A Data-Driven Formulation of the Maximal Admissible Set and the Data-Enabled Reference Governor

Hamid R. Ossareh, *Senior Member, IEEE*

***Abstract*—The maximal admissible set (MAS) of a stable LTI system characterizes the set of all initial conditions and constant inputs for which the output satisfies pre-specified state/output constraints for all time. The MAS (or its finitely-determined, polytopic approximations) is often employed in set-theoretic methods in control and for constraint management, for example in the Reference Governor (RG) algorithm. The existing MAS (and consequently RG) formulations require a state-space model of the dynamics to characterize the MAS. In this letter, we offer an alternative, data-driven perspective: we leverage output predictors from the behavioral system theory and subspace predictive control literature to formulate a data-driven version of the MAS. As we show, the proposed set is polytopic and has finite complexity, similar to its model-based counterpart, but resides in a higher dimensional space and may have higher complexity. We present the properties of the data-driven MAS including its admissibility index, and compare the data-driven MAS against its model-based formulation, where we show that the two sets are related via a linear map under mild assumptions. Finally, we use the data-driven MAS to introduce a data-enabled RG for constraint management of closed-loop control systems. Numerical simulations are presented to illustrate the results.**

***Index Terms*—Maximal admissible sets, reference governors, data-driven control, behavioral system theory.**

I. Introduction

The maximal admissible set (MAS)[1] of a dynamical system is a positively-invariant set that characterizes the set of all initial conditions and constant inputs for which the response satisfies pre-specified state/output constraints for all time [1], [2]. For discrete-time linear time-invariant systems subject to polytopic constraints, the MAS is polytopic, though it may be of infinite complexity. The MAS (or its finite-complexity, polytopic approximations) has been broadly employed in the control literature, for example, as a terminal constraint in the Model Predictive Control (MPC) optimization problem to guarantee closed-loop stability [3], [4], or in the analysis of constrained systems and in set-theoretic methods in control, see, e.g., [5], [6]. It has also been employed in constraint management algorithms such as the Reference Governor (RG) to guarantee infinite-horizon constraint satisfaction [7], [8].

The existing MAS formulations require a parametric (i.e., state-space) model of the dynamics to form output predictions and, thus, characterize the MAS. Motivated by the recent trends in decision making for complex systems and in data-driven control [9], this letter introduces an alternative, *data-driven* formulation of the MAS. Specifically, we take a representation-free and non-parametric perspective, and define the MAS using data-driven predictors that map offline input-output data directly onto online output predictions. Our approach is grounded in Willem's fundamental lemma [10] and is inspired by the literature on behavioral system theory [11], subspace methods [12], [13], and recent data-driven control algorithms [14], [15]. As we show, similar to the model-based MAS, the proposed data-driven MAS is positively invariant and can be characterized by a polytope with finite complexity. However, the data-driven MAS exists in a higher dimensional space, may be of higher complexity, and may not be compact. In this letter, we will assume that the offline input/output data is noise-free, reserving the consideration of the noisy case for future work.

Using this data-driven MAS, this letter then formulates a data-enabled RG. Block diagrams of a closed-loop system controlled by the traditional RG and the proposed data-enabled RG are depicted in Fig. 1. In both cases, the RG modifies the reference signal to the (pre-stabilized) closed-loop control system to enforce pointwise-in-time constraints and, thus, decouples the problem of tracking/stabilization and constraint management. However, in contrast to the model-based RG, the proposed data-enabled RG uses the data-driven version of the MAS and, as such, it bypasses the system identification step and does not require an observer. For this reason, it is simpler to implement and provides an end-to-end solution to the practitioner. The disadvantage as compared to the model-based RG is that it may have a higher computational complexity, requires access to clean data to construct the MAS, and can only be applied to constrain the measured output, not the internal states. We present numerical simulations to illustrate the results.

The notation throughout this letter is as follows. The sets $\mathbb{Z}^+$, $\mathbb{R}$, $\mathbb{R}^n$, and $\mathbb{R}^{n\times m}$ denote the set of non-negative integers, real numbers, n-dimensional vectors of real numbers, and $n\times m$ matrices with real entries, respectively. For a set A, the interior is denoted by $\mathrm{Int}(A)$. For a sequence of matrices $X_1, \dots, X_n$ with the same number of columns, we denote $[X_1^\top, \dots, X_n^\top]^\top$

Manuscript received 15 September 2023; accepted 23 October 2023. Date of publication 6 November 2023; date of current version 1 December 2023. This work was supported in part by NIST under Grant 70NANB21H133; in part by NASA under Grant VT-80NSSC20M0213; and in part by NSF under Grant CMMI-2238424. Recommended by Senior Editor S. Olaru.

The author is with the Electrical Engineering Program, University of Vermont, Burlington, VT 05405 USA (e-mail: hossareh@uvm.edu).

Digital Object Identifier 10.1109/LCSYS.2023.3330386

[1]Also called the maximal output admissible set or maximal positively invariant set.

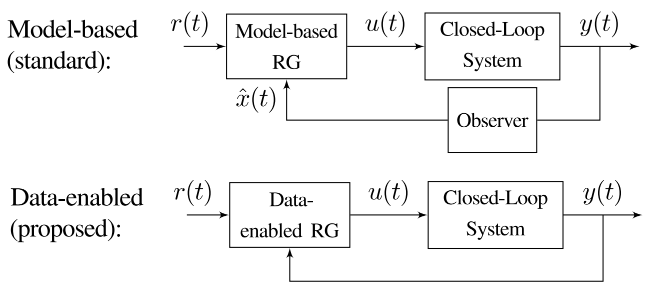

Fig. 1. Model-based and data-enabled Reference Governor (RG) block diagrams.

by $\mathbf{col}(X_1, \ldots, X_n)$. The vectors $\mathbf{1}_T$ and $\mathbf{0}_T$ denote vectors of all ones and all zeros of length T, respectively, and we drop the subscript if the dimensionality can be inferred from the context. The matrix $\mathbf{0}_{n\times m}$ is the n by m zero matrix. For a signal $u(t) \in \mathbb{R}^m$, $t = 0, \ldots, T-1$, we use boldface to denote the signal as a vector over the horizon T, i.e., $\mathbf{u} = \mathbf{col}(u(0), u(1), \ldots, u(T-1))$.

II. Preliminaries and Problem Statement

Consider the top block diagram shown in Fig. 1, in which the "closed-loop system" is described by the LTI model:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{1}$$

where $t \in \mathbb{Z}^+$ is the discrete time index, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, and $y(t) \in \mathbb{R}^p$. In typical RG applications, $u(t)$ represents the reference command. The output, $y(t)$, is to be constrained as follows:

$$y(t) \in \mathbb{Y} \triangleq \{y{:}Sy \leq s\}, \tag{2}$$

where $S \in \mathbb{R}^{q\times p}$ and $s \in \mathbb{R}^q$ are given. Here, q is the number of inequalities describing (2). In this letter, we assume that:

Assumption 1: System (1) is asymptotically stable (A is Schur), controllable, and observable. Furthermore, the constraint set is compact and satisfies $0 \in \mathrm{Int}(\mathbb{Y})$.

We first introduce the "maximal admissible set" (MAS), denoted by O_∞, which is the set of all states and constant control inputs that satisfy (2) for all time:

$$O_\infty = \{(x_0, u_0) \in \mathbb{R}^n \times \mathbb{R}^m{:}\quad y(t) \in \mathbb{Y}, \quad \forall t \in \mathbb{Z}^+\}, \tag{3}$$

where $y(t) = CA^t x_0 + (C(I-A^t)(I-A)^{-1}B + D)u_0$. This set is generally not finitely determined, i.e., it cannot be described by a finite number of inequalities. However, it is shown in [1] that a finitely-determined inner approximation, denoted by $\widetilde{O}_\infty$ and with similar invariance properties, can be obtained by imposing a tightened version of the constraint on the steady-state output, that is:

$$\begin{aligned} \widetilde{O}_\infty = \{(x_0, u_0){:}\quad & y(\infty) \in (1-\epsilon)\mathbb{Y}, \quad y(t) \in \mathbb{Y}, \\ & t = 0, \ldots, t^*\}, \end{aligned} \tag{4}$$

where $\epsilon \in (0, 1)$ is a small number and $y(\infty) := P_0 u_0$, with P_0 being the DC gain matrix, i.e., $P_0 = C(I-A)^{-1}B + D$. The *admissibility index*, t^*, is the smallest prediction horizon to fully characterize $\widetilde{O}_\infty$, i.e., the inequalities corresponding to $y(t) \in \mathbb{Y}$ are redundant for $t > t^*$. The set $\widetilde{O}_\infty$ is thus a polytope of the form:

$$\widetilde{O}_\infty = \{(x_0, u_0){:}\ SP_0u_0 \leq (1-\epsilon)s, \quad H_x x_0 + H_u u_0 \leq h\}, \tag{5}$$

where the i-th block rows of H_x, H_u, and h are, respectively, given by SCA^{i-1}, $SC(I-A^{i-1})(I-A)^{-1}B + SD$, and s.

The RG leverages $\widetilde{O}_\infty$ to find an optimal, constraint-admissible value of u at every timestep. The update law of the RG is as follows:

$$u(t) = u(t-1) + \kappa(r(t) - u(t-1)), \tag{6}$$

where $\kappa \in [0, 1]$ is obtained by solving the linear program (LP)

$$\begin{aligned} \underset{\kappa\in[0,1]}{\text{maximize}} \quad & \kappa \\ \text{s.t.} \quad & u_0 = u(t-1) + \kappa(r(t) - u(t-1)) \\ & x_0 = x(t) \\ & (x_0, u_0) \in \widetilde{O}_\infty \end{aligned} \tag{7}$$

If the reference $r(t)$ is constraint-admissible, then $\kappa = 1$ and, therefore, $u(t) = r(t)$. Otherwise, $u(t)$ will be closer to $u(t-1)$. Positive invariance of $\widetilde{O}_\infty$ implies that $\kappa = 0$ is always a feasible solution of the above LP, as long as it is feasible at the initial time.

Problem Statement: Consider now the bottom block diagram shown in Fig. 1, where the closed-loop system is again described by (1). Suppose that the matrices A, B, C, and D are unavailable for the calculation of $\widetilde{O}_\infty$ or for the purpose of RG design and implementation. However, suppose that M length-T input-output, noise-free trajectories are available. Let the i-th input and output trajectory be denoted by $u_i^d(t)$ and $y_i^d(t)$, where $t = 0, \ldots, T-1$; $i = 1, \ldots, M$. The superscript d refers to "data". In general, the M trajectories may come from independent experiments. However, they can also be different sections of the same, long trajectory, $u^d(t), y^d(t)$. In this situation, $u_i^d(t)$ and $y_i^d(t)$ could be chosen as: $u_k^d(i) = u^d(k+i)$ and $y_k^d(i) = y^d(k+i)$ for all $i \in \{0, \ldots, T-1\}$ and $k \in \{1, \ldots, M\}$.

The problem addressed in this letter is to characterize $\widetilde{O}_\infty$ using the input-output trajectories described above without identifying a parametric model of the system, and to use this data-driven $\widetilde{O}_\infty$ to formulate a data-enabled RG.

III. A Data-Driven Formulation of the Maximal Admissible Set (MAS)

Eq. (4) shows that, to develop a data-driven version of $\widetilde{O}_\infty$, two key components are required: the output predictions, $y(t)$, and the steady-state output, $y(\infty)$. In this section, we present mappings from the offline data onto $y(t)$ and $y(\infty)$, and formally introduce a data-driven version of the MAS, along with its associated properties.

To begin, define the extended observability and convolution matrices:

$$O_t = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{t-1} \end{bmatrix}, \quad C_t = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \vdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ CA^{t-2}B & \cdots & CB & D \end{bmatrix} \tag{8}$$

Let the integer L be any positive time horizon. Starting from any initial condition, x_0, and input sequence $\mathbf{u} \in \mathbb{R}^{mL}$ over the horizon L, the output of (1), $\mathbf{y} \in \mathbb{R}^{pL}$, over the same horizon is given by:

$$\mathbf{y} = O_L x_0 + C_L \mathbf{u} \tag{9}$$

Here, boldface denotes the signal as a vector, consistent with the notation in the Introduction. We introduce the following definitions.

Definition 1: The lag (or observability index) of system (1), denoted by ℓ, is defined as the smallest integer such that O_ℓ has rank n (i.e., full column rank), where n is the system order.

Remark 1: ℓ exists per Assumption 1.

Definition 2: For any positive integer L, we define the "restricted behavior", $\mathscr{B}_L$, of system (1) as

$$\mathscr{B}_L = \{\mathbf{col}(\mathbf{u}, \mathbf{y}) \in \mathbb{R}^{(m+p)L} : \exists\ x_0 \in \mathbb{R}^n \text{ s.t. (9) holds}\}$$

Any length-L input-output sequence that satisfies $\mathbf{col}(\mathbf{u}, \mathbf{y}) \in \mathscr{B}_L$ is an "admissible trajectory" of system (1).

Remark 2: Suppose $\mathbf{col}(\mathbf{u}, \mathbf{y}) \in \mathscr{B}_L$ for some $L > 0$. The initial condition x_0 can be **uniquely** determined from (9) if and only if $L \geq \ell$. Proof of this fact follows from the definition of ℓ.

Now, consider the M length-T offline input-output trajectories introduced in the "Problem Statement" in Section II. Per the literature of data-driven prediction in the Behavioral setting (see [11] for an overview), we construct the "data matrix", $\mathcal{H}$, whose columns consist of the input and output trajectories stacked on top of one another:

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}_u \\ \mathcal{H}_y \end{bmatrix} \in \mathbb{R}^{T(m+p)\times M}$$

where

$$\mathcal{H}_u = \begin{bmatrix} u_1^d(0) & u_2^d(0) & \cdots & u_M^d(0) \\ \vdots & \vdots & \ddots & \vdots \\ u_1^d(T-1) & u_2^d(T-1) & \cdots & u_M^d(T-1) \end{bmatrix}$$

and $\mathcal{H}_y$ can be similarly defined by replacing all instances of u by y in the above.

If the data matrix satisfies the so-called "low-rank condition", also called the "generalized persistency of excitation condition" [16]:

$$\text{rank}(\mathcal{H}) = mT + n, \tag{10}$$

for $T \geq \ell$, then any length-T admissible trajectory $\mathbf{col}(\mathbf{u}, \mathbf{y}) \in \mathbb{R}^{(m+p)T}$ will belong to the column space of $\mathcal{H}$. Said differently, there must exist a (non-unique) vector g, such that

$$\mathcal{H}g = \begin{bmatrix} \mathbf{u} \\ \mathbf{y} \end{bmatrix} \tag{11}$$

This idea has been used in the literature for the purpose of data-driven simulation and prediction. This is typically done by partitioning the output trajectory, $\mathbf{y}$, into two parts, one with a horizon length of T_{ini} that serves to implicitly fix the initial condition, and another of horizon length T_{pred} that serves as the predicted output: $\mathbf{y} = \mathbf{col}(\mathbf{y_{ini}}, \mathbf{y_{pred}})$, where $\mathbf{y_{ini}} \in \mathbb{R}^{pT_{ini}}$ and $\mathbf{y_{pred}} \in \mathbb{R}^{pT_{pred}}$, and $T_{ini} + T_{pred} = T$. We will discuss how to select T_{ini} and T_{pred} later. Similarly, partition the input as $\mathbf{u} = \mathbf{col}(\mathbf{u_{ini}}, \mathbf{u_{pred}})$, and accordingly the data matrix as $\mathcal{H}_u = \mathbf{col}(U_p, U_f)$ and $\mathcal{H}_y = \mathbf{col}(Y_p,\ Y_f)$. We can then express (11) as:

$$\begin{bmatrix} U_p \\ U_f \\ Y_p \\ Y_f \end{bmatrix} g = \begin{bmatrix} \mathbf{u_{ini}} \\ \mathbf{u_{pred}} \\ \mathbf{y_{ini}} \\ \mathbf{y_{pred}} \end{bmatrix} \tag{12}$$

As discussed in Remark 2, if $\mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}}) \in \mathscr{B}_{T_{ini}}$ and $T_{ini} \geq \ell$, one can uniquely solve for the latent initial condition and, thus, uniquely compute $\mathbf{y_{pred}}$. That being said, the solution for g is generally not unique. To solve for $\mathbf{y_{pred}}$, the minimum-norm solution to g is often chosen, which results in:

$$\mathbf{y_{pred}} = Y_f \begin{bmatrix} U_p \\ U_f \\ Y_p \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{u_{ini}} \\ \mathbf{u_{pred}} \\ \mathbf{y_{ini}} \end{bmatrix}, \tag{13}$$

where † is the Moore–Penrose pseudoinverse. These ideas will now be leveraged to define and construct a data-driven formulation of $\widetilde{O}_\infty$ defined in (4). As mentioned earlier, the first component needed for doing so is a predictor for $y(t)$, $t = 0, \ldots, t^*$, directly from data. To this end, we choose $T_{pred} > t^*$ and, consistent with the definition of $\widetilde{O}_\infty$, let $u(t)$ be a constant over the horizon, that is $u(t) = u_0$, $t = 0, \ldots, T_{pred} - 1$. We then form $\mathbf{u_{pred}}$, a block vector with blocks equal to u_0, i.e., $\mathbf{u_{pred}} = \mathbf{1}_{T_{pred}} \otimes u_0$, where $\mathbf{1}_{T_{pred}} \in \mathbb{R}^{T_{pred}}$ is a vector of all ones and $\otimes$ is the Kronecker product. We then leverage Eq. (13) to find $\mathbf{y_{pred}}$:

$$\mathbf{y_{pred}} = \underbrace{Y_f \begin{bmatrix} U_p \\ U_f \\ Y_p \end{bmatrix}^{\dagger}}_{\mathcal{F}} \begin{bmatrix} \mathbf{u_{ini}} \\ \mathbf{1}_{T_{pred}} \otimes u_0 \\ \mathbf{y_{ini}} \end{bmatrix} \tag{14}$$

Now, to simplify (14), we define three matrices $\mathcal{F}_u$, $\mathcal{F}_y$, and $\mathcal{F}_0$ as follows: $\mathcal{F}_u$ is the first mT_{ini} columns of $\mathcal{F}$, $\mathcal{F}_y$ is the last pT_{ini} columns of $\mathcal{F}$, and $\mathcal{F}_0$ is a matrix whose i-th column is the sum of T_{pred} columns of $\mathcal{F}$ starting from column i and skipping every m columns, that is:

$$\mathcal{F}_0 = \begin{bmatrix} \sum_{i=0}^{T_{pred}-1} \mathcal{F}^{mi+1} & \ldots & \sum_{i=0}^{T_{pred}-1} \mathcal{F}^{mi+m} \end{bmatrix},$$

where $\mathcal{F}^j$ denotes the j-th column of $\mathcal{F}$. With this notation, (14) can be conveniently expressed as:

$$\mathbf{y_{pred}} = \mathcal{F}_u \mathbf{u_{ini}} + \mathcal{F}_y \mathbf{y_{ini}} + \mathcal{F}_0 u_0 \tag{15}$$

This equation provides predictions of the output, $y(t)$ from $t = 0$ to $t = T_{pred} - 1 \geq t^*$, using the offline data (encoded in the matrices $\mathcal{F}_u$, $\mathcal{F}_y$, and $\mathcal{F}_0$), data from the recent T_{ini} samples of the input and output (encoded in $\mathbf{u_{ini}}$ and $\mathbf{y_{ini}}$), and the constant input u_0, without the use of a parametric model. Finally, constraint (2) can be recast as a constraint on $\mathbf{y_{pred}}$: $\widetilde{S}\mathbf{y_{pred}} \leq \mathbf{1}_{T_{pred}} \otimes s$, where $\widetilde{S} \in \mathbb{R}^{(qT_{pred})\times(pT_{pred})}$ is a block diagonal matrix with S on all the block diagonals. Replacing $\mathbf{y_{pred}}$ from (15) leads to

$$\widetilde{S}\mathcal{F}_u \mathbf{u_{ini}} + \widetilde{S}\mathcal{F}_y \mathbf{y_{ini}} + \widetilde{S}\mathcal{F}_0 u_0 \leq \mathbf{1}_{T_{pred}} \otimes s \tag{16}$$

This inequality will be used to define the data-driven MAS.

As mentioned previously, the second component required to define a data-driven MAS is the steady-state output, i.e., $y(\infty) = P_0 u_0$, where $P_0 \in \mathbb{R}^{p\times m}$ is the DC gain matrix. To find P_0 using data, we leverage the fact that if the input to an LTI system is constant and equal to e_i, where e_i is the i-th column of the identity matrix, then the *steady-state* output will be the i-th column of P_0, denoted by P_0^i. We can use this fact to compute the columns of P_0 one by one directly from data. Specifically, for each i, we set $u(t) = e_i$ and $y(t) = P_0^i$ for $t = 0, \ldots, T-1$ to enforce steady-state. We then form $\mathbf{u} = \mathbf{1}_T \otimes e_i$ and $\mathbf{y} = \mathbf{1}_T \otimes P_0^i$ and plug into Eq. (11) to obtain:

$$\begin{bmatrix} \mathcal{H}_u \\ \mathcal{H}_y \end{bmatrix} g = \begin{bmatrix} \mathbf{1}_T \otimes e_i \\ \mathbf{1}_T \otimes P_0^i \end{bmatrix} \Rightarrow \begin{bmatrix} \mathcal{H}_u \\ \mathcal{H}_y \end{bmatrix} G = \begin{bmatrix} \mathbf{1}_T \otimes I_m \\ \mathbf{1}_T \otimes P_0 \end{bmatrix}, \tag{17}$$

where we have combined the m different equations into one, with $G = [g_1 \ldots g_m]$. This equation can be equivalently expressed as:

$$\begin{bmatrix} \mathbf{0}_{mT\times p} & \mathcal{H}_u \\ -\mathbf{1}_T \otimes I_p & \mathcal{H}_y \end{bmatrix} \begin{bmatrix} P_0 \\ G \end{bmatrix} = \begin{bmatrix} \mathbf{1}_T \otimes I_m \\ \mathbf{0}_{pT\times m} \end{bmatrix} \tag{18}$$

This equation has a unique solution for P_0 (but not for G) under the rank condition for $\mathcal{H}$ and $T \geq \ell + 1$. Using the minimum-norm solution given by the pseudoinverse, P_0 can be readily obtained:

$$P_0 = \begin{bmatrix} I_p & \mathbf{0}_{p\times M}\end{bmatrix} \begin{bmatrix} \mathbf{0}_{mT\times p} & \mathcal{H}_u \\ -\mathbf{1}_T \otimes I_p & \mathcal{H}_y \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{1}_T \otimes I_m \\ \mathbf{0}_{pT\times m}\end{bmatrix} \quad (19)$$

The approach presented is similar to the one in [17] for computing the frequency response. Note that, by construction, P_0 given by (19) coincides with $C(I - A^{-1})B + D$ obtained using a model.

Remark 3: In some applications, the DC gain may be readily available (for example, when step response data has been collected). For these applications, the empirical P_0 can be used instead of the one calculated using Eq. (19) to avoid unnecessary computations.

We are now ready to explain why in this letter we present a data-driven version of $\widetilde{O}_\infty$ in (4), not of O_∞ in (3). Recall that O_∞ is, in general, not finitely determined. Thus, a data-driven version of O_∞ will necessitate $T_{pred} = \infty$, which calls for an infinite amount of offline data and an infinite-dimensional data matrix, neither of which is possible. Therefore, we instead formulate a data-driven version of $\widetilde{O}_\infty$, as introduced next.

Definition 3: Let $T_{ini} \geq \ell$ and $T_{pred} > t^*$. The data-driven version of $\widetilde{O}_\infty$, denoted by $\widetilde{O}^D_\infty$, is defined as the set of all recent T_{ini} inputs/outputs and future constant inputs such that $y(\infty) \in (1-\epsilon)\mathbb{Y}$ and $y(t) \in \mathbb{Y}, \quad t = 0, \ldots, T_{pred} - 1$, that is:

$$\begin{aligned} \widetilde{O}^D_\infty = \{(\mathbf{u_{ini}}, \mathbf{y_{ini}}, u_0) \in \mathbb{R}^{mT_{ini}} \times \mathbb{R}^{pT_{ini}} \times \mathbb{R}^m: \\ y(\infty) \in (1-\epsilon)\mathbb{Y}, \quad y(t) \in \mathbb{Y}, \\ t = 0, \ldots, T_{pred} - 1\} \end{aligned} \quad (20)$$

Using (16) and (19), this set can be equivalently characterized as the following polytope:

$$\begin{aligned} \widetilde{O}^D_\infty = \{(\mathbf{u_{ini}}, \mathbf{y_{ini}}, u_0) : SP_0 u_0 \leq (1-\epsilon)s, \\ \widetilde{S}\mathcal{F}_u \mathbf{u_{ini}} + \widetilde{S}\mathcal{F}_y \mathbf{y_{ini}} + \widetilde{S}\mathcal{F}_0 u_0 \leq \mathbf{1}_{T_{pred}} \otimes s\} \end{aligned} \quad (21)$$

Note that this set is not identical to $\widetilde{O}_\infty$. For one, it lives in a higher dimensional space. This is because, in the data-driven setting, $\mathbf{u_{ini}}$ and $\mathbf{y_{ini}}$ serve the role of the initial condition, instead of x_0. Less obviously, the $(\mathbf{u_{ini}}, \mathbf{y_{ini}})$ pair in (21) are not required to satisfy $\mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}}) \in \mathscr{B}_{T_{ini}}$, i.e., be admissible trajectories of (1) (enforcing such condition will not allow for a polytopic H-representation of $\widetilde{O}^D_\infty$, which is suitable for optimization-based control strategies; see Remark 4 for details). The relationship between $\widetilde{O}_\infty$ and $\widetilde{O}^D_\infty$ is further elaborated in the following theorem.

Theorem 1: Consider system (1). If $T_{ini} \geq \ell$, $T_{pred} > t^*$, and $\mathcal{H}$ satisfies (10) with $T = T_{ini} + T_{pred}$, then $\widetilde{O}_\infty$ and $\widetilde{O}^D_\infty$ are related as follows:

1) If $(\mathbf{u_{ini}}, \mathbf{y_{ini}}, u_0) \in \widetilde{O}^D_\infty$ and $\mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}}) \in \mathscr{B}_{T_{ini}}$, then there exists a unique x_0 such that $(x_0, u_0) \in \widetilde{O}_\infty$.
2) If $(x_0, u_0) \in \widetilde{O}_\infty$, then there exist (non-unique) $\mathbf{u_{ini}}$ and $\mathbf{y_{ini}}$ such that $\mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}}) \in \mathscr{B}_{T_{ini}}$ and $(\mathbf{u_{ini}}, \mathbf{y_{ini}}, u_0) \in \widetilde{O}^D_\infty$.
3) If $T_{pred} = t^* + 1$, then the vectors $\widetilde{S}\mathcal{F}_0$ and $\mathbf{1}_{T_{pred}} \otimes s$ in (21) are identical to H_u and h in (5).

Proof: Define the following matrices:

$$M_t = \begin{bmatrix} I_{tm} & \mathbf{0}_{tm\times n} \\ C_t & O_t \end{bmatrix}, \quad N_t = \begin{bmatrix} A^{t-1}B \cdots AB \; B \end{bmatrix}$$

We begin by proving part 1. First, note that P_0 obtained using (19) is the exact DC gain matrix. Thus, the constraint $SP_0u_0 \leq (1-\epsilon)s$ is common between the definitions of $\widetilde{O}_\infty$ and $\widetilde{O}^D_\infty$. Now, suppose $(\mathbf{u_{ini}}, \mathbf{y_{ini}}, u_0) \in \widetilde{O}^D_\infty$ and $\mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}}) \in \mathscr{B}_{T_{ini}}$. From (9) and Remark 2, there exists an initial condition $x_{ini} = x(-T_{ini})$ such that $\mathbf{y_{ini}} = O_{T_{ini}} x_{ini} + C_{T_{ini}} \mathbf{u_{ini}}$, which we can rewrite as:

$$\mathbf{col}(\mathbf{u_{ini}}, \; \mathbf{y_{ini}}) = M_{T_{ini}} \mathbf{col}(\mathbf{u_{ini}}, \; x_{ini}) \quad (22)$$

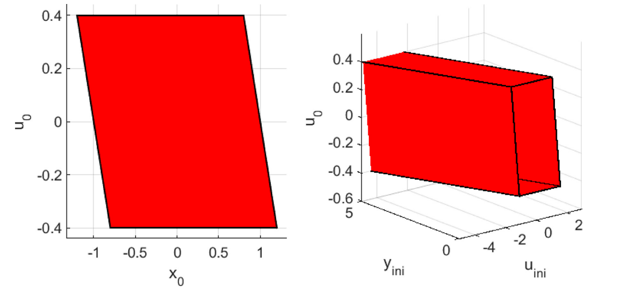

Fig. 2. $\widetilde{O}_\infty$ (left) and $\widetilde{O}^D_\infty$ (right). The latter set is unbounded, for example along the line defined by $u_0 = 0$ and $\mathbf{y_{ini}} = -1.5\mathbf{u_{ini}}$.

Since $T_{ini} \geq \ell$, $M_{T_{ini}}$ has full column rank, and so we can uniquely solve: $\mathbf{col}(\mathbf{u_{ini}}, x_{ini}) = M^{\dagger}_{T_{ini}} \mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}})$. Starting from initial condition x_{ini} at time $t = -T_{ini}$, one can solve for the system state at time $t = 0$: $x_0 = A^{T_{ini}} x_{ini} + N_{T_{ini}} \mathbf{u_{ini}}$, or

$$x_0 = \begin{bmatrix} N_{T_{ini}} & A^{T_{ini}} \end{bmatrix} M^{\dagger}_{T_{ini}} \mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}}), \quad (23)$$

which is uniquely defined. The output predictions, $y(t)$, starting from this x_0 and with constant input u_0 are the same as those obtained from (15). Since the data-driven predictions satisfy the constraints by our assumption of inclusion in $\widetilde{O}^D_\infty$, the model-based ones do too, and so (x_0, u_0) must belong to $\widetilde{O}_\infty$. This proves the first part.

To prove the second part, suppose $(x_0, u_0) \in \widetilde{O}_\infty$. Any arbitrary $\mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}}) \in \mathscr{B}_{T_{ini}}$ consistent with (23) results in the same initial state x_0 and the same predictions over the prediction horizon. This second statement thus follows.

To prove the third part, we note that the matrix H_u represents the step response of the system with zero initial conditions, and $\widetilde{S}\mathcal{F}_0$ represents the same. Furthermore, the two vectors have the same number of elements, and thus are identical. The equality of $\mathbf{1}_{T_{pred}} \otimes s$ and h can be shown by comparing them element-wise. ■

As seen from the proof of the theorem, if $T_{ini} \geq \ell$, $T_{pred} > t^*$, and $\mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}}) \in \mathscr{B}_{T_{ini}}$, then, for any u_0, the slices of $\widetilde{O}_\infty$ and $\widetilde{O}^D_\infty$ are related through (23), which is a linear map.

Remark 4: Without the condition $\mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}}) \in \mathscr{B}_{T_{ini}}$ in Theorem 1, there is no direct relationship between $\widetilde{O}_\infty$ and $\widetilde{O}^D_\infty$. However, in practice, $\mathbf{u_{ini}}$ and $\mathbf{y_{ini}}$ are data collected from the system and thus represent admissible trajectories. Hence, in a (noise-free) implementation setting, this condition is fulfilled.

Remark 5: In the proof of Theorem 1, it is shown that the (latent) initial condition, x_0, can be recovered from $\mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}})$ using the matrices M_t and N_t, which require the knowledge of A, B, C, and D matrices. This makes sense because, in order to determine x_0, a coordinate system must first be established, and the model realization through the A, B, C, D matrices serves this purpose.

We now illustrate the above theorem and the properties of $\widetilde{O}^D_\infty$ using an example. Consider a first-order system with matrices $A = 0.5$, $B = 1$, $C = 1$, $D = 0.5$, and constraint $-1 \leq y(t) \leq 1$. We let $\epsilon = 0.01$ and compute $\widetilde{O}_\infty$ using the method presented in [1]. For this system, $\ell = 1$ and $t^* = 0$. To compute $\widetilde{O}^D_\infty$, we collect offline input-output data (with random noise as input), and form the data matrix with $T_{ini} = T_{pred} = 1$. For this first-order example, any $\mathbf{col}(\mathbf{u_{ini}}, \mathbf{y_{ini}})$ is admissible, so the assumptions in Theorem 1 are all fulfilled. Finally, we compute $\widetilde{O}^D_\infty$ using (21). The polytopes corresponding to $\widetilde{O}_\infty$ and $\widetilde{O}^D_\infty$ are illustrated in Fig. 2. As can be seen, $\widetilde{O}^D_\infty$ lives in $\mathbb{R}^3$ while $\widetilde{O}_\infty$ lives in $\mathbb{R}^2$. However, both sets are characterized by exactly four inequalities (i.e., bounded by four facets). Thus, while $\widetilde{O}_\infty$ is compact, $\widetilde{O}_\infty^D$ is not. Moreover, Eq. (23) for this system can be written as $x_0 = 0.75\mathbf{u_{ini}} + 0.5\mathbf{y_{ini}}$. That is, slices of $\widetilde{O}_\infty^D$ for constant u_0 are related to those of $\widetilde{O}_\infty$ through multiplication by the non-square matrix $[0.75 \quad 0.5]$.

As shown above, unlike $\widetilde{O}_\infty$, $\widetilde{O}_\infty^D$ is not compact. However, as the following theorem shows, $\widetilde{O}_\infty^D$ is positively invariant and contains the origin in its interior – two properties that it shares with $\widetilde{O}_\infty$. These properties allow us to leverage $\widetilde{O}_\infty^D$ for constraint management over infinite horizon, for example within the RG framework.

Theorem 2: Consider system (1). If $T_{ini} \geq \ell$, $T_{pred} > t^*$, and $\mathcal{H}$ satisfies (10) with $T = T_{ini} + T_{pred}$, then the set $\widetilde{O}_\infty^D$ as defined in (20) has the following properties:

1) $0 \in \text{Int}(\widetilde{O}_\infty^D)$;
2) Let $t_0 \in \mathbb{Z}^+$ and $\mathbf{col}(\mathbf{u_{ini}},\mathbf{y_{ini}})$ be input-output trajectories collected from the immediate past (i.e., from time $t_0 - T_{ini}$ to time $t_0 - 1$). If $u(t) = u_0, \forall t \geq t_0$, where $(\mathbf{u_{ini}}, \mathbf{y_{ini}}, u_0) \in \widetilde{O}_\infty^D$, then $y(t) \in \mathbb{Y}$, $\forall t \geq t_0$. Furthermore, $(\bar{u}_{ini}, \bar{y}_{ini}, u_0) \in \widetilde{O}_\infty^D$, where $\mathbf{col}(\bar{u}_{ini},\bar{y}_{ini})$ is the input-output trajectory from time $t_0 - T_{ini} + 1$ to time t_0; that is, $\widetilde{O}_\infty^D$ is positively invariant with respect to the dynamics in (1).

Proof: Part 1 follows from the fact that $\mathbf{y_{pred}}$ is linear in $\mathbf{u_{ini}}$, $\mathbf{y_{ini}}$, and u_0, and the fact that $0 \in \text{Int}(\mathbb{Y})$. To prove part 2, note that since $\mathbf{u_{ini}}$ and $\mathbf{y_{ini}}$ are collected from the immediate past, they are admissible trajectories. The result follows by invoking part 1 of Theorem 1 and using the fact that the model-based $\widetilde{O}_\infty$ is positively invariant and constraint-admissible [1]. ■

The above theorems assume that $T_{ini} \geq \ell$ and $T_{pred} > t^*$. The attentive reader may note that, without a model, the system lag, ℓ, and the admissibility index, t^*, may not be known *a priori*, so these assumptions may appear difficult to check. To overcome this challenge as it pertains to ℓ, note that ℓ satisfies: $\ell \leq n$, where n is the system order. Thus, if an upper bound, $\bar{n}$, on the system order is known, one can select $T_{ini} = \bar{n}$, which satisfies $T_{ini} \geq \ell$.

Similarly, if an upper bound, $\bar{t}^*$, on t^* is available, T_{pred} can be selected as $T_{pred} = \bar{t}^* + 1$. In many situations, however, such upper bound is not available. As the following theorem shows, even though t^* may not be known, one can easily check if T_{pred} has been chosen such that $T_{pred} \geq t^* + 2$.

Theorem 3: Consider system (1). Suppose $T_{ini} \geq \ell$, $T_{pred} > 1$, and $\mathcal{H}$ satisfies (10) with $T = T_{ini} + T_{pred}$. If the last q inequalities in (21) are all redundant (i.e., removing them does not change the set), then $T_{pred} \geq t^* + 2$. Otherwise, $T_{pred} \leq t^* + 1$.

Proof: Recall from the definition of t^* that, if $y(\infty) \in (1-\epsilon)\mathbb{Y}$ and $y(t) \in \mathbb{Y}$ for $t = 0, \ldots, t^*$, then $y(t) \in \mathbb{Y}$ for all $t > t^*$, i.e., the inequalities corresponding to $y(t) \in \mathbb{Y}$ are redundant for all $t \geq t^* + 1$. An implication of this is that, if there exists a $t_0 > 0$ such that all the inequalities corresponding to $y(t_0) \in \mathbb{Y}$ are redundant subject to $y(\infty) \in (1-\epsilon)\mathbb{Y}$ and $y(t) \in \mathbb{Y}$, $t = 0, \ldots, t_0 - 1$, then t_0 must satisfy $t_0 \geq t^* + 1$; see [1] for details. To employ this idea in the data-driven setting, we use the fact that $\mathbf{y_{pred}}$ defined in (15) provides output predictions from time $t = 0$ up to $t = T_{pred} - 1$. Thus, if the last q inequalities in (21) are all redundant, then T_{pred} must satisfy $T_{pred} \geq t^* + 2$, which follows from the above by letting $t_0 = T_{pred} - 1$. ■

Note that if the last q inequalities, as described in the Theorem, are not redundant, then $T_{pred} \leq t^* + 1$ and so it is likely that $T_{pred} \leq t^*$ as well. In this situation, one must select a larger T_{pred} (for example double it), recreate the data matrix, and repeat this process until redundancy is achieved, as described in the theorem.

Note that the above does not check whether $T_{pred} > t^*$; rather, it checks whether $T_{pred} \geq t^* + 2$. Thus, $\widetilde{O}_\infty^D$ will contain redundant inequalities, i.e., non-minimal complexity. However, this can be alleviated as discussed in the following remark.

Remark 6: Choosing non-minimal values for T_{ini} and T_{pred} has negative computational consequences. If T_{ini} is too large, then $\widetilde{O}_\infty^D$ lives in an unnecessarily high-dimensional space. If T_{pred} is too large, then the set may have non-minimal complexity (i.e., contain redundant inequalities). The latter can be alleviated by removing redundant constraints after constructing the set, using an algorithm similar to the one presented in [1].

IV. Data-Enabled Reference Governor (RG)

Here, we develop a data-enabled version of the RG, which leverages $\widetilde{O}_\infty^D$ instead of $\widetilde{O}_\infty$. The update law of the data-enabled RG remains the same as (6), but $\kappa \in [0, 1]$ is obtained by solving the following linear program (compare with (7)):

$$\begin{aligned} \underset{\kappa \in [0,1]}{\text{maximize}} \quad & \kappa \\ \text{s.t.} \quad & u_0 = u(t-1) + \kappa(r(t) - u(t-1)) \\ & \mathbf{u_{ini}} = \mathbf{col}(u(t - T_{ini}), \ldots, u(t-1)) \\ & \mathbf{y_{ini}} = \mathbf{col}(y(t - T_{ini}), \ldots, y(t-1)) \\ & (\mathbf{u_{ini}}, \mathbf{y_{ini}}, u_0) \in \widetilde{O}_\infty^D \end{aligned} \tag{24}$$

Assuming $\mathbf{u_{ini}}$ and $\mathbf{y_{ini}}$ are updated at every timestep using the input-output data collected from the system, Theorem 1 can be invoked to show that the data-enabled RG is *equivalent* to the standard, model-based RG; that is, it produces the same output $u(t)$ at every timestep. This RG, thus, *inherits* the properties of the standard RG [7], as summarized in the following theorem.

Theorem 4: The data-enabled RG enjoys the following properties:

1) Suppose that (24) is feasible at $t = 0$. Then, (24) is feasible for all $t > 0$.
2) The signal $u(t)$ computed using (6), (24) is bounded.
3) For a constant $r(t) = r$, $u(t)$ converges in finite time. Moreover, if $r(t)$ is steady-state constraint-admissible, then $u(t)$ converges to $r(t)$.

The proposed data-enabled RG has several advantages over its model-based counterpart, but also several disadvantages:

Pros: The data-enabled RG does not require a parametric model. Thus, it does not require a system identification step, nor does it require an observer for state estimation. For this reason, it provides more flexibility and an end-to-end solution to the practitioner.

Cons: First, only the measured output can be constrained, whereas in the model-based RG, any linear combination of the inputs and states can be constrained. Second, notice that the data-enabled RG enforces the constraints starting at $t = 0$, but requires online data from time $t = -T_{ini}$ to $t = -1$. Thus, the constraints can only be enforced after a delay of T_{ini} at the system startup. Third, since the data-enabled RG leverages a higher dimensional MAS, it requires more compute power and memory to store the data/matrices as compared to the model-based RG. To compare the memory usage, consider the efficient case where instead of storing the entire matrices

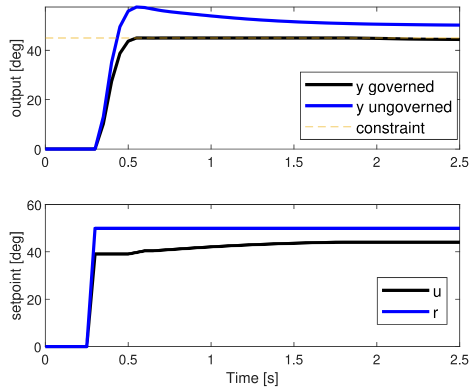

Fig. 3. Simulation results. In the top subplot, the solid black curve shows the output of the system driven by the proposed data-driven RG, and blue curve shows the response of the system without a reference governor (i.e., with $u(t)=r(t)$). In the bottom figure, the governed reference and the ungoverned reference are shown. See the bottom block diagram of Fig. 1 for the signal details.

$\mathbf{1}_{T_{pred}} \otimes s$ in (21) and h in (5), we only store ϵ and s. The total floating point numbers that must be stored in memory for the model-based and data-driven formulations of MAS/RG are:

$$\text{Model-based: } q\big((t^*+1)(n+m)+m+1\big)+1$$

$$\text{Data-enabled: } q\big(T_{pred}((m+p)T_{ini}+m)+m+1\big)+1$$

Since $T_{pred} \geq t^*+1$ and $T_{ini} \geq \ell$, the difference between the two, in the best case scenario where $T_{pred} = t^*+1$ and $T_{ini} = \ell$, is $q(t^*+1)((m+p)\ell - n)$, which is always strictly positive, meaning the data-driven version requires more memory. Of course, the standard RG requires an observer as well, which adds to the memory and computational requirements. So, a fair comparison can only be made on an application-to-application basis.

To illustrate the data-enabled RG, consider a DC motor described by $P(s) = \frac{1380}{s(0.15s+1)}$, which represents a Quanser QUBE Servo 2 experiment. The input to this model is voltage and the output is shaft angle. The goal is to design a closed-loop controller to track the setpoint, $u(t)$, and a data-enabled RG to govern $u(t)$ to enforce the constraint $y(t) \leq 45°$. To this end, we discretize $P(s)$ at $T_s = 0.05s$ using the sample and hold approach. We then choose a discrete PID controller with gains $K_p = 0.00755$, $K_i = 0.01$, $K_d = 0.000863$, which result in a rise time of 0.1s and phase margin of 60°. The closed loop system can be described by (1), where

$$A = \begin{bmatrix} 2.455 & -2.233 & 0.927 & -0.159 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$C = [\,0.261\ -0.2\ -0.21\ 0.159\,]$, and $D = 0$. To design an RG, we first choose $\epsilon = 0.02$ and compute the model-based $\widetilde{O}_\infty$, which has $t^* = 115$. We then collect a single long sequence of input-output offline data, by driving the system with random noise sampled from the uniform distribution. We use this input-output data to create the data matrix, $\mathcal{H}$, with $T_{pred} = 116$ and $T_{ini} = 4$ (for simplicity, we assume ℓ and t^* are known here). The characterizations of $\widetilde{O}_\infty$ and $\widetilde{O}^D_\infty$ require the storage of 583 and 1047 floating point numbers, respectively. We simulate the response of the closed-loop system to a step change in $r(t)$. The results are shown in Fig. 3. As can be seen, the response satisfies the constraint for all time, as desired.

V. Conclusion and Future Work

This letter proposed a data-driven formulation of the maximal admissible set (MAS) and the reference governor (RG). The formulations leverage output predictors from the behavioral system theory and subspace predictive control literature, and rely on offline clean data instead of parametric models. We presented the properties of the proposed data-driven MAS and RG and compared them with their model-based counterparts. Numerical simulations illustrated the results. Future work will study the robust formulations of data-driven MAS and RG in the case of systems affected by set-bounded disturbances and measurement noise. We will also study the impact of noise in the offline data, which degrades the performance of our data-driven predictors. Finally, we will study the extension of the methods herein to nonlinear and/or infinite dimensional LTI systems.

Acknowledgment

The author wishes to thank Miss Yasaman Pedari, and Drs. Ivan Markovsky, Florian Dörfler, and Ilya Kolmanovsky for fruitful discussions and feedback on this letter.

References

[1] E. G. Gilbert and K. T. Tan, "Linear systems with state and control constraints: The theory and application of maximal output admissible sets," *IEEE Trans. Autom. Control*, vol. 36, no. 9, pp. 1008–1020, Sep. 1991.

[2] K. T. Tan, "Maximal output admissible sets and the nonlinear control of linear discrete-time systems with state and control constraints," Ph.D. dissertation, Univ. Michigan, Ann Arbor, MI, USA, 1991.

[3] E. F. Camacho and C. B. Alba, *Model Predictive Control*. Heidelberg, Germany: Springer, 2013.

[4] J. B. Rawlings, D. Q. Mayne, and M. Diehl, *Model Predictive Control: Theory, Computation, and Design*. vol. 2. Madison, WI, USA: Nob Hill Publ., 2017.

[5] F. Blanchini, "Set invariance in control," *Automatica*, vol. 35, no. 11, pp. 1747–1767, 1999.

[6] F. Blanchini and S. Miani, *Set-Theoretic Methods in Control*. vol. 78. Boston, MA, USA: Springer, 2008.

[7] E. G. Gilbert, I. Kolmanovsky, and K. T. Tan, "Discrete-time reference governors and the nonlinear control of systems with state and control constraints," *Int. J. Robust Nonlin. Control*, vol. 5, no. 5, pp. 487–504, 1995.

[8] E. Garone, S. Di Cairano, and I. Kolmanovsky, "Reference and command governors for systems with constraints: A survey on theory and applications," *Automatica*, vol. 75, pp. 306–328, Jan. 2017.

[9] F. Lamnabhi-Lagarrigue et al., "Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges," *Annu. Rev. Control*, vol. 43, pp. 1–64, 2017.

[10] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. M. De Moor, "A note on persistency of excitation," *Syst. Control Lett.*, vol. 54, no. 4, pp. 325–329, 2005.

[11] I. Markovsky and F. Dörfler, "Behavioral systems theory in data-driven analysis, signal processing, and control," *Annu. Rev. Control*, vol. 52, pp. 42–64, 2021.

[12] W. Favoreel, B. De Moor, and M. Gevers, "SPC: Subspace predictive control," *IFAC Proc. Vol.*, vol. 32, no. 2, pp. 4004–4009, 1999.

[13] M. Ikeda, Y. Fujisaki, and N. Hayashi, "A model-less algorithm for tracking control based on input-output data," *Nonlin. Anal., Theory, Methods Appl.*, vol. 47, no. 3, pp. 1953–1960, 2001.

[14] F. Dörfler, J. Coulson, and I. Markovsky, "Bridging direct & indirect data-driven control formulations via regularizations and relaxations," *IEEE Trans. Autom. Control*, vol. 68, no. 2, pp. 883–897, Feb. 2023.

[15] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the DeePC," in *Proc. 18th Eur. Control Conf. (ECC)*, 2019, pp. 307–312.

[16] I. Markovsky and F. Dörfler, "Identifiability in the behavioral setting," *IEEE Trans. Autom. Control*, vol. 68, no. 3, pp. 1667–1677, Mar. 2023.

[17] I. Markovsky and H. Ossareh, "Finite-data nonparametric frequency response evaluation without leakage," *Automatica*, vol. 159, Jan. 2024, Art. no. 111351.