2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Data-Driven Model Reduction by Moment Matching for Linear Systems Driven by an Unknown Implicit Signal Generator

Debraj Bhattacharjee and Alessandro Astolfi

Abstract— **We study the model reduction by moment matching problem for linear systems in a data-driven framework. We show that reduced-order models can be directly computed from data without knowledge of the structure of the signal generator or of its internal state. The reduced-order models thus obtained match the moments of the unknown underlying system asymptotically. Our construction provides a simple way to enforce additional constraints in the reduced-order model. We demonstrate the applicability of the results using data from a high-dimensional model of a building.**

I. Introduction

Mathematical models have played a pivotal part in ensuring the efficient operation of large-scale dynamical systems. These models have been extensively used in the design, analysis, and subsequent control of such systems. In recent years, however, the increased complexity in these systems has made it more difficult to model the underlying process accurately, while also demanding large amounts of computational power once aggregated models are available. Therefore, it is necessary to obtain 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*.

Model reduction algorithms have been extensively used in various 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 [1]; weather forecast models, which deal with large amounts of atmospheric data [2]; and the modeling of complex mechanical systems, where the effect of flexible modes is often neglected [3].

Significant attention has been devoted to model reduction for linear systems. Two classes of methods have predominantly found popularity: algorithms based on the Singular Value Decomposition (SVD) and algorithms based on moment matching methods. The Hankel operator [4] and the celebrated balanced realization [5] belong to the first group, whereas methods based on interpolation theory belong to the latter group [6]–[8]. Moment matching methods 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.

Model reduction approaches, typically, 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 high-performance computing platforms, large-scale monitoring and collection of data from such systems is relatively easy [9]. 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 (POD) and Dynamic Mode Decomposition (DMD), which are based on the SVD, have found extensive usage, among many others, in the domain of fluid systems [10]–[13]. In the moment matching framework, a data-driven version of the Loewner framework has been presented for model reduction using frequency-domain measurements [6]. However, the use of such observables limits the extension of this procedure 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 [14]. 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 procedure requires the knowledge of the structure of the signal generator that drives the system as well as the knowledge of its internal states. However, in several applications, only qualitative information may be known about the signal generator. 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.

To deal with the 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 [7], [14] to derive reduced-order models directly from input-output data. The rest of the paper is structured as follows. The steady-state notion of moment and an algorithm to compute it from data are revisited in Section II. A procedure for computing reduced-order models without computing the moment of the system is introduced and is illustrated via an example in Section III. Finally, some concluding remarks are presented in Section IV.

Debraj Bhattacharjee and Alessandro Astolfi are with the Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ, London, U.K. (email:[d.bhattacharjee22, a.astolfi]@imperial.ac.uk)

Alessandro Astolfi is also with the Dipartimento di Ingegneria Civile e Ingegneria Informatica, University of Rome "Tor Vergata", 00133, Rome, Italy.

This work has been partially supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 739551 (KIOS CoE), by the Italian Ministry for Research in the framework of the 2017 and 2020 Program for Research Projects of National Interest (PRIN), Grants no. 2017YKXYXJ and 2020RTWES4, and by the EPSRC Grants EP/W005557 "Model reduction from data" and EP/X033546.

Notation: 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, and $\mathbb{C}_{\geq 0}$ denotes the set $\mathbb{C} \setminus \mathbb{C}_{<0}$. 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 vectorization operator of a matrix A is denoted by $\mathrm{vec}(A)$ and is obtained by stacking A's columns.

II. Preliminaries

In this section, we first recall the steady-state based description of moment as described in [7], [15]. We then revisit the procedure to compute the moment of the system from data given in [14]. Finally, we introduce a formulation that allows constructing reduced-order models without the knowledge of the structure of the signal generator.

A. The steady-state notion of moment

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

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ y &= Cx, \end{aligned} \tag{1}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, and $y(t) \in \mathbb{R}$. Let the associated transfer function $W(s) = C(sI - A)^{-1}B$ be minimal, i.e., the triple (C, A, B) is controllable and observable.

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

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

The notion of moment for system (1) can be described in terms of a Sylvester equation, as outlined below.

Lemma 1: [7] Consider system (1), let $s_i \in \mathbb{C}$ be such that $s_i \notin \sigma(A)$, for all $i = 1, \cdots, \eta$. Then there exists a one-to-one relation between the moments $\eta_0(s_1), \cdots, \eta_{k_1-1}(s_1), \cdots, \eta_0(s_\eta), \cdots, \eta_{k_\eta-1}(s_\eta)$ and the matrix $C\Pi$, where Π is the unique solution of the Sylvester equation

$$A\Pi + BL = \Pi S, \tag{2}$$

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

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

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

This formulation allows a relationship to be established, through the Sylvester equation, between the moments of a system and its steady-state output response, as outlined next.

Theorem 1: [7] Consider system (1). Let $s_i \in \mathbb{C}$ be such that $s_i \notin \sigma(A)$, for all $i = 1, \cdots, n$, and assume that $\sigma(A) \subset \mathbb{C}_{<0}$. Let $S \in \mathbb{R}^{\nu\times\nu}$ be any non-derogatory matrix with characteristic polynomial as defined in (3). Consider the interconnection of system (1) with the signal generator

$$\begin{aligned} \dot{\omega} &= S\omega, \\ u &= L\omega, \end{aligned} \tag{4}$$

such that the pair (L, S) is observable and the triple $(L, S, \omega(0))$ is minimal. Then there exists a one-to-one relation between the moments $\eta_0(s_1), \cdots, \eta_{k_1-1}(s_1), \cdots, \eta_0(s_\eta), \cdots, \eta_{k_\eta-1}(s_\eta)$ and the steady-state output of the interconnected system (1) and (4).

Exploiting Lemma 1 and Theorem 1, the notion of reduced-order model can be defined as follows. The system described by the equations

$$\begin{aligned} \dot{\xi} &= F\xi + Gu, \\ \psi &= H\xi, \end{aligned} \tag{5}$$

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

$$C\Pi = HP, \tag{6}$$

where Π is the unique solution of the the Sylvester equation (2), such that the pair (L, S) is observable, and P is the unique solution of the Sylvester equation

$$FP + GL = PS. \tag{7}$$

Equation (6) is the so-called *moment matching* condition. In addition, system (5) is said to be a reduced-order model of system (1) if $\nu < n$.

B. Computing moment from data

While deriving reduced-order models, one usually starts with a high-dimensional model of the underlying system. This, of course, assumes the knowledge of such a 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. We now recall a data-driven procedure that utilizes input-output data to compute the moment of the system with subsequent construction of reduced-order models.

Theorem 2: [14] Let the time snapshots $Q_k \in \mathbb{R}^{w\times n\nu}$ and $\rho_k \in \mathbb{R}^w$, with $w \geq n\nu$, be defined as

$$Q_k = \begin{bmatrix} (w(0)^T \otimes C)(e^{S^T t_{k-w+1}} \otimes I - I \otimes e^{A^T t_{k-w+1}}) \\ \vdots \\ (w(0)^T \otimes C)(e^{S^T t_{k-1}} \otimes I - I \otimes e^{A^T t_{k-1}}) \\ (w(0)^T \otimes C)(e^{S^T t_k} \otimes I - I \otimes e^{A^T t_k}) \end{bmatrix},$$

and

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

respectively. Assume that the matrix Q_k has full column rank. Then

$$\text{vec}(C\Pi_k) = (Q_k^T Q_k)^{-1} Q_k^T \rho_k. \tag{8}$$

Equation (8) is solved over a set of sample times given by $T_k^w = \{t_{k-w+1}, \ldots, t_{k-1}, t_k\}$ such that $0 \leq t_0 < t_1 < \ldots < t_{k-w} < \ldots < t_k < \ldots < t_v$. T_k^w is a moving window of w sample times where $w \geq 0$ and $v \geq w$. $C\Pi_k$ is the estimate of the matrix $C\Pi$ at T_k^w, meaning that the estimate computed at time t_k uses data sampled from the last w instants of time.

The formulation defined in Theorem 2 requires the knowledge of the system matrices and the initial conditions of the underlying model for computing the moment of the system. This requirement can be circumvented by employing the following approximation.

Theorem 3: [14] Let the time snapshots $\tilde{Q}_k \in \mathbb{R}^{w \times \nu}$ and $\tilde{\rho}_k \in \mathbb{R}^w$, with $w \geq \nu$, be defined as

$$\tilde{Q}_k = \left[\omega(t_{k-w+1}) \ldots \omega(t_{k-1}) \; \omega(t_k)\right]^T,$$

and

$$\tilde{\rho_k} = \left[y(t_{k-w+1}) \ldots y(t_{k-1}) \; y(t_k)\right]^T.$$

Assume that the matrix $\tilde{Q}_k$ has full column rank. Then

$$\text{vec}(\widetilde{C\Pi_k}) = (\tilde{Q}_k^T \tilde{Q}_k)^{-1} \tilde{Q}_k^T \tilde{\rho}_k$$

is an approximation of the on-line estimate $C\Pi_k$, namely there exists a sequence $\{t_k\}$ such that $\lim_{k\to\infty}(\widetilde{C\Pi_k}) = C\Pi$. Choosing $P = I$ in equations (6) and (7) yields a family of reduced-order models for system (1) at S that achieve moment matching [7]. The state-space realization of this family of reduced-order models is described as in equation (5) with

$$F = S - GL, \quad H = C\Pi.$$

C. Problem formulation

The procedure for obtaining reduced-order models described in Section II-B requires full information about the structure of the signal generator via the use of S and L. However, 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, consider the following assumptions and properties.

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

Assumption 2: The signal generator defined by equation (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.

Assumption 3: The elements of T_k^v are selected such that $\text{rank}\Big(\big[\omega(t_{k-v+1}) \; \ldots \; \omega(t_k)\big]\Big) = \nu$, for all k.

Assumption 2 guarantees that the signals generated by (4) are persistently exciting [16]–[18]. Assumption 3 ensures that $\tilde{Q}_k$ (defined in the previous section) is a full-rank square matrix. Similar to Assumption 1, we impose a stability condition on the reduced-order model (5).

Assumption 4: System (5) is asymptotically stable, i.e. $\sigma(F) \subset \mathbb{C}_{<0}$ and the pair (F, G) is controllable.

Finally, the time response of both the full-order and of the reduced-order models can be obtained as follows.

Proposition 1: [7], [14] Consider the interconnection of system (1) and the signal generator (4) and suppose that Assumptions 1 and 2 hold. Then system (1) has a global invariant manifold described by $\mathcal{M} = \{(x, \omega) \in \mathbb{R}^n \times \mathbb{R}^\nu \,|\, x = \Pi\omega)\}$, that is for all $t \in \mathbb{R}_{\geq 0}$, we have that

$$x(t) - \Pi\omega(t) = e^{At}(x(0) - \Pi\omega(0)). \tag{9}$$

Proposition 2: [7], [14] Consider the interconnection of system (5) and the signal generator (4) and suppose that Assumptions 2 and 4 hold. Then system (5) has a global invariant manifold described by $\mathcal{M} = \{(\xi, \omega) \in \mathbb{R}^\nu \times \mathbb{R}^\nu \,|\, \xi = P\omega)\}$, that is for all $t \in \mathbb{R}_{\geq 0}$, we have that

$$\xi(t) - P\omega(t) = e^{Ft}(\xi(0) - P\omega(0)). \tag{10}$$

III. Direct computation of reduced-order models from data

In this section, we outline a procedure to directly obtain reduced-order models from input-output data when 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, we show that reduced-order models can be derived from input-output data without the knowledge of the structure of the signal generator or of its internal state. We note again that our 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.

A. Batch estimation of reduced-order models

We begin this section by showing how reduced-order models can be computed from a batch of input-output data. To begin with, let H_k be the estimate of the matrix H at T_k^w.

Theorem 4: Consider system (1), system (5), and the signal generator (4). Suppose that Assumptions 1-4 hold. Let the time snapshots $R_k \in \mathbb{R}^{w \times \nu}$ and $\gamma_k \in \mathbb{R}^w$, with $w \geq \nu$, be defined as

$$R_k = \begin{bmatrix} (P\omega(t_{k-w+1}))^T \\ \vdots \\ (P\omega(t_{k-1}))^T \\ (P\omega(t_k))^T \end{bmatrix},$$

and

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

respectively, and $\bar{x}(0) = x(0) - \Pi\omega(0)$. Assume that the matrix R_k has full column rank. Then

$$\text{vec}(H_k) = (R_k^T R_k)^{-1} R_k^T \gamma_k. \tag{11}$$

Proof: From equation (6) we have that

$$C\Pi = HP,$$

which can be rewritten as

$$C\Pi\omega(t) = HP\omega(t). \tag{12}$$

From Proposition 1, we know that

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

which leads to

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

Multiplying by C on both sides, we obtain

$$C\Pi\omega(t) = y(t) - Ce^{At}\bar{x}(0).$$

Using the moment matching condition (12), we have that

$$HP\omega(t) = y(t) - Ce^{At}\bar{x}(0). \tag{13}$$

Finally, we use the Kronecker product and the vectorization operator to obtain

$$\text{vec}(HP\omega(t)) = \text{vec}(y(t) - Ce^{At}\bar{x}(0)),$$

which leads to

$$(P\omega(t))^T\text{vec}(H) = \text{vec}(y(t) - Ce^{At}\bar{x}(0)). \tag{14}$$

By computing equation (14) at all elements of T_k^w, we obtain

$$R_k\text{vec}(H_k) = \gamma_k. \tag{15}$$

Since by assumption the matrix R_k has full column rank, we can compute H_k from equation (15), yielding equation (11). ■

We solve equation (11) over a set of sample times given by T_k^w. T_k^w is a moving window of w sample times where $w \geq 0$ and $v \geq w$ (see Section II-B). As discussed at the beginning of this section, the estimate of H computed at time t_k uses data sampled from the last w instants of time.

The result presented in Theorem 4 contains terms that depend on the system model and the initial conditions of the underlying system and of the signal generator. However, by Assumption 1 we know that $\sigma(A) \subset \mathbb{C}_{<0}$, implying that these terms exponentially decay to zero. We now present an approximate version of the result presented in Theorem 4.

To this end, suppose that Assumptions 1-4 hold. Let the time snapshots $\tilde{R}_k \in \mathbb{R}^{w\times\nu}$ and $\tilde{\gamma}_k \in \mathbb{R}^w$, with $w \geq \nu$, be defined as

$$\tilde{R}_k = \left[(P\omega(t_{k-w+1})) \ \cdots \ (P\omega(t_{k-1})) \ (P\omega(t_k))\right]^T,$$

and

$$\tilde{\gamma}_k = \left[y(t_{k-w+1}) \ \cdots \ y(t_{k-1}) \ y(t_k)\right]^T,$$

respectively. Assume that the matrix $\tilde{R}_k$ has full column rank. Then

$$\text{vec}(\tilde{H}_k) = (\tilde{R}_k^T\tilde{R}_k)^{-1}\tilde{R}_k^T\tilde{\gamma}_k. \tag{16}$$

Although ostensibly the computation of the term $P\omega$ requires the knowledge of P and ω, as can be seen from equation (10), one does not need information about either quantity in practice. This is because by Assumption 4, $\sigma(F) \subset \mathbb{C}_{<0}$, which implies that $P\omega$ converges to the state of the reduced-order model asymptotically as shown in the next statement.

Lemma 2: Consider the interconnection of system (5) and the signal generator (4), and suppose that Assumptions 2-4 hold. Then for any sequence $\{t_k\}$, $\lim_{t_k\to\infty}(P\omega(t_k) - \xi(t_k)) = 0$.

Proof: By Proposition 2 we know that, for all $t \in \mathbb{R}_{\geq 0}$,

$$\xi(t) - P\omega(t) = e^{Ft}(\xi(0) - P\omega(0)).$$

Therefore, since $\sigma(F) \subset \mathbb{C}_{<0}$ by Assumption 4, for any sequence $\{t_k\}$ we obtain

$$\lim_{t_k\to\infty}(\xi(t_k) - P\omega(t_k)) = \lim_{t_k\to\infty} e^{Ft_k}(\xi(0) - P\omega(0)) = 0.$$

■

Therefore, one can use the state of the reduced-order model as a proxy for $P\omega$ in equation (16). In fact, one can control the speed of convergence of $P\omega$ and ξ by appropriately selecting the eigenvalues of F. We now show that $\tilde{H}_k$ converges to H.

Lemma 3: Consider system (1), system (5), and the signal generator (4). Suppose that Assumptions 1-4 hold. Then there exists a matrix $\bar{H}$ such that $\lim_{k\to\infty}\tilde{H}_k = \bar{H}$.

Proof: By Assumptions 1 and 2, there exists a matrix $\bar{H}$ such that the steady-state output response y^{ss} of the interconnection of system (1) with the signal generator (4) is given by $y^{ss}(t) = \bar{H}P\omega(t)$, for all $t \geq 0$. By substituting $\gamma_k^{ss} = \tilde{R}_k\text{vec}(\bar{H})$ in equation (16), we obtain

$$\lim_{k\to\infty}\text{vec}(\tilde{H}_k) = (\tilde{R}_k^T\tilde{R}_k)^{-1}\tilde{R}_k^T\gamma_k^{ss} = \text{vec}(\bar{H}),$$

hence the claim. ■

We are now ready to state and prove the main result of this section.

Theorem 5: Let H be the solution of equation (6). Suppose that Assumptions 1-4 hold. Then there exist sequences $\{t_k\}$ such that $\lim_{k\to\infty}\tilde{H}_k = H$.

Proof: From equation (13) we have

$$HP\omega(t_k) = y(t_k) - Ce^{At_k}\bar{x}(0). \tag{17}$$

Similarly, from equation (16) we have

$$\tilde{H}_kP\omega(t_k) = y(t_k). \tag{18}$$

Using equations (17) and (18), we obtain

$$(\tilde{H}_k - H)P\omega(t_k) = Ce^{At_k}\bar{x}(0).$$

By Assumptions 2-4, there exist sequences $\{t_k\}$ with $\lim_{k\to\infty} t_k = \infty$, such that for any $t_i \in \{t_k\}$, $P\omega(t_i) \neq 0$ (see [19]). Therefore, since $\sigma(A) \subset \mathbb{C}_{<0}$ by Assumption 1, we obtain

$$\lim_{k\to\infty}(\tilde{H}_k - H)P\omega(t_k) = \lim_{k\to\infty} Ce^{At_k}\bar{x}(0) = 0.$$

By Assumption 3 and Lemma 3, we obtain

$$\lim_{k\to\infty}(\tilde{H}_k - H) = \lim_{k\to\infty}(\bar{H} - H) = 0.$$

This implies that $\tilde{H}_k$ asymptotically converges to H. ■

B. Online estimation of reduced-order models

As seen in equation (16), the computation of $\mathrm{vec}(\tilde{H}_k)$ at each time instant involves the inversion of a matrix of dimension $\nu \times \nu$. 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 introduce a recursive version of (16) that alleviates this issue.

Theorem 6: Let $\alpha_k = P\omega(t_k)$ and assume that $\Phi_k = (\tilde{R}_k^T \tilde{R}_k)^{-1}$ and $\Upsilon_k = ((\tilde{R}_{k-1}^T \tilde{R}_{k-1})^{-1} + \alpha_k \alpha_k^T)^{-1}$ are full rank for all $t \geq t_r$, where $t_r > t_w$. Given $\mathrm{vec}(\tilde{H}_r)$, Φ_r, α_r, and Υ_r, the recursive least-square formulation for $\mathrm{vec}(\tilde{H}_k)$ for all $t \geq t_r$ is given by

$$\begin{aligned}\mathrm{vec}(\tilde{H}_k) &= \mathrm{vec}(\tilde{H}_{k-1}) + \Phi_k \alpha_k (y(t_k) - \alpha_k^T \mathrm{vec}(\tilde{H}_{k-1})) \\ &\quad - \Phi_k \alpha_{k-w} (y(t_{k-w}) - \alpha_{k-w}^T \mathrm{vec}(\tilde{H}_{k-1})),\end{aligned} \tag{19}$$

where

$$\Phi_k = \Upsilon_k - \Upsilon_k \alpha_{k-w} (\alpha_{k-w}^T \Upsilon_k \alpha_{k-w} - I)^{-1} \alpha_{k-w}^T \Upsilon_k, \tag{20}$$

and

$$\Upsilon_k = \Phi_{k-1} - \Phi_{k-1} \alpha_k (I + \alpha_k^T \Phi_{k-1} \alpha_k)^{-1} \alpha_k^T \Phi_{k-1}. \tag{21}$$

Proof: The proof of this result is similar to that of the online algorithm described in [14], and is therefore omitted. ■

We note that the matrices that need to be inverted in the recursive formulation are in fact scalars, which can be computed easily and more efficiently as compared to the computation in the batch formulation. Even if the number of interpolation points (and correspondingly, the size of the reduced-order model) were to be increased, matrix inversions would not be affected. If one has knowledge about the structure of the signal generator, P and ω (and consequently $P\omega$) are obtained readily. In the case in which the structure of the signal generator is unknown, one can use ξ in its stead as $P\omega$ converges to ξ asymptotically, as shown in Lemma 2.

When compared to existing results in the literature (see [14] for example), our results are advantageous on three fronts while maintaining the same computational benefits. First, using existing methods, one cannot construct reduced-order models if the signal generator is not known while, in our work, we do not need to know the structure of the signal generator or its internal state. Secondly, in our method, there is no need to compute the moment of the full-order system as an intermediate step while obtaining reduced-order models. Finally, in our formulation, both F and G are free parameters that can be chosen freely to preserve properties of the underlying system such as matching with prescribed eigenvalues, matching with prescribed relative degree, etc. We conclude this section with an example.

C. Example: Hospital building model

In this section, we apply the results developed in the previous sections to the model of the Los Angeles Hospital building. This building has 8 floors, each having three degrees of freedom [14], [20]. The model of this building can be represented in state-space form such as (1), with $n = 48$. The motion of the first coordinate is the output of the system. As part of the simulation environment, we design the implicit signal generator such that the interpolation points are $\pm 5.22\,i, \pm 13.8\,i, \pm 25.2\,i$, corresponding to $\nu = 6$. Similarly, we set the parameter L of the signal generator to $L = \begin{bmatrix} 1\,1\,1\,1\,1\,1 \end{bmatrix}$. We note that the quantities S, L, and ω are not exposed to our algorithm. The only observables that we utilize are u and y. In the remainder of this example, we show how one can match the dynamics of the underlying system by exploiting the structure of the reduced-order model.

1) A naive reduced-order model: In the first part of this example, we choose our reduced-order model with the assumption that we do not know anything about the dynamics of the underlying system. Therefore, without any other prior information about the full-order model, our reduced-order model is naively chosen with

$$F = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 & -6 \end{bmatrix}, \tag{22}$$

$$G = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T.$$

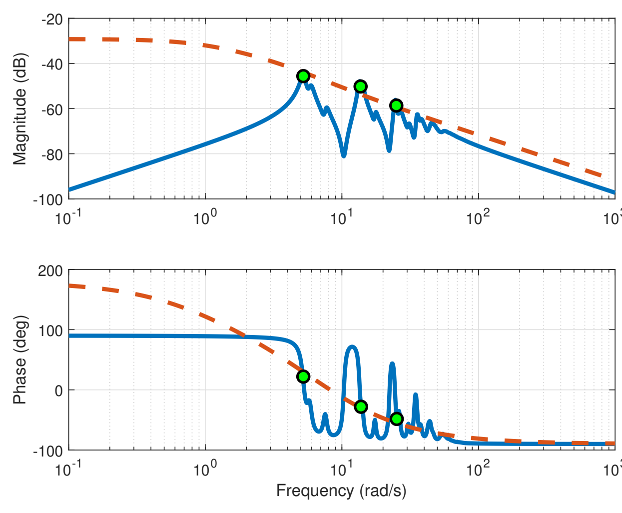

Fig. 1. Bode plot of the building model (solid blue line) and the estimated reduced-order model (dashed orange line) presented in Section III-C.1. The green dots indicate the interpolation points.

We apply the result presented in Theorem 6 for estimating the reduced-order model from u and y. The estimation procedure is terminated when $||\tilde{H}_{k+1} - \tilde{H}_k||$ is below a user-specified threshold. Figure 1 shows the Bode plot of the actual building model and the resulting reduced-order model along with the interpolation points (marked with green dots). As can be seen, the moment of the reduced-order model coincides with that of the underlying system at the interpolation points even though the dynamics of both systems significantly differ at other frequencies. To improve matching, we can incorporate prior knowledge of the underlying system in the reduced-order model, as we show next.

2) Matching with prescribed eigenvalues: If we have prior knowledge about the underlying system, we can integrate this knowledge in our reduced-order model using the results presented in the previous section. Specifically, in this part, we demonstrate how we can create a reduced-order model with prescribed eigenvalues. Suppose that from previous knowledge or other experiments, we are able to ascertain that the building model has eigenvalues at $-0.27 \pm 7.63\,i$, $-0.40 \pm 17.55\,i$, $-0.61 \pm 26.45\,i$ (the original building model has 48 complex eigenvalues, six of which are shown here). We can now very easily construct our reduced-order model to include this information. We design a new matrix F, defined in (22), such that its eigenvalues are the ones that we want to match while keeping G to be the same. Following the result presented in Theorem 6, we run the estimation procedure as we did in the first part of this example. Figure 2 shows the Bode plot of the building model and the resulting reduced-order model, with a newly designed F, along with the interpolation points (marked with green dots).

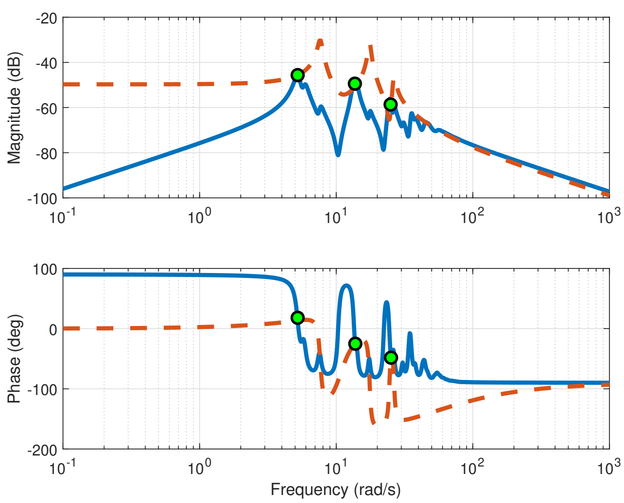

Fig. 2. Bode plot of the building model (solid blue line) and the estimated reduced-order model (dashed orange line) presented in Section III-C.2. The green dots indicate the interpolation points.

It is evident from Figure 2 that incorporation of this additional information in the construction of the reduced-order model greatly improves the shape of its frequency response when compared to the original building model itself. Similarly other properties of the full-order system can be preserved easily by utilizing the results that we have presented in this paper.

IV. Conclusion

We have presented a data-driven procedure to compute reduced-order models directly from input-output data without necessitating the need to compute the moment of the underlying system. The reduced-order models thus obtained asymptotically match the moments of the high-dimensional full-order system to be reduced. In addition, our formulation allows an easy way of enforcing additional constraints on the reduced-order model. The results have been presented in a batch formulation as well as in a recursive one. The recursive formulation overcomes computational challenges that may be encountered in the batch formulation. Finally, we highlight the advantages in our formulation when compared to existing methods.

References

[1] P. Benner, M. Hinze, and E. J. W. Ter Maten, *Model Reduction for Circuit Simulation*. Springer, 2011, vol. 74.

[2] 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 Modeling & Simulation*, vol. 6, no. 4, pp. 1125–1145, 2008.

[3] P. Koutsovasilis and M. Beitelschmidt, "Comparison of model reduction techniques for large mechanical systems: a study on an elastic rod," *Multibody System Dynamics*, vol. 20, pp. 111–128, 2008.

[4] K. Glover, "All optimal hankel-norm approximations of linear multivariable systems and their L^∞-error bounds," *International journal of control*, vol. 39, no. 6, pp. 1115–1193, 1984.

[5] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE transactions on automatic control*, vol. 26, no. 1, pp. 17–32, 1981.

[6] A. Mayo and A. Antoulas, "A framework for the solution of the generalized realization problem," *Linear algebra and its applications*, vol. 425, no. 2-3, pp. 634–662, 2007.

[7] A. Astolfi, "Model reduction by moment matching for linear and nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2321–2336, 2010.

[8] G. Scarciotti and A. Astolfi, "Model reduction by matching the steady-state response of explicit signal generators," *IEEE Transactions on Automatic Control*, vol. 61, no. 7, pp. 1995–2000, 2015.

[9] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*. SIAM, 2016.

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

[11] 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.

[12] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of fluid mechanics*, vol. 656, pp. 5–28, 2010.

[13] ——, "Dynamic mode decomposition and its variants," *Annual Review of Fluid Mechanics*, vol. 54, pp. 225–254, 2022.

[14] G. Scarciotti and A. Astolfi, "Data-driven model reduction by moment matching for linear and nonlinear systems," *Automatica*, vol. 79, pp. 340–351, 2017.

[15] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. SIAM, 2005.

[16] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems: Theory—Implementation—Applications*. Springer Science & Business Media, 2012.

[17] K. J. Åström and B. Wittenmark, *Adaptive Control*. Courier Corporation, 2013.

[18] A. Padoan, G. Scarciotti, and A. Astolfi, "A geometric characterization of the persistence of excitation condition for the solutions of autonomous systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5666–5677, 2017.

[19] J. Z. Hearon, "Nonsingular solutions of $TA - BT = C$," *Linear Algebra and its applications*, vol. 16, no. 1, pp. 57–63, 1977.

[20] A. Antoulas, J. Ball, J. Kang, and J. Willems, "On the solution of the minimal rational interpolation problem," *Linear Algebra and its Applications*, vol. 137, pp. 511–573, 1990.