

# Data-Driven Internal Model Control of Second-Order Discrete Volterra Systems

Juan G. Rueda-Escobedo<sup>1</sup> and Johannes Schiffer<sup>1</sup>

**Abstract**—The increase in system complexity paired with a growing availability of operational data has motivated a change in the traditional control design paradigm. Instead of modeling the system by first principles and then proceeding with a (model-based) control design, the data-driven control paradigm proposes to directly characterize the controller from data. By exploiting a fundamental result of Willems and collaborators, this approach has been successfully applied to linear systems, yielding data-based formulas for many classical linear controllers. In the present paper, the data-driven approach is extended to a class of nonlinear systems, namely second-order discrete Volterra systems. Two main contributions are made for this class of systems. At first, we show that - under a necessary and sufficient condition on the input data excitation - a data-based system representation can be derived from input-output data and used to replace an explicit system model. That is, the fundamental result of Willems *et al.* is extended to this class of systems. Subsequently a data-driven internal model control formula for output-tracking is derived. The approach is illustrated via two simulation examples.

## I. INTRODUCTION

The steady drive for digitization of engineering processes has greatly increased the amount of available system input-output data [1]–[3]. At the same time, the system complexity has also increased [1], [3], hindering the model derivation from first principles. This has motivated a change in the manner in which the control design is approached [1]. Instead of modeling the system by using first principles and then proceeding to the control design, the new paradigm is to take advantage of the available data to directly characterize a controller, whose structure is typically assumed known. This approach is referred to as *data-driven control* [1], [4].

However, in order to provide standard closed-loop performance guarantees, such as stability and robustness, it is still necessary to assume a certain structure of the system to be controlled. Clearly, a natural first approach is to assume that the underlying system dynamics are linear. In this respect, data-driven results for linear quadratic tracking [5], [6], dynamical feedback [7], state-feedback and optimal control [8]–[13] as well as model predictive control [14], [15] have been developed. The collection of these contributions practically cover a wide range of classical design techniques for linear systems. Therefore, to further progress in the consolidation of the data-driven control theory, it is necessary to take a step forward and start developing methods allowing to deal with nonlinear systems.

This motivates the present note, which is devoted to the extension of the data-driven theory to nonlinear Discrete Volterra Systems (DVSs). DVSs have been classically used to approximate the behavior of nonlinear systems and they play an important role in nonlinear system identification theory [16]–[18]. Many classical nonlinear systems, like Hammerstein and Wiener systems with smooth nonlinearities, can be described by DVSs [19, Sec. 2]. Likewise, under mild assumptions on the output nonlinearity, Luré type systems can be represented by DVSs [19, Sec. 2]. Additionally, multilayer neural networks can be modeled as the interconnection of several Wiener systems and can thus be approximated by high-order DVSs [20]–[23]. In the area of process control, DVSs have shown to be of great usefulness for modeling and control design [24]–[27] and also be employed to represent a class of communication systems [28]. Very recently, a trajectory-based approach to represent Hammerstein and Wiener systems has been proposed in the context of data-driven control [29] with future applications in model predictive control. Hence, there already is an interest in extending the data-driven theory to systems contained in the class of DVSs.

An outstanding subclass within the broad class of DVSs are second-order DVSs. This (sub)class of systems has a strong connection with the bispectrum [30], which helps to characterize stochastic signals beyond Gaussian sequences, and has applications, for example, in amplitude and frequency modulation [31]. Another important application for second-order DVSs is the approximation of the input-output response of bilinear systems when using explicit Runge-Kutta formulas for the discretization [24],[19, Sec. 4.7]. Finally, working with this (sub)class of systems represents a “first step beyond linearity” while keeping a reasonable dimensionality and number of parameters [19, Sec. 3].

Many of the abovementioned results on data-driven control for linear systems exploit a fundamental lemma derived by Willems and collaborators within the behavioral systems framework [32]. The importance of Willems *et al.*'s result in a data-driven context is that it establishes a necessary and sufficient condition on the required excitation of the input data of a linear system, such that all its possible trajectories can be represented solely by means of the system's input-output data. Building upon this fact, existing data-driven control approaches seek to directly parametrize the controller through data, see e.g. [9]. As a consequence, in order to extend the data-driven control approach from linear to nonlinear systems similar tools are required.

In this context, the present work contains the following

<sup>1</sup>J. G. Rueda-Escobedo and J. Schiffer are with the Fachgebiet Regelungssysteme und Netzleittechnik, Brandenburgische Technische Universität Cottbus - Senftenberg {ruedaesc, schiffer}@b-tu.de.

two main contributions for data-driven control of DVSs:

- 1) A data-based representation theory for second-order DVSs is developed, hence extending the fundamental result for linear systems [32, Theo. 1], [9, Lem. 2] to this class of nonlinear systems. More precisely, under a necessary and sufficient condition on the required excitation of the input data, our result enables the replacement of an explicit system model by input-output data.
- 2) A data-driven internal model control (IMC) formula for output-tracking in second-order DVSs is given. In general, even if the system dynamics are known, designing an output-tracking control for DVSs is challenging due to their nonlinear characteristics<sup>1</sup>. A popular class of controllers that can achieve the desired output tracking are IMCs [19], [25], [33]. Thus, our second main contribution is to derive an IMC formula, which is directly parametrized through data. This is accomplished by using the data-based representation results from 1).

The paper organization is as follows. In Section II the notation used along the note and the class of systems under study are introduced. In Section III the data-based representation theory for second-order DVSs is presented. A data-driven IMC formula for output-tracking in second-order DVSs are developed in Section IV. Finally, in Section V, the application of the methods developed in Section III and Section IV is illustrated via numerical simulations by controlling an ideal DVS and a bilinear system.

## II. PRELIMINARIES

### A. Notation

Along the note,  $\mathbb{Z}$  represents the set of integer numbers and  $\mathbb{R}$  the set of real numbers. Let  $\mathbb{F}$  be either  $\mathbb{Z}$  or  $\mathbb{R}$ . Then  $\mathbb{F}_{>0}$  ( $\mathbb{F}_{\geq 0}$ ) denotes the set of all elements of  $\mathbb{F}$  greater than (or equal to) zero.  $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ , with  $n \in \mathbb{Z}_{>0}$ , denotes the  $n \times n$  identity matrix. The Kronecker product of two matrices  $A \in \mathbb{R}^{n \times m}$  and  $B \in \mathbb{R}^{r \times s}$  is denoted by  $A \otimes B$ .

Let  $A \in \mathbb{R}^{n \times n}$  be a symmetric matrix. We denote by  $\text{vech}(A) \in \mathbb{R}^{\frac{n}{2}(n+1)}$  the column vector containing the entries in the lower triangular part of  $A$ , i.e.,

$$\text{vech}(A) = [ a_{(1,1)} \quad a_{(2,1)} \quad a_{(2,2)} \quad \cdots \quad a_{(n,n)} ]^\top.$$

Given a signal  $u : \mathbb{Z} \rightarrow \mathbb{R}^m$ , the column vector  $u_{[k,k+T]} \in \mathbb{R}^{m \cdot T}$ , with  $k \in \mathbb{Z}$  and  $T \in \mathbb{Z}_{>0}$ , corresponds to

$$u_{[k,k+T]} = [ u^\top(k) \quad u^\top(k+1) \quad \cdots \quad u^\top(k+T-1) ]^\top.$$

Given a signal  $u : \mathbb{Z} \rightarrow \mathbb{R}^m$ , the Hankel matrix  $U_{\{k,L,N\}} \in \mathbb{R}^{(m \cdot L) \times N}$  associated to  $u(k)$ , with  $k \in \mathbb{Z}$ ,  $L \in \mathbb{Z}_{\geq 0}$  and

$N \in \mathbb{Z}_{>0}$  is given by

$$U_{\{k,L,N\}} = \begin{bmatrix} u(k) & u(k+1) & \cdots & u(k+N-1) \\ u(k+1) & u(k+1) & \cdots & u(k+N) \\ \vdots & \vdots & \ddots & \vdots \\ u(k+L-1) & u(k+L) & \cdots & u(k+L+N-2) \end{bmatrix}. \quad (1)$$

Consider a vector  $u \in \mathbb{R}^m$ ,  $u = [u_1 \ u_2 \ \cdots \ u_m]^\top$ . By  $u^2 \in \mathbb{R}^{\frac{m}{2}(m+1)}$  we denote the vector

$$u^2 = \text{vech}(uu^\top) = [ u_1^2 \quad u_1 u_2 \quad u_2^2 \quad \cdots \quad u_m^2 ]^\top. \quad (2)$$

For a matrix  $A \in \mathbb{R}^{n \times m}$ ,  $A^\dagger$  denotes the Moore-Penrose pseudoinverse of  $A$ . If  $A$  has either full-row rank or full-column rank, it follows that [34, Prop. 6.1.5]

$$A^\dagger = \begin{cases} (A^\top A)^{-1} A^\top & \text{if } \text{rank}(A) = m, \\ A^\top (A A^\top)^{-1} & \text{if } \text{rank}(A) = n. \end{cases}$$

### B. Second-order discrete Volterra systems (DVSs)

In this note, we focus on data-driven control for *second order M-dimensional discrete Volterra systems (DVSs)*. These systems are described by the input-output equation [19], [35]

$$y(k) = \sum_{i=0}^M \alpha_{i+1} u(k-i) + \sum_{i=0}^M \sum_{j=0}^M \beta_{(i+1,j+1)} u(k-i) u(k-j), \quad (3)$$

with  $k \in \mathbb{Z}$ ,  $M \in \mathbb{Z}_{>0}$  the memory length of the system,  $u(k) \in \mathbb{R}$  the system input signal,  $y(k) \in \mathbb{R}$  the system output signal and  $\alpha_i \in \mathbb{R}$  and  $\beta_{(i,j)} \in \mathbb{R}$  the system parameters, with  $i, j = \{1, 2, \dots, M+1\}$ . Following standard practice [19, Sec. 3.1], we assume symmetricity of the system *kernels* from which it follows that  $\beta_{i,j} = \beta_{j,i}$ .

The system (3) can be understood as the combination of two operators applied to the input signal  $u(k)$ . Denote these operators by  $\mathbf{P}_1$  and  $\mathbf{P}_2$ . Then, the system (3) can be equivalently described as

$$y(k) = \mathbf{P}_1(u(k)) + \mathbf{P}_2(u(k)), \quad (4)$$

with

$$\begin{aligned} \mathbf{P}_1(u(k)) &= \theta_1^\top \mu(k), \\ \mathbf{P}_2(u(k)) &= \theta_2^\top \mu^2(k), \\ \mu(k) &= [ u(k) \quad u(k-1) \quad \cdots \quad u(k-M) ]^\top, \\ \theta_1 &= [ \alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_{M+1} ]^\top, \\ \theta_2 &= [ \beta_{(1,1)} \quad 2\beta_{(1,2)} \quad \cdots \quad \beta_{(M+1,M+1)} ]^\top, \end{aligned} \quad (5)$$

where the vector  $\mu(k)$  contains the input  $u(k)$  and its first  $M$  delays and  $\mu^2(k)$  follows the convention introduced in (2). As can be seen from (5),  $\mathbf{P}_1$  is linear with respect to the input  $u(k)$  while  $\mathbf{P}_2$  it is not. Furthermore,  $\mathbf{P}_1$  is a causal and stable operator.

<sup>1</sup>DVSs are inherently bounded-input-bounded-output (BIBO) stable systems. Therefore, internal stability is guaranteed and thus only output-tracking is relevant for control [19], [25], [33].

### III. DATA-BASED SYSTEM REPRESENTATION OF DVSS

In this section we provide an input-output data-based representation of the second-order DVS (4). The proofs of all results contained in this section are given in the Appendix.

For the presentation of the results and in analogy to the linear case, see e.g. [9, Def. 1], it is convenient to introduce a persistence of excitation concept *ad hoc* to the class of systems under study, i.e., the DVS (4). To this end, consider an observed input sequence  $u_{ob,[-M,T]}$  and its associated sequences  $\mu_{ob,[0,T]}$  and  $\mu_{ob,[0,T]}^2$  where, for  $k \in [0, T-1]$ , their elements have to be understood as in (5) and (2). Consider the Hankel matrices corresponding to the associated sequences  $\mu_{ob}(k)$  and  $\mu_{ob}^2(k)$ , i.e.,

$$M_{ob,\{0,L,T-L+1\}} = \begin{bmatrix} \mu_{ob}(0) & \mu_{ob}(1) & \cdots & \mu_{ob}(T-L) \\ \mu_{ob}(1) & \mu_{ob}(2) & \cdots & \mu_{ob}(T-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{ob}(L-1) & \mu_{ob}(L) & \cdots & \mu_{ob}(T-1) \end{bmatrix}, \quad (6)$$

$$M_{ob}^2_{ob,\{0,L,T-L+1\}} = \begin{bmatrix} \mu_{ob}^2(0) & \mu_{ob}^2(1) & \cdots & \mu_{ob}^2(T-L) \\ \mu_{ob}^2(1) & \mu_{ob}^2(2) & \cdots & \mu_{ob}^2(T-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{ob}^2(L-1) & \mu_{ob}^2(L) & \cdots & \mu_{ob}^2(T-1) \end{bmatrix}, \quad (7)$$

as well as the block matrix

$$\mathcal{M}_{ob,\{0,L,T-L+1\}} = \begin{bmatrix} M_{ob,\{0,L,T-L+1\}} \\ M_{ob}^2_{ob,\{0,L,T-L+1\}} \end{bmatrix}. \quad (8)$$

We introduce the following notion, cf. [9, Def. 1].

**Definition 1.** Consider the sequence  $u_{ob,[-M,T]}$  and associated sequences  $\mu_{ob,[0,T]}$  and  $\mu_{ob,[0,T]}^2$  built from  $u_{ob,[-M,T]}$  following (5) and (2). The sequence  $u_{ob,[-M,T]}$  is persistently exciting of order  $L > 0$  w.r.t. the system (4) if the matrix  $\mathcal{M}_{ob,\{0,L,T-L+1\}}$  given in (8) has full-row rank, i.e., if it has rank equal to  $\frac{1}{2}L(M+1)(M+4)$ .

Now we are in the position of introducing the main data-based representation result of this note, which is a natural counterpart to the representation theorems for linear systems reported in [9, Lem. 2], [32, Theo. 1].

**Theorem 1 (Data-based trajectory representation of a second-order DVS).** Consider the system (4) and observed input-output signals  $u_{ob}(k)$  and  $y_{ob}(k)$ . Construct the  $T$ -long sequences  $\mu_{ob,[0,T]}$ ,  $\mu_{ob,[0,T]}^2$  and  $y_{ob,[0,T]}$  according to (5) and (2). In addition, follow (6) and (7), to build the associated Hankel matrices  $M_{ob,\{0,L,T-L+1\}}$ ,  $M_{ob}^2_{ob,\{0,L,T-L+1\}}$  and  $Y_{ob,\{0,L,T-L+1\}}$ , with  $T \gg L > 0$ .

- i) *Iff  $u_{ob,[-M,T]}$  is persistently exciting of order  $L$  w.r.t. the system (4), then for every  $L$ -long sequences  $\mu_{[0,L]}$ ,  $\mu_{[0,L]}^2$  and  $y_{[0,L]}$  associated to the system (4), there*

exists  $g \in \mathbb{R}^{T-L+1}$ , such that

$$\begin{bmatrix} \mu_{[0,L]} \\ \mu_{[0,L]}^2 \\ y_{[0,L]} \end{bmatrix} = \begin{bmatrix} M_{ob,\{0,L,T-L+1\}} \\ M_{ob}^2_{ob,\{0,L,T-L+1\}} \\ Y_{ob,\{0,L,T-L+1\}} \end{bmatrix} g. \quad (9)$$

- ii) *Consider an arbitrary sequence  $v : [-M, L-1] \rightarrow \mathbb{R}$  and corresponding vectors  $\nu(\tau) = [v(\tau) v(\tau-1) \cdots v(\tau-M)]^\top$  and  $\nu^2(\tau)$  according to (2), for  $\tau = \{0, 1, \dots, L-1\}$ . Iff  $u_{ob,[-M,T]}$  is persistently exciting of order  $L$  w.r.t. the system (4), then the  $L$ -long sequences  $\mu_{[0,L]}$ ,  $\mu_{[0,L]}^2$  and  $y_{[0,L]}$  generated by*

$$\begin{bmatrix} \mu_{[0,L]} \\ \mu_{[0,L]}^2 \\ y_{[0,L]} \end{bmatrix} = \begin{bmatrix} M_{ob,\{0,L,T-L+1\}} \\ M_{ob}^2_{ob,\{0,L,T-L+1\}} \\ Y_{ob,\{0,L,T-L+1\}} \end{bmatrix} g, \quad (10)$$

with  $g$  given by

$$g = \mathcal{M}_{ob,\{0,L,T-L+1\}}^\dagger \begin{bmatrix} \nu_{[0,L]} \\ \nu_{[0,L]}^2 \end{bmatrix} + \left( \mathbf{I}_{T-L+1} - \mathcal{M}_{ob,\{0,L,T-L+1\}}^\dagger \mathcal{M}_{ob,\{0,L,T-L+1\}} \right) w, \quad (11)$$

where  $\mathcal{M}_{ob,\{0,1,T\}}$  is given in (8) and  $w \in \mathbb{R}^{T-L+1}$  is an arbitrary vector, correspond to trajectories of the system (4).

The claims in Theorem 1 exhibit two important differences with respect to the related results for linear systems [9], [32]. First, Theorem 1, part ii) establishes that not every linear combination of the input-output data contained in the Hankel matrices  $M_{ob,\{0,L,T-L+1\}}$ ,  $M_{ob}^2_{ob,\{0,L,T-L+1\}}$  and  $Y_{ob,\{0,L,T-L+1\}}$ , actually corresponds to a trajectory  $\{\mu_{[0,L]}, \mu_{[0,L]}^2, y_{[0,L]}\}$  of the system (4). Second, every trajectory  $\{\mu_{[0,L]}, \mu_{[0,L]}^2, y_{[0,L]}\}$  of the system can be represented as a linear combination of the Hankel matrices  $M_{ob,\{0,L,T-L+1\}}$ ,  $M_{ob}^2_{ob,\{0,L,T-L+1\}}$  and  $Y_{ob,\{0,L,T-L+1\}}$ , a characteristic that is reflected in Theorem 1, part i).

Now, as a direct consequence of Theorem 1, and for  $L = 1$ , a result analogous to [9, Theo. 1] is presented. It enables the replacement of the model-based representation of the system in (3) or (4) by data.

**Corollary 1 (Data-based representation of a second-order DVS).** Consider the system (4) and observed input-output sequences  $u_{ob,[-M,T]}$  and  $y_{ob,[0,T]}$ . Iff  $u_{ob,[-M,T]}$  is persistently exciting of order  $L = 1$  w.r.t. the system (4), then the system (4) can be equivalently represented through data as

$$\begin{aligned} \mu(k) &= [u(k) \quad u(k-1) \quad \cdots \quad u(k-M)]^\top, \\ y(k) &= Y_{ob,\{0,1,T\}} \mathcal{M}_{ob,\{0,1,T\}}^\dagger \begin{bmatrix} \mu(k) \\ \mu^2(k) \end{bmatrix}, \end{aligned} \quad (12)$$

where  $\mathcal{M}_{ob,\{0,1,T\}}$  is defined in (8) and  $Y_{ob,\{0,1,T\}}$  is the associated Hankel matrix to  $y_{ob,[0,T]}$ .

The two results of this section, Theorem 1 and Corollary 1, are analogous to the classical representation theorems for linear systems. Theorem 1 provides a representation of the input-output trajectories of the system (4) and, in this

respect, describes the *behavior* of the system. Corollary 1, on the other hand, is intended to represent the input-output behavior via data, thus, to replace the system model. It is also important to remark that the excitation condition required in Corollary 1 corresponds to classical requirements for the parameter identification in this class of systems [19, Sec. 4.1].

For the data-driven control design tackled in the next section, it is convenient to independently represent the linear and nonlinear operators  $\mathbf{P}_1$  and  $\mathbf{P}_2$ , respectively. This is achieved via the following specialization of Corollary 1, which can be obtained by splitting  $\mathcal{M}_{ob,\{0,1,T\}}$  in (12) in two separate matrices.

**Corollary 2 (Data-based representation of  $\mathbf{P}_1$  and  $\mathbf{P}_2$ ).** *Consider the system (4) and observed input-output sequences  $u_{ob,[-M,T]}$  and  $y_{ob,[0,T]}$ . Iff  $u_{ob,[-M,T]}$  is persistently exciting of order  $L = 1$  w.r.t. the system (4), then the operators  $\mathbf{P}_1$  and  $\mathbf{P}_2$  in (4) can be represented through data as*

$$\begin{aligned} \mu(k) &= [ u(k) \quad u(k-1) \quad \cdots \quad u(k-M) ]^\top, \\ \mathbf{P}_1(u(k)) &= P_1 \mu(k), \\ \mathbf{P}_2(u(k)) &= P_2 \mu^2(k), \end{aligned} \quad (13)$$

where the matrices  $P_1$  and  $P_2$  are defined in (14).

#### IV. DATA-DRIVEN FORMULA FOR INTERNAL MODEL CONTROL OF DVSS

Next, the results of the previous section are specialized in order to be applied for output-tracking of second-order DVSSs by means of a data-driven IMC algorithm. As the name suggests, IMCs are model-based and, essentially, rely on inverting the system input-output map. More precisely, as illustrated in Figure 1, the fundamental idea of the IMC principle is as follows [19], [25], [33]. Given a reference trajectory  $y_r : \mathbb{Z} \rightarrow \mathbb{R}$ , use the system model represented by the operator  $\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2$  (see (5)) to produce a feed-forward term and add it to the tracking error  $\tilde{y}$  in order to cancel the known part of the system response. Then, the inverse of the system model  $\mathbf{P}^{-1}$  is employed to produce the input that, in the absence of uncertainties, will achieve the desired output.

Hence, given a reference trajectory  $y_r : \mathbb{Z} \rightarrow \mathbb{R}$  for the system (4), the corresponding IMC algorithm reads [19], [25]

$$u = \mathbf{P}^{-1}(\mathbf{P}(u) - y + y_r). \quad (15)$$

To derive a data-driven IMC formula for the second-order DVS (4), we recall that  $\mathbf{P}_1$  in (4) is a linear causal stable operator and make the following two standard assumptions [19, Sec. 6.2].

##### Assumption 1.

- i) *The linear operator  $\mathbf{P}_1$  in (4) has minimum phase, i.e., all its zeros lay strictly inside the unit circle.*
- ii) *The inverse of the operator  $\mathbf{I} + \mathbf{P}_1^{-1}\mathbf{P}_2$  exists, where  $\mathbf{I}$  corresponds to the identity operator.*

Under Assumption 1, the inverse operator  $\mathbf{P}_1^{-1}$  is stable, and the inverse of the whole system  $\mathbf{P}^{-1}$  can be implemented

following Figure 2, see [19, Sec. 6]. It follows that, compared to the general IMC structure, for the implementation of (15) only  $\mathbf{P}_1^{-1}$  is needed and not the inverse of the nonlinear operator  $\mathbf{P}_2$ .

To streamline our main result, the following preliminary lemma is helpful, in which a data-based representation of the inverse operator  $\mathbf{P}_1^{-1}$  is provided. This requires to introduce the part of the observed output data corresponding to  $\mathbf{P}_1$ , which we denote by  $y_{1ob,[0,T]}$ . By using Corollary 2,  $y_{1ob,[0,T]}$  can be extracted following any of the next two expressions:

$$\begin{aligned} y_{1ob,[0,T]} &= (\mathbf{I}_T \otimes P_1) \mu_{ob,[0,T]} \\ &= y_{ob,[0,T]} - (\mathbf{I}_T \otimes P_2) \mu_{ob,[0,T]}^2, \end{aligned} \quad (16)$$

with  $P_1$  and  $P_2$  as in (14). Let  $u(k)$  be the output of  $\mathbf{P}_1^{-1}$  when applied to  $y_1(k)$ , i.e.,  $u(k) = \mathbf{P}_1^{-1}(y_1(k))$ . The following result characterizes  $\mathbf{P}_1^{-1}$  through data. Its proof is given in the Appendix.

**Lemma 1 (Data-based representation of  $\mathbf{P}_1^{-1}$ ).** *Consider the system (4) and observed input-output sequences  $u_{ob,[-M,T]}$  and  $y_{ob,[0,T]}$ . Assume that  $u_{ob,[-M,T]}$  is persistently exciting of order  $L = 1$  w.r.t. to the system (4). Build  $y_{1ob,[0,T]}$  following (16). Iff  $y_{1ob,[0,T]}$  is persistently exciting of order one in the linear sense, i.e., the Hankel matrix  $Y_{1ob,\{0,1,T\}}$  has full-row rank, then the inverse operator  $\mathbf{P}_1^{-1}$  can be represented through data as*

$$\begin{aligned} u(k) &= U_{ob,\{0,1,T\}} \left[ \frac{Y_{1ob,\{0,1,T\}}}{X_{ob,\{0,1,T\}}} \right]^\dagger \begin{bmatrix} y(k) \\ \chi(k) \end{bmatrix}, \\ \chi(k) &= [ u(k-1) \quad u(k-2) \quad \cdots \quad u(k-M) ]^\top, \end{aligned} \quad (17)$$

where  $U_{ob,\{0,1,T\}}$ ,  $X_{ob,\{0,1,T\}}$  and  $Y_{1ob,\{0,1,T\}}$  are the Hankel matrices associated to  $u_{ob,[0,T]}$ ,  $\chi_{ob,[0,T]}$  and  $y_{1ob,[0,T]}$ , respectively.

It is important to emphasize that  $\chi(k)$  in (17) stores the past outputs and plays the role of internal state of the inverse operator  $\mathbf{P}_1^{-1}$ . We are now in the position to state our main result. Recall the matrices  $U_{ob,\{0,1,T\}}$ ,  $Y_{1ob,\{0,1,T\}}$  and  $X_{ob,\{0,1,T\}}$  defined in Lemma 1 as well as the matrix  $P_1$  defined in (14).

**Theorem 2 (Data-driven IMC formula for output tracking).** *Consider the system (4) with Assumption 1 and observed input-output sequences  $u_{ob,[-M,T]}$  and  $y_{ob,[0,T]}$ . Let  $y_r : \mathbb{Z} \rightarrow \mathbb{R}$  be a desired reference trajectory. Iff  $u_{ob,[-M,T]}$  is persistently exciting of order  $L = 1$  w.r.t. the system (4), then the IMC algorithm in (15) corresponding to the block diagrams in Figures 1 and 2 can be implemented through the observed data as*

$$\begin{aligned} u(k) &= U_{ob,\{0,1,T\}} \left[ \frac{Y_{1ob,\{0,1,T\}}}{X_{ob,\{0,1,T\}}} \right]^\dagger \begin{bmatrix} P_1 \mu(k) + \tilde{y}(k) \\ \chi(k) \end{bmatrix}, \\ \chi(k) &= [ u(k-1) \quad u(k-2) \quad \cdots \quad u(k-M) ]^\top, \\ \mu(k) &= [ u(k) \quad u(k-1) \quad \cdots \quad u(k-M) ]^\top, \\ \tilde{y}(k) &= y_r(k) - y(k). \end{aligned} \quad (18)$$

$$P_1 = Y_{ob,\{0,1,T\}} \left( \mathbf{I}_T - (M_{ob,\{0,1,T\}}^2)^\dagger M_{ob,\{0,1,T\}}^2 \right) M_{ob,\{0,1,T\}}^\dagger \left( \mathbf{I}_{M+1} - M_{ob,\{0,1,T\}} (M_{ob,\{0,1,T\}}^2)^\dagger M_{ob,\{0,1,T\}}^2 M_{ob,\{0,1,T\}}^\dagger \right)^{-1},$$

$$P_2 = Y_{ob,\{0,1,T\}} \left( \mathbf{I}_T - M_{ob,\{0,1,T\}}^\dagger M_{ob,\{0,1,T\}} \right) (M_{ob,\{0,1,T\}}^2)^\dagger \left( \mathbf{I}_{\frac{1}{2}(M+1)(M+2)} - M_{ob,\{0,1,T\}}^2 M_{ob,\{0,1,T\}}^\dagger M_{ob,\{0,1,T\}} (M_{ob,\{0,1,T\}}^2)^\dagger \right)^{-1} \quad (14)$$

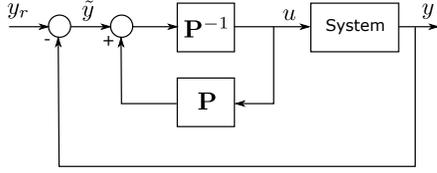


Fig. 1. Basic structure of the Internal Model Control (IMC) with reference  $y_r$  and tracking error  $\tilde{y}$ , based on [19, Sec. 6].

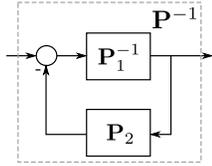


Fig. 2. Implementation of the system inverse  $\mathbf{P}^{-1}$  using the nonlinear operator  $\mathbf{P}_2$  and the inverse of the linear operator  $\mathbf{P}_1$ , both defined in (5).

*Proof.* Consider the block diagram in Figure 1 and replace  $\mathbf{P}^{-1}$  by its implementation in Figure 2. By reducing the block diagram at the summation point and noting that  $\mathbf{P}(u) - \mathbf{P}_2(u) = \mathbf{P}_1(u)$ , the input  $u$  results in

$$u = \mathbf{P}_1^{-1}(\mathbf{P}_1(u) - y + y_r) = \mathbf{P}_1^{-1}(\mathbf{P}_1(u) + \tilde{y}). \quad (19)$$

Thus, for the implementation of (19) only the linear operator  $\mathbf{P}_1$  and its inverse are needed. The data-based representations of the operators  $\mathbf{P}_1$  and  $\mathbf{P}_1^{-1}$  are obtained from Corollary 2 and Lemma 1, respectively. The combination of expressions (13) and (17) yields (18), i.e., the equivalent data-driven formula to (19).  $\square$

The standard approach to implement the IMC (15) is to pursue a two-step design procedure: In the first step, the parameters  $\alpha_i$  and  $\beta_{(i,j)}$  in (3) are identified. Then, in the second step the inverse of the linear part is computed using, e.g., the  $Z$ -transform and the IMC algorithm is implemented. In contrast, Theorem 2 follows the spirit of data-driven control [1], [4], [9] by providing the concise formula (18), that enables a direct implementation of the IMC (15) from data in a single step.

## V. SIMULATION EXAMPLE

To illustrate the application of the proposed data-driven IMC formula (18), two scenarios are considered. At first, we investigate the control performance for an ideal second-order DVS of the form (4). In the second example, we then evaluate the robustness of the derived methodology via its application to a continuous-time bilinear system, which does not correspond exactly to a second-order DVS. In addition, the impact of noise on the control performance is investigated.

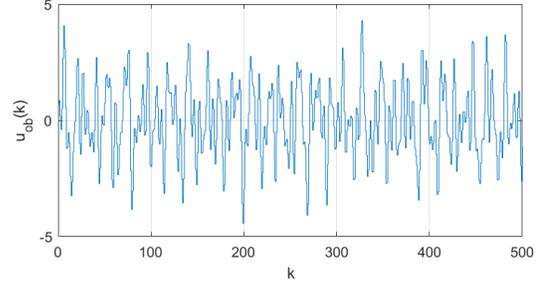


Fig. 3. Example 1: Input signal used to characterize the second-order DVS.

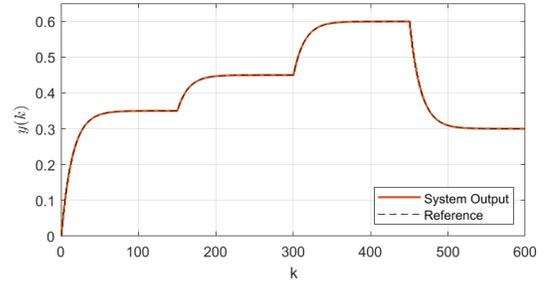


Fig. 4. Example 1: By using the proposed data-driven IMC formula (18), perfect reference tracking of the second-order DVS is achieved.

**Example 1.** For the first scenario, a second-order DVS (4), (5) of dimension  $M = 5$  is considered with the parameters

$$\theta_1^\top = [ 4 \quad 3 \quad 0.82 \quad 0.156 \quad -0.014 \quad -0.006 ],$$

$$\theta_2^\top = [ 0.8147 \quad 0.9058 \quad 0.127 \quad 0.9134 \quad 0.6324 \quad 0.0975 \\ 0.2785 \quad 0.5469 \quad 0.9575 \quad 0.9649 \quad 0.1576 \quad 0.9706 \\ 0.9572 \quad 0.4854 \quad 0.8003 \quad 0.1419 \quad 0.4218 \quad 0.9157 \\ 0.7922 \quad 0.9595 \quad 0.6557 ].$$

The parameters are chosen such that the conditions in Assumption 1 are satisfied. The system is excited with a superposition of discrete sinusoidal signals, shown in Figure 3. After verifying that the signal is persistently exciting, see Definition 1, the collected input-output data is used to compute the IMC using directly the formula (18). The output reference for the tracking process is shown in Figure 4. As can be seen, perfect reference tracking is achieved. This is possible because the considered system coincides exactly with a second-order DVS and the use of the inverse model in the IMC imposes this behavior.

**Example 2.** Now, the following bilinear system is considered

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) + x_2(t)u(t), \\ \dot{x}_2(t) &= -x_1(t) - 4x_2(t) - x_1(t)u(t) + u(t), \\ y(t) &= x_1(t) - x_2(t). \end{aligned} \quad (20)$$

It is known that the response of bilinear systems can be approximated by a DVS, though in general the DVS order is significantly higher than two [19, Sec. 4.7], [35]. Hence, this represents a non ideal scenario for our derived results.

The system (20) is simulated using a fourth-order Runge-Kutta method with a step size of  $1 \times 10^{-4}$  [s]. At the input and output of the system (20), a zero-order hold with a sampling time of 1.5 [s] is introduced in order to obtain sampled input-output data. Additionally, uniform noise in the interval  $[-0.05, 0.05]$  (which corresponds to approximately 20% of the maximum output signal amplitude) is added to the measured output. As in Example 1, the data is generated using a superposition of discrete sinusoidal input signals. The system initial conditions for this process are  $x_1(0) = 1$  and  $x_2(0) = 0$ . The observed input-output data is shown in Figure 5.

To implement a data-driven output-tracking IMC for the system (20), its dynamics is approximated by the second-order DVS (4), (5). Then the controller formula (18) is used. Thereby, the key design parameter is the dimension  $M$ , which determines the number of parameters of the second-order DVS, see (5). It follows from Definition 1 that there is an inherent trade-off between the magnitude of  $M$  and the amount of required data needed to verify persistence of excitation. To evaluate the influence of  $M$  on both the data requirements and the control performance, we pursue three IMC designs with  $M = 3$ ,  $M = 4$  and  $M = 5$ .

In the present case, we can verify that the generated input signal is persistently exciting for all our three choices of  $M$ . Hence, we proceed to the IMC design via (18). The resulting tracking performance for zero initial conditions and a smooth reference signal is shown in Figure 6. The corresponding tracking errors are given in Figure 7. Since the bilinear system (20) is not a second-order DVS, exact reference tracking is not achieved. However, all three controllers are capable of tracking the reference with acceptable accuracy, showing that the approach can be used even when the system does not exactly corresponds to a DVS. However, it is clear from Figure 7 that the controller for  $M = 5$  has higher error peaks and exhibits oscillations. This is an indicative of overfitting and thus, it seems recommendable to use the controller with the lowest dimension.

To robustify the IMC against model uncertainties and external disturbances, it is conventional to add a filter in between the error signal and the system inverse [19]. Furthermore, to deal with non-minimum phase systems, an inner-outer factorization of the linear part is required. We are currently exploring, how to incorporate such approaches in a data-driven setting.

## VI. CONCLUSIONS

In this work, the fundamental result of Willems and collaborators on data-based representation of the input-output behavior of linear systems has been extended to an important class of nonlinear systems, namely second-order DVSs. That is, we have provided an explicit data-dependent representation of a second-order DVS requiring only finite, but rich

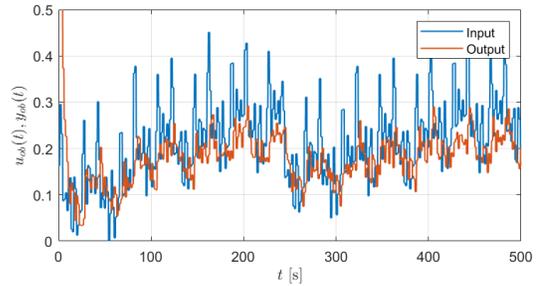


Fig. 5. Example 2: Observed input-output data from the bilinear system (20).

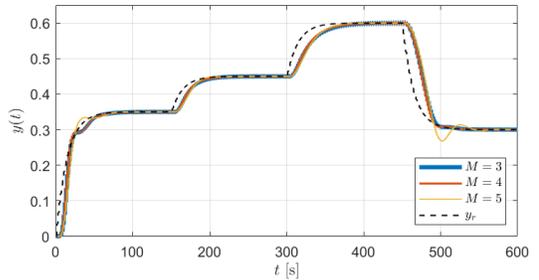


Fig. 6. Example 2: Output reference tracking with the data-driven IMC (18) using a second-order DVS of dimension  $M = 3$ ,  $M = 4$  and  $M = 5$ , respectively.

enough, data.

By using this new characterization, a *data-driven internal model control* for output-tracking of this class of systems has been proposed. Compared to the conventional two-step “system identification/control design” procedure, we provide a single formula, in which the controller itself is directly parametrized through data. The performance and robustness of this approach have been illustrated by controlling a bilinear system with noisy measurements.

The methodology can be extended to DVSs of arbitrary order, work that is under development. Additionally, it is under investigation how to address non-minimum phase systems and how to attenuate model mismatches, two problems that frequently appear when the methodology is applied to nonlinear systems that do not correspond exactly to DVSs.

## APPENDIX

### A. Proof of Theorem 1

**Part i)** Define  $\sigma_1 = L(M + 1)$  and  $\sigma_2 = \frac{1}{2}L(M + 1)(M + 2)$ . Any  $L$ -long sequences  $\mu_{[0,L]}$ ,  $\mu_{[0,L]}^2$  and  $y_{[0,L]}$  corresponding to the system (4) can be represented as

$$\begin{bmatrix} \mu_{[0,L]} \\ \mu_{[0,L]}^2 \\ y_{[0,L]} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{\sigma_1} & 0_{\sigma_1 \times \sigma_2} \\ 0_{\sigma_2 \times \sigma_1} & \mathbf{I}_{\sigma_2} \\ \mathbf{I}_L \otimes \theta_1^\top & \mathbf{I}_L \otimes \theta_2^\top \end{bmatrix} \begin{bmatrix} \mu_{[0,L]} \\ \mu_{[0,L]}^2 \end{bmatrix}, \quad (22)$$

with  $\theta_1$  and  $\theta_2$  as in (5). If  $u_{ob,[ -M, T]}$  is persistently exciting of order  $L$ , i.e., the rank condition in Definition 1 holds, then

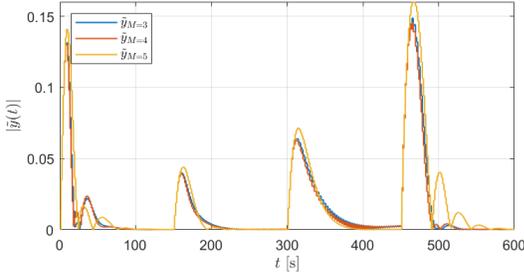


Fig. 7. Example 2: Tracking error of the the data-driven IMC (18) using a second-order DVS of dimension  $M = 3$ ,  $M = 4$  and  $M = 5$ , respectively.

it follows that

$$\begin{aligned} \text{rank}(\mathcal{M}_{ob,\{0,L,T-L+1\}}) &= \\ \text{rank} \left( \begin{bmatrix} \mathcal{M}_{ob,\{0,L,T-L+1\}} & \begin{bmatrix} \mu_{[0,L]} \\ \mu_{[0,L]}^2 \end{bmatrix} \end{bmatrix} \right) &= \\ = \frac{1}{2}L(M+1)(M+4). & \end{aligned} \quad (23)$$

Since the rank does not change when the sequences of  $\mu_{[0,L]}$  and  $\mu_{[0,L]}^2$  are incorporated into  $\mathcal{M}_{ob,\{0,1,T\}}$ , there exists a  $g \in \mathbb{R}^{T-L+1}$  such that

$$\begin{bmatrix} \mu_{[0,L]} \\ \mu_{[0,L]}^2 \end{bmatrix} = \mathcal{M}_{ob,\{0,L,T-L+1\}} g. \quad (24)$$

By replacing (24) in (22), one obtains

$$\begin{bmatrix} \mu_{[0,L]} \\ \mu_{[0,L]}^2 \\ y_{[0,L]} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{\sigma_1} & 0_{\sigma_1 \times \sigma_2} \\ 0_{\sigma_2 \times \sigma_1} & \mathbf{I}_{\sigma_2} \\ \mathbf{I}_L \otimes \theta_1^\top & \mathbf{I}_L \otimes \theta_2^\top \end{bmatrix} \mathcal{M}_{ob,\{0,L,T-L+1\}} g \quad (25)$$

$$= \begin{bmatrix} M_{ob,\{0,L,T-L+1\}} \\ M_{ob,\{0,L,T-L+1\}}^2 \\ Y_{ob,\{0,L,T-L+1\}} \end{bmatrix} g. \quad (26)$$

In this manner, the result of the first part is recovered.

The necessity stems from the fact that if  $\mathcal{M}_{ob,\{0,L,T-L+1\}}$  does not have full-row rank, the existence of  $g$  cannot be ensured in general since the rank of (23) may change.

**Part ii)** From (22), one has

$$\begin{bmatrix} \mu_{[0,L]} \\ \mu_{[0,L]}^2 \\ y_{[0,L]} \end{bmatrix} = \begin{bmatrix} M_{ob,\{0,L,T-L+1\}} \\ M_{ob,\{0,L,T-L+1\}}^2 \\ Y_{ob,\{0,L,T-L+1\}} \end{bmatrix} g \quad (27)$$

$$= \begin{bmatrix} \mathbf{I}_{\sigma_1} & 0_{\sigma_1 \times \sigma_2} \\ 0_{\sigma_2 \times \sigma_1} & \mathbf{I}_{\sigma_2} \\ \mathbf{I}_L \otimes \theta_1^\top & \mathbf{I}_L \otimes \theta_2^\top \end{bmatrix} \mathcal{M}_{ob,\{0,L,T-L+1\}} g. \quad (28)$$

By direct substitution of  $g$  as defined in (11), it follows that

$$\begin{bmatrix} \mu_{[0,L]} \\ \mu_{[0,L]}^2 \\ y_{[0,L]} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{\sigma_1} & 0_{\sigma_1 \times \sigma_2} \\ 0_{\sigma_2 \times \sigma_1} & \mathbf{I}_{\sigma_2} \\ \mathbf{I}_L \otimes \theta_1^\top & \mathbf{I}_L \otimes \theta_2^\top \end{bmatrix} \begin{bmatrix} \nu_{[0,L-1]} \\ \nu_{[0,L-1]}^2 \end{bmatrix}, \quad (29)$$

thus, corroborating that the trajectory (10) corresponds to the system (4). The necessity follows from the fact that if  $\mathcal{M}_{ob,\{0,L,T-L+1\}}$  does not have full-row rank, then

$\mathcal{M}_{ob,\{0,L,T-L+1\}} \mathcal{M}_{ob,\{0,L,T-L+1\}}^\dagger \neq \mathbf{I}_{L(\sigma_1+\sigma_2)}$ . Therefore, in general, the trajectory (10) will no correspond to the system (4).

### B. Proof of Corollary 1

If the observed input sequence  $u_{ob,[-M,T]}$  is persistently exciting of order one with respect to the system (4),  $\mathcal{M}_{ob,\{0,1,T\}}$  has full-row rank. Then Theorem 1 part i) implies that there exists a  $g \in \mathbb{R}^T$ , such that

$$y(k) = Y_{ob,\{0,1,T\}} g.$$

The proof is completed by choosing  $g$  as in Theorem 1 part ii) with  $\nu_{[0,1]} = \mu(k)$ , which yields (12).

### C. Proof of Corollary 2

Define  $\sigma_1 = M + 1$  and  $\sigma_2 = \frac{1}{2}(M + 1)(M + 2)$ . One has that

$$\begin{aligned} \mathcal{M}_{ob,\{0,1,T\}} \mathcal{M}_{ob,\{0,1,T\}}^\dagger &= \begin{bmatrix} M_{ob,\{0,1,T\}} \\ M_{ob,\{0,1,T\}}^2 \end{bmatrix} \mathcal{M}_{ob,\{0,1,T\}}^\dagger \\ &= \mathbf{I}_{\sigma_1+\sigma_2}. \end{aligned} \quad (30)$$

Let  $\mathcal{M}_{ob,\{0,1,T\}}^\dagger$  be split in two matrices  $D_1$  and  $D_2$ , i.e.,  $\mathcal{M}_{ob,\{0,1,T\}}^\dagger = [D_1 \ D_2]$ . Then, (30) can be written equivalently as

$$\begin{aligned} \begin{bmatrix} M_{ob,\{0,1,T\}} \\ M_{ob,\{0,1,T\}}^2 \end{bmatrix} [D_1 \ D_2] &= \\ &= \begin{bmatrix} M_{ob,\{0,1,T\}} D_1 & M_{ob,\{0,1,T\}} D_2 \\ M_{ob,\{0,1,T\}}^2 D_1 & M_{ob,\{0,1,T\}}^2 D_2 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I}_{\sigma_1} & 0_{\sigma_1 \times \sigma_2} \\ 0_{\sigma_2 \times \sigma_1} & \mathbf{I}_{\sigma_2} \end{bmatrix}. \end{aligned}$$

Solving for  $D_1$  and  $D_2$  yields (21).

By substituting the relation  $\mathcal{M}_{ob,\{0,1,T\}}^\dagger = [D_1 \ D_2]$  in (12), we obtain

$$y(k) = Y_{ob,\{0,1,T\}} D_1 \mu(k) + Y_{ob,\{0,1,T\}} D_2 \mu^2(k).$$

From the last expression,  $P_1$  and  $P_2$  in (14) follow, completing the proof.

### D. Proof of Lemma 1

From (4) it follows that

$$y_1(k) = \theta_{1,1} u(k) + \theta_{1,2} u(k-1) + \dots + \theta_{1,M+1} u(k-M),$$

where  $\theta_{1,i}$  denotes the  $i$ -th element of  $\theta_1$ . In order to represent the input  $u(k)$  as a function of the output  $y(k)$ , the relation above can be reorganized as

$$u(k) = \frac{1}{\theta_{1,1}} y_1(k) - \frac{\theta_{1,2}}{\theta_{1,1}} u(k-1) \dots - \frac{\theta_{1,M+1}}{\theta_{1,1}} u(k-M).$$

Consider the following short-hands:

$$\bar{\theta} = \begin{bmatrix} -\frac{\theta_{1,2}}{\theta_{1,1}} & -\frac{\theta_{1,3}}{\theta_{1,1}} & \dots & -\frac{\theta_{1,M+1}}{\theta_{1,1}} \end{bmatrix}^\top, \quad d = \frac{1}{\theta_{1,1}}.$$

Hence, the inverse system is described by the linear system

$$\begin{aligned} \chi(k) &= [u(k-1) \ u(k-2) \ \dots \ u(k-M)]^\top, \\ u(k) &= \bar{\theta}^\top \chi(k) + d y_1(k). \end{aligned} \quad (31)$$

$$D_1 = \left( I_T - (M_{ob,\{0,1,t\}}^2)^\dagger M_{ob,\{0,1,T\}}^2 \right) M_{ob,\{0,1,T\}}^\dagger \left( I_{M+1} - M_{ob,\{0,1,T\}} (M_{ob,\{0,1,T\}}^2)^\dagger M_{ob,\{0,1,T\}}^2 M_{ob,\{0,1,T\}}^\dagger \right)^{-1}, \quad (21)$$

$$D_2 = \left( I_T - M_{ob,\{0,1,t\}}^\dagger M_{ob,\{0,1,T\}} \right) (M_{ob,\{0,1,T\}}^2)^\dagger \left( I_{\frac{M+1}{2}(M+2)} - M_{ob,\{0,1,T\}}^2 M_{ob,\{0,1,T\}}^\dagger M_{ob,\{0,1,T\}} (M_{ob,\{0,1,T\}}^2)^\dagger \right)^{-1}.$$

It follows that any  $L$ -long input-output sequence of (31) can be represented as

$$\begin{bmatrix} y_{1,[0,L]} \\ u_{[0,L]} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_L & 0_{L \times M} \\ d\mathbf{I}_L & \mathbf{I}_L \otimes \bar{\theta}^\top \end{bmatrix} \begin{bmatrix} y_{1,[0,L]} \\ \chi_{[0,L]} \end{bmatrix}. \quad (32)$$

Since  $y_{1ob,[0,T]}$  is persistently exciting, there exists  $g \in \mathbb{R}^T$  such that

$$\begin{bmatrix} y_{1,[0,L]} \\ \chi_{[0,L]} \end{bmatrix} = \begin{bmatrix} Y_{1ob,\{0,L,T-L+1\}} \\ X_{ob,\{0,L,T-L+1\}} \end{bmatrix} g. \quad (33)$$

A particular  $g$  can be found using the right inverse. By replacing (33) and (34) in (32), we obtain

$$\begin{bmatrix} y_{1,[0,L]} \\ u_{[0,L]} \end{bmatrix} = \begin{bmatrix} Y_{1ob,\{0,L,T-L+1\}} \\ X_{ob,\{0,L,T-L+1\}} \end{bmatrix} \begin{bmatrix} Y_{1ob,\{0,L,T-L+1\}} \\ X_{ob,\{0,L,T-L+1\}} \end{bmatrix}^\dagger \begin{bmatrix} y_{1,[0,L]} \\ \chi_{[0,L]} \end{bmatrix}.$$

The expression (17) follows by setting  $L = 1$  above.

#### REFERENCES

- [1] Z.-S. Hou and Z. Wang, "From model-based control to data-driven control: Survey, classification and perspective," *Information Sciences*, vol. 235, pp. 3 – 35, 2013.
- [2] S. Yin, S. X. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 11, pp. 6418–6428, 2014.
- [3] F. Lamnabhi-Lagarrigue, A. Annaswamy, S. Engell, A. Isaksson, P. Khargonekar, R. M. Murray, H. Nijmeijer, T. Samad, D. Tilbury, and P. V. den Hof, "Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges," *Annual Reviews in Control*, vol. 43, pp. 1 – 64, 2017.
- [4] A. Bazanella, L. Campestri, and E. D., *Data-Driven Controller Design - The H<sub>2</sub> Approach*. Dordrecht, The Netherlands: Springer Netherlands, 2012.
- [5] I. Markovsky and P. Rapisarda, "On the linear quadratic data-driven control," in *2007 European Control Conference (ECC)*, 2007, pp. 5313–5318.
- [6] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," *International Journal of Control*, vol. 81, no. 12, pp. 1946–1959, 2008.
- [7] U. S. Park and M. Ikeda, "Stability analysis and control design of LTI discrete-time systems by the direct use of time series data," *Automatica*, vol. 45, no. 5, pp. 1265 – 1271, 2009.
- [8] B. Pang, T. Bian, and Z. Jiang, "Data-driven finite-horizon optimal control for linear time-varying discrete-time systems," in *2018 IEEE Conference on Decision and Control (CDC)*, 2018, pp. 861–866.
- [9] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality and robustness," *IEEE Transactions on Automatic Control*, 2019.
- [10] S. Tu and B. Recht, "The gap between model-based and model-free methods on the linear quadratic regulator: An asymptotic viewpoint," *arXiv preprint arXiv:1812.03565*, 2018. [Online]. Available: arxiv.org/abs/1812.03565
- [11] M. Rotulo, C. De Persis, and P. Tesi, "Data-driven linear quadratic regulation via semidefinite programming," *arXiv preprint arXiv:1911.07767*, 2019. [Online]. Available: arxiv.org/abs/1911.07767
- [12] J. Berberich, A. Romer, C. W. Scherer, and F. Allgöwer, "Robust data-driven state-feedback design," *arXiv preprint arXiv:1909.04314*, 2019. [Online]. Available: arxiv.org/abs/1909.04314
- [13] S. Dean, S. Tu, N. Matni, and B. Recht, "Safely learning to control the constrained linear quadratic regulator," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 5582–5588.
- [14] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the DeePC," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 307–312.
- [15] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *arXiv preprint arXiv:1906.04679*, 2019. [Online]. Available: arxiv.org/abs/1906.04679
- [16] A. Janczak, *Identification of Nonlinear Systems Using Neural Networks and Polynomial Models: A Block-Oriented Approach*, ser. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, 2004.
- [17] T. Ogunfunmi, *Adaptive Nonlinear System Identification: The Volterra and Wiener Model Approaches*, ser. Signals and Communication Technology. Springer US, 2007.
- [18] C. Schmidt, S. Biagiola, J. Cousseau, and J. Figueroa, "Volterra-type models for nonlinear systems identification," *Applied Mathematical Modelling*, vol. 38, no. 9, pp. 2414 – 2421, 2014.
- [19] F. Doyle, R. Pearson, and B. Ogunnaile, *Identification and control using Volterra models*. London: Springer, 2002.
- [20] G. Govind and P. Ramamoorthy, "Multi-layer neural networks and Volterra series: The missing," in *IEEE International Conference on Systems Engineering*, 1990.
- [21] G. Davis and L. Gasperi, "ANN modeling of Volterra systems," in *Seattle International Joint Conference on Neural Networks*, 1991.
- [22] V. Z. Marmarelis and X. Zhao, "Volterra models and three-layer perceptrons," *IEEE Transactions on Neural Networks*, vol. 8, no. 6, pp. 1421–1433, 1997.
- [23] A. Janczak, *Identification of Nonlinear Systems Using Neural Networks and Polynomial Models: A Block-Oriented Approach*, ser. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, 2009.
- [24] B. R. Maner, F. J. Doyle, B. A. Ogunnaik, and R. K. Pearson, "Nonlinear model predictive control of a simulated multivariable polymerization reactor using second-order volterra models," *Automatica*, vol. 32, no. 9, pp. 1285 – 1301, 1996.
- [25] W.-M. Ling and D. E. Rivera, "Control relevant model reduction of volterra series models," *Journal of Process Control*, vol. 8, no. 2, pp. 79 – 88, 1998.
- [26] A. Vargas and F. Allgöwer, "Model reduction for process control using iterative nonlinear identification," in *American Control Conference*, 2004.
- [27] C. Medina-Ramos and H. Nieto-Chaupis, "Testing Laguerre-Volterra model in a batch reactor," in *IEEE International Conference on Industrial Technology*, 2010.
- [28] D. Antunes, J. Hespanha, and C. Silvestre, "Volterra integral approach to impulsive renewal systems: Application to networked control," *IEEE Transactions on Automatic Control*, 2012.
- [29] J. Berberich and F. Allgöwer, "A trajectory-based framework for data-driven system analysis and control," *arXiv preprint arXiv:1903.10723*, 2019. [Online]. Available: arxiv.org/abs/1903.10723
- [30] C. Nikias and A. Petropulu, *Higher-order Spectra Analysis: A Nonlinear Signal Processing Framework*, ser. Prentice Hall signal processing series. PTR Prentice Hall, 1993.
- [31] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "On amplitude and frequency demodulation using energy operators," *IEEE Transactions on Signal Processing*, vol. 41, no. 4, pp. 1532–1550, 1993.
- [32] J. Willems, P. Rapisarda, I. Markovsky, and B. D. Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325 – 329, 2005.
- [33] F. J. Doyle III, B. A. Ogunnaik, and R. K. Pearson, "Nonlinear model-based control using second-order volterra models," *Automatica*, vol. 31, no. 5, pp. 697–714, 1995.
- [34] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*. New Jersey, USA: Princeton University Press, 2009.
- [35] W. Rugh, *Nonlinear system theory: The Volterra/Winner approach*. Baltimore: The Johns Hopkins University Press, 2002/1981.