 to 2022, he was the Head of the Control and Power Group, Imperial College London. From 1998 to 2003, he was an Associate Professor with the Department of Electronics and Information, Politecnico di Milano, Milano, Italy. Since 2005, he has also been a Professor with the Dipartimento di Ingegneria Civile e Ingegneria Informatica, University of Rome Tor Vergata, Rome. His research interests include mathematical control theory and control applications, with special emphasis for the problems of discontinuous stabilization, robust and adaptive control, observer design, optimal control, game theory, and model reduction.

Dr. Astolfi was the recipient of the IEEE CSS A. Ruberti Young Researcher Prize (2007), the IEEE RAS Googol Best New Application Paper Award (2009), the IEEE CSS George S. Axelby Outstanding Paper Award (2012), the Automatica Best Paper Award (2017), IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY Outstanding Paper Award (2023), and the Institute of Measurement and Control Sir Harold Hartley Medal for "Outstanding contributions to the technology of measurement and control". He is a "Distinguished Member" of the IEEE CSS, IFAC Fellow, IET Fellow, and Member of the Academia Europaea. He was an Associate Editor, Senior Editor, and Area Editor for several journals, and Editor-in-Chief for the *European Journal of Control*. He is the Editor-in-Chief of IEEE TRANSACTIONS ON AUTOMATIC CONTROL (since 2018). He was the Chair of the IEEE CSS Conference Editorial Board (2010–2017), the Chair of the IEEE CSS Antonio Ruberti Young Researcher Prize (2015–2021), and a Member of the IEEE Fellow Committee in 2016 and from 2019 to 2022. He is the Vice Chair of the IFAC Technical Board (2020–2026), a Member of the IEEE PSPB Strategic Planning Committee, and a Member of the IEEE Thesaurus Editorial Board.