Willems' fundamental lemma for linear descriptor systems and its use for data-driven output-feedback MPC

Philipp Schmitz, Timm Faulwasser, and Karl Worthmann

***Abstract*— In this paper we investigate data-driven predictive control of discrete-time linear descriptor systems. Specifically, we give a tailored variant of Willems' fundamental lemma, which shows that for descriptor systems the non-parametric modelling via a Hankel matrix requires less data compared to linear time-invariant systems without algebraic constraints. Moreover, we use this description to propose a data-driven framework for optimal control and predictive control of discrete-time linear descriptor systems. For the latter, we provide a sufficient stability condition for receding-horizon control before we illustrate our findings with an example.**

***Index Terms*— Data-driven control, descriptor systems, discrete time, Hankel matrix, MPC, Willems' fundamental lemma, predictive control, non-parametric system description, optimal control**

I. Introduction

Recently, data-driven control—and in particular Willems' fundamental lemma [1]—is subject to substantial research interest. This includes non-parametric system representations for deterministic discrete-time linear time-invariant (LTI) systems [2] and linear parameter-varying (LPV) systems [3], stochastic LTI systems [4], as well as extensions to polynomial and non-polynomial nonlinear systems [5], [6]. These non-parametric representations enable system identification [7], control design [8], and also the implementation of predictive control [9], [10].

In the context of modelling of dynamical systems, continuous-time and discrete-time descriptor systems are of tremendous relevance in applications [11]. However, system-theoretic analysis as well as controller design for such systems face several challenges which range from existence of solutions [12], stability and controllability [13], to feedback design [14], [15]. In the context of model predictive control (MPC), early works on descriptor systems include [16]–[18], while more recent results can be found in [19], [20]. However, to the best of the authors' knowledge, only little has been done in terms of data-driven analysis and control of discrete-time descriptor systems. One of the few exceptions is [21], wherein identification and data-driven feedback design are discussed.

In the present paper, we show that the behavioral approach allows the consideration of linear discrete-time descriptor systems. To this end, we give a variant of the fundamental lemma tailored to such systems. Interestingly, it turns out that—compared to the usual LTI case—the necessary amount of data in the Hankel matrix is reduced for regular descriptor systems while the persistency of excitation requirements for the input signals do not change. Moreover, we leverage the developed non-parametric system description to derive a data-driven predictive control framework for LTI descriptor systems. We give a stability proof based on terminal constraints and illustrate the scheme with a numerical example.

The remainder of the paper is structured as follows: Section II recalls the basics of discrete-time linear descriptor systems such as their representation in quasi-Weierstraß form as well as specific controllability and observability notions. Section III presents and discusses a fundamental lemma for discrete-time descriptor systems, while Section IV turns towards data-driven predictive control tailored to this system class. In Section IV-C our findings are illustrated by an example before conclusions are drawn in Section V.

Notation: $\mathbb{N}_0$, $\mathbb{N}$ denote the natural numbers with and without zero, respectively. Moreover, for two numbers $a, b \in \mathbb{N}_0$ with $a \leq b$, the non-empty interval $[a, b] \cap \mathbb{N}_0$ is denoted by $[a : b]$. The identity and the zero matrix in $\mathbb{R}^{n\times m}$ are denoted by I_n and $0_{n\times n}$, respectively. For a matrix $A \in \mathbb{R}^{m\times n}$ we denote by $\operatorname{rk}(A)$ and $\operatorname{im}(A)$ the rank and the image of A, respectively. Further, for $k \in \mathbb{N}$ let $\operatorname{diag}_k(A) = I_k \otimes A$, where $\otimes$ denotes the Kronecker product.

For a function $f : \Omega \to \Gamma$, we denote the restriction of f to $\Omega_0 \subset \Omega$ by $f|_{\Omega_0}$. Considering a map $f : [t : T-1] \to \mathbb{R}^k$ with $t < T$, we denote the vectorization of f by

$$\mathbf{f}_{[t,T-1]} \doteq \begin{bmatrix} f(t)^\top & \dots & f(T-1)^\top \end{bmatrix}^\top \in \mathbb{R}^{k(T-t)}$$

and, for $L \in \mathbb{N}$ with $L \leq T - t$, the corresponding Hankel matrix $H_L(f_{[t,T-1]}) \in \mathbb{R}^{kL\times(T-t-L+1)}$ is defined by

$$H_L(\mathbf{f}_{[t,T-1]}) \doteq \begin{bmatrix} f(t) & \dots & f(T-L) \\ \vdots & \ddots & \vdots \\ f(t+L-1) & \dots & f(T-1) \end{bmatrix}.$$

Given a symmetric positive-definite matrix Q we define the norm $\|x\|_Q \doteq (x^\top Qx)^{1/2}$.

II. Basics of Linear Descriptor Systems

We consider discrete-time linear descriptor systems

$$Ex(t+1) = Ax(t) + Bu(t), \tag{1a}$$

$$y(t) = Cx(t) + Du(t), \tag{1b}$$

with (consistent) initial condition $(Ex)(0) = x^0$, where $A, E \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$. We

K. Worthmann gratefully acknowledges funding by the German Research Foundation DFG (WO 2056/6-1).

T. Faulwasser is with TU Dortmund University, Germany (e-mail: timm.faulwasser@ieee.org).

P. Schmitz and K. Worthmann, are with Technische Universität Ilmenau, Germany (e-mail: [philipp.schmitz, karl.worthmann]@tu-ilmenau.de).

assume that $\det(\lambda E - A) \neq 0$ for some $\lambda \in \mathbb{C}$, i.e., regularity of system (1a). Particularly, we are interested in the case where the matrix E is singular, i.e., $\mathrm{rk}(E) < n$.

We rely on the behavior notion given in [22, Definition 1.3.4], i.e., the trajectories of the system (1) are collected in the *full behavior*,

$$\mathfrak{B}_{\mathrm{f}} \doteq \left\{ (x,u,y) : \mathbb{N}_0 \to \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \;\middle|\; \begin{array}{c} x,u,y \text{ satisfy (1)} \\ \text{for all } t \in \mathbb{N}_0 \end{array} \right\}.$$

Further, we consider the input-output trajectories associated to the full behavior, i.e., the so-called *manifest behavior*

$$\mathfrak{B}_{\mathrm{m}} \doteq \left\{ (u,y) : \mathbb{N}_0 \to \mathbb{R}^m \times \mathbb{R}^p \;\middle|\; \begin{array}{c} \exists\, x : \mathbb{N}_0 \to \mathbb{R}^n : \\ (x,u,y) \in \mathfrak{B}_{\mathrm{f}} \end{array} \right\}. \quad (2)$$

For $t, T \in \mathbb{N}_0$, $t \leq T$, we denote the restrictions of the behaviors to the finite time interval $[t,T]$ by $\mathfrak{B}_{\mathrm{f}}[t,T] \doteq \{ b|_{[t,T]} \,|\, b \in \mathfrak{B}_{\mathrm{f}} \}$ and $\mathfrak{B}_{\mathrm{m}}[t,T] \doteq \{ b|_{[t,T]} \,|\, b \in \mathfrak{B}_{\mathrm{m}} \}$, respectively. The *consistent* initial values of the system (1) are collected in

$$\mathfrak{V} \doteq \{ x^0 \in \mathbb{R}^n \,|\, \exists (x,u,y) \in \mathfrak{B}_{\mathrm{f}} \text{ with } (Ex)(0) = x^0 \}. \quad (3)$$

Since the descriptor system (1a) is regular, there exist invertible matrices P, $S \in \mathbb{R}^{n \times n}$ such that

$$\begin{aligned} SEP &= \begin{bmatrix} I_q & 0 \\ 0 & N \end{bmatrix}, \quad & SAP &= \begin{bmatrix} A_1 & 0 \\ 0 & I_r \end{bmatrix}, \\ SB &= \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, & CP &= \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \end{aligned} \quad (4)$$

where $N \in \mathbb{R}^{r \times r}$ is nilpotent with nilpotency index s, and $A_1 \in \mathbb{R}^{q \times q}$, $B_1 \in \mathbb{R}^{q \times m}$, $B_2 \in \mathbb{R}^{r \times m}$, $C_1 \in \mathbb{R}^{p \times q}$, $C_2 \in \mathbb{R}^{p \times r}$ with $q + r = n$, cf. [23] and [13, Section 8.2]. Upon introduction of the coordinate change $z = P^{-1}x$, system (1) can equivalently be written in quasi-Weierstraß form, i.e.

$$\begin{bmatrix} I_q & 0 \\ 0 & N \end{bmatrix} z(t+1) = \begin{bmatrix} A_1 & 0 \\ 0 & I_r \end{bmatrix} z(t) + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(t) \quad (5a)$$

$$y(t) = \begin{bmatrix} C_1 & C_2 \end{bmatrix} z(t) + Du(t). \quad (5b)$$

Although the quasi-Weierstraß form is not unique, the nilpotency index s and the state dimensions q and r do not depend on the particular transformation matrices P and S [14, Lemma 2.10]. Put differently, the indices q and s (and, thus, $r = n - q$) are invariants of the original system (1) preserved in the quasi-Weierstraß form. Similarly to before, we consider the full and manifest behavior as well as their restrictions to finite time intervals for system (5). Specifically, we denote these behaviors by $\mathfrak{B}'_{\mathrm{f}}$, $\mathfrak{B}'_{\mathrm{m}}$, $\mathfrak{B}'_{\mathrm{f}}[t,T]$, and $\mathfrak{B}'_{\mathrm{m}}[t,T]$, respectively. Note that $\mathfrak{B}'_{\mathrm{f}} = \{ (z,u,y) \,|\, (Pz,u,y) \in \mathfrak{B}_{\mathrm{f}} \}$. Further, observe that the manifest behaviors of (1) and (5) coincide, i.e., $\mathfrak{B}_{\mathrm{m}} = \mathfrak{B}'_{\mathrm{m}}$ and $\mathfrak{B}_{\mathrm{m}}[t,T] = \mathfrak{B}'_{\mathrm{m}}[t,T]$.

Given an input trajectory $u : \mathbb{N}_0 \to \mathbb{R}^m$ and an initial value $z_1^0 \in \mathbb{R}^q$ there is a unique trajectory $(z,u,y) \in \mathfrak{B}'_{\mathrm{f}}$ such that the state

$$z = \begin{bmatrix} z_1^\top & z_2^\top \end{bmatrix}^\top : \mathbb{N}_0 \to \mathbb{R}^{q+r}$$

satisfies $z_1(0) = z_1^0$. This state $z(t)$, $t \in \mathbb{N}_0$, is given by

$$z_1(t) = \quad A_1^t z_1(0) + \sum_{k=1}^{t} A_1^{t-k} B_1 u(k-1) \quad (6a)$$

$$z_2(t) = -\sum_{k=0}^{s-1} N^k B_2 u(t+k). \quad (6b)$$

Observe that to determine the state z at time t one needs the future inputs $u(t), \ldots, u(t+s-1)$. Respectively, the future inputs need to satisfy (6b). Hence, system (5) can be regarded as non-causal.

The set of consistent initial values of system (5) in quasi-Weierstraß form is given by $\mathfrak{V}' = S\mathfrak{V}$ and can be equivalently characterized as

$$\mathfrak{V}' = \left\{ \begin{bmatrix} z_1^0 \\ z_2^0 \end{bmatrix} \in \mathbb{R}^{q+r} \;\middle|\; \begin{array}{c} \exists u \in [0 : s-2] \to \mathbb{R}^m \text{ s.t.} \\ z_2^0 = -\displaystyle\sum_{k=0}^{s-2} N^{k+1} B_2 u(k) \end{array} \right\}. \quad (7)$$

The characterization (7) together with the transformation P gives rise to an equivalent description of the set of consistent initial values $\mathfrak{V}$ of the original system (1), cf. the concept of an input index in the continuous-time setting [24].

Next we recall the concepts of R-controllability and R-observability, established in [13], see also [25], [26].

Definition 1 (R-controllability and R-observability [13]): The descriptor system (1) is called *R-controllable* if

$$\mathrm{rk}\left(\begin{bmatrix} \lambda E - A & B \end{bmatrix} \right) = n \quad (8a)$$

holds for all $\lambda \in \mathbb{C}$. System (1) is called *R-observable* if

$$\mathrm{rk}\left(\begin{bmatrix} \lambda E - A \\ C \end{bmatrix} \right) = n \quad (8b)$$

holds for all $\lambda \in \mathbb{C}$. □

Remark 2 (Controllability/Observability conditions): The R-controllability property is equivalent to the usual Kalman controllability rank condition for z_1 in (5a)

$$\mathrm{rk}\left(\begin{bmatrix} B_1 & A_1 B_1 & \ldots & A_1^{q-1} B_1 \end{bmatrix} \right) = q, \quad (9a)$$

see [25]. Similarly, R-observabilty is equivalent to

$$\mathrm{rk}\left(\begin{bmatrix} C_1 \\ C_1 A_1 \\ \vdots \\ C_1 A_1^{q-1} \end{bmatrix} \right) = q. \quad (9b)$$

□

The next lemma provides a lower bound on the length of an input-output trajectory to guarantee uniqueness of the corresponding internal state.

Lemma 3 (Uniqueness of state trajectories): Consider system (1) let the corresponding values of q and s be known. Assume that (1) is R-observable. If two trajectories (x,u,y), $(\tilde{x},\tilde{u},\tilde{y}) \in \mathfrak{B}_{\mathrm{f}}[0, q+s-2]$ satisfy $u|_{[0,q+s-2]} = \tilde{u}|_{[0,q+s-2]}$ and $y|_{[0,q+s-2]} = \tilde{y}|_{[0,q+s-2]}$, then $x|_{[0,q-1]} = \tilde{x}|_{[0,q-1]}$.

Proof: We consider the corresponding trajectories (z,u,y), $(\tilde{z},\tilde{u},\tilde{y}) \in \mathfrak{B}'_{\mathrm{f}}[0,q+s-2]$ of the equivalent system (5), that is $z = P^{-1}x$, $\tilde{z} = P^{-1}\tilde{x}$. According to (6) we have

$$C_1\Big(A_1^t\big(z_1(0)-\tilde{z}_1(0)\big)+\sum_{k=1}^{t}A_1^{t-k}B_1\big(u(k-1)-\tilde{u}(k-1)\big)\Big)$$
$$-C_2\sum_{k=0}^{s-1}N^kB_2\big(u(t+k)-\tilde{u}(t+k)\big)+D\big(u(t)-\tilde{u}(t)\big)$$
$$=y(t)-\tilde{y}(t)=0$$

for $t = 1,\dots,q-1$, This implies

$$C_1A_1^t\big(z_1(0)-\tilde{z}_1(0)\big)=0 \tag{10}$$

for all $t = 0,\dots,q-1$. With (9b) this yields $z_1(0) = \tilde{z}_1(0)$. Moreover, (6b) implies $z_2(0) = \tilde{z}_2(0)$. By evolving the states z_1 and z_2 via (6) up to the time $q-1$ we find $z_1|_{[0,q-1]} = \tilde{z}_1|_{[0,q-1]}$ and $z_2|_{[0,q-1]} = \tilde{z}_2|_{[0,q-1]}$. The assertion follows with $x = Pz$ and $\tilde{x} = P\tilde{z}$. ■

III. The Fundamental lemma for descriptor systems

We recall the notion of persistency of excitation.

Definition 4 (Persistency of excitation): A function $u : [0:T-1] \to \mathbb{R}^m$ is said to be *persistently exciting* of order L if the Hankel matrix $H_L(\mathbf{u}_{[0,T-1]})$ has rank mL.□
Note that $(m+1)L-1 \le T$ is necessary for persistency of excitation. Further, persistent excitation of order L implies persistent excitation of lower order $\tilde{L}$, $\tilde{L} \le L$.

The next result shows that the vector space $\mathfrak{B}_{\mathrm{m}}[0,L-1]$ of input-output trajectories with finite-time horizon is spanned by a Hankel matrix built from input-output data. The result is implicitly included in the original fundamental lemma by Willems et al. [1], whose original proof heavily relies on algebraic concepts and is formulated in behavioral notation. Based on a result for explicit LTI systems [27] we give a proof in terms of state-space descriptions. This proof allows to deduce further insights, especially regarding the amount of data needed in the Hankel matrix. The basic idea for descriptor systems is that only (A_1,B_1) subsystem of the quasi-Weierstraß form (5a), which is an explicit LTI system, has to be persistently excited to reconstruct trajectories.

Lemma 5 (Fundamental lemma for descriptor systems): Suppose that the system (1) is R-controllable and regular. Let $(\bar{u},\bar{y}) \in \mathfrak{B}_{\mathrm{m}}[0,T-1]$ such that $\bar{u}$ is persistently exciting of order $L+q+s-1$ and $T,L \in \mathbb{N}$ satisfy $(m+1)(L+q+s)-1 \le T$. Then $(u,y) \in \mathfrak{B}_{\mathrm{m}}[0,L-1]$ if and only if there is $\alpha \in \mathbb{R}^{(m+p)L\times(T-s-L+2)}$ such that

$$\begin{bmatrix} H_L(\bar{\mathbf{u}}_{[0,T-s]}) \\ H_L(\bar{\mathbf{y}}_{[0,T-s]}) \end{bmatrix}\alpha = \begin{bmatrix} \mathbf{u}_{[0,L-1]} \\ \mathbf{y}_{[0,L-1]} \end{bmatrix}. \tag{11}$$

□

Proof: Without loss of generality, we assume that system (1a) is given in quasi-Weierstraß form (5a). The proof proceeds in two steps.

Step 1. Consider $\mathcal{S} \in \mathbb{R}^{qL\times q}$, $\mathcal{T} \in \mathbb{R}^{qL\times mL}$, $\mathcal{R} \in \mathbb{R}^{rL\times m(L+s-1)}$

$$\mathcal{S}=\begin{bmatrix} I_q \\ A_1 \\ \vdots \\ A_1^{L-1}\end{bmatrix},\quad \mathcal{T}=\begin{bmatrix} 0 & \dots & 0 & 0 \\ B_1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ A_1^{L-2}B_1 & \dots & B_1 & 0\end{bmatrix},$$
$$\mathcal{R}=\begin{bmatrix} B_2 & \dots & N^{s-1}B_2 & 0 & \dots & 0 \\ 0 & B_2 & \dots & N^{s-1}B_2 & \dots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \vdots \\ 0 & \dots & 0 & B_2 & \dots & N^{s-1}B_2 \end{bmatrix}$$

and $\mathcal{U} \in \mathbb{R}^{(n+m)L\times(q+m(L+s-1))}$, $\mathcal{V} \in \mathbb{R}^{(n+m+p)L\times(n+m)L}$

$$\mathcal{U} \doteq \left[\begin{array}{c|cc} \mathcal{S} & \mathcal{T} & 0_{qL\times m(s-1)} \\ 0_{rL\times q} & \multicolumn{2}{c}{-\mathcal{R}} \\ 0_{mL\times q} & I_{mL} & 0_{mL\times m(s-1)} \end{array}\right]$$
$$\mathcal{V} \doteq \begin{bmatrix} I_{qL} & 0_{qL\times rL} & 0_{qL\times mL} \\ 0_{rL\times qL} & I_{rL} & 0_{rL\times mL} \\ 0_{mL\times qL} & 0_{mL\times rL} & I_{mL} \\ \mathrm{diag}_L(C_1) & \mathrm{diag}_L(C_2) & \mathrm{diag}_L(D) \end{bmatrix}.$$

We show that $(z,u,y) \in \mathfrak{B}'_{\mathrm{f}}[0,L-1]$ if and only if

$$\begin{bmatrix} \mathbf{z}_{1\,[0,L-1]} \\ \mathbf{z}_{2\,[0,L-1]} \\ \mathbf{u}_{[0,L-1]} \\ \mathbf{y}_{[0,L-1]} \end{bmatrix} \in \mathrm{im}(\mathcal{V}\mathcal{U}) \tag{12}$$

holds, where $z(t)$ is composed of two vectors $z_1(t) \in \mathbb{R}^q$ and $z_2(t) \in \mathbb{R}^r$. To this end, let $(z,u,y) \in \mathfrak{B}'_{\mathrm{f}}[0,L-1]$. Then there exists $(z^*,u^*,y^*) \in \mathfrak{B}'_{\mathrm{f}}$ with $z^*|_{[0,L-1]} = z$, $u^*|_{[0,L-1]} = u$ and $y^*|_{[0,L-1]} = y$. The explicit solution (6) of (5a) gives

$$\begin{bmatrix} \mathbf{z}^*_{1\,[0,L-1]} \\ \mathbf{z}^*_{2\,[0,L-1]} \\ \mathbf{u}^*_{[0,L-1]} \end{bmatrix} = \mathcal{U}\begin{bmatrix} z_1^*(0) \\ \mathbf{u}^*_{[0,L+s-2]} \end{bmatrix},$$

which together with (5b) yields

$$\begin{bmatrix} \mathbf{z}_{1\,[0,L-1]} \\ \mathbf{z}_{2\,[0,L-1]} \\ \mathbf{u}_{[0,L-1]} \\ \mathbf{y}_{[0,L-1]} \end{bmatrix} = \begin{bmatrix} \mathbf{z}^*_{1\,[0,L-1]} \\ \mathbf{z}^*_{2\,[0,L-1]} \\ \mathbf{u}^*_{[0,L-1]} \\ \mathbf{y}^*_{[0,L-1]} \end{bmatrix} = \mathcal{V}\mathcal{U}\begin{bmatrix} z_1^*(0) \\ \mathbf{u}^*_{[0,L+s-2]} \end{bmatrix}. \tag{13}$$

This implies (12).

On the other hand, if (12) holds for some $(z,u,y) : [0:L-1] \to \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$, then there exists $(z^*,u^*,y^*) \in \mathfrak{B}'_{\mathrm{f}}[0,L-1]$ such that (13) holds. This implies $z^*|_{[0,L-1]} = z$, $u^*|_{[0,L-1]} = u$ and $y^*|_{[0,L-1]} = y$, which shows $(z,u,y) \in \mathfrak{B}'_{\mathrm{f}}[0,L-1]$.

Step 2. Consider $(\bar{u},\bar{y}) \in \mathfrak{B}'_{\mathrm{m}}[0,T-1]$. There exists $\bar{z} = \begin{bmatrix}\bar{z}_1^\top & \bar{z}_2^\top\end{bmatrix}^\top : [0:T-1] \to R^{q+r}$ such that $(\bar{z},\bar{u},\bar{y}) \in \mathfrak{B}'_{\mathrm{f}}[0,T-1]$. By assumption (A_1,B_1) from (5a) is controllable (cf. Remark 2) and $\bar{u}$ is persistently exciting of order $L+q+s-1$. As a consequence of [27, Thm. 1 (i)]

$$\mathcal{H} \doteq \begin{bmatrix} H_1(\bar{\mathbf{z}}_{1\,[0,T-L-s+1]}) \\ H_{L+s-1}(\bar{\mathbf{u}}_{[0,T-1]}) \end{bmatrix},$$

where $\mathcal{H} \in \mathbb{R}^{(q+m(L+s-1))\times(T-L-s+2)}$, has rank $q+m(L+s-1)$. Therefore, $\mathrm{im}(\mathcal{VU}) = \mathrm{im}(\mathcal{VUH})$.

Similar to (13) one sees that for the jth column of the matrix $\mathcal{H}$, where $j \in \{0,\ldots,T-L-s+1\}$,

$$\begin{bmatrix} \bar{\mathbf{z}}_{1\,[j,j+L-1]} \\ \bar{\mathbf{z}}_{2\,[j,j+L-1]} \\ \bar{\mathbf{u}}_{[j,j+L-1]} \\ \bar{\mathbf{y}}_{[j,j+L-1]} \end{bmatrix} = \mathcal{V} \begin{bmatrix} \bar{\mathbf{z}}_{1\,[j,j+L-1]} \\ \bar{\mathbf{z}}_{2\,[j,j+L-1]} \\ \bar{\mathbf{u}}_{[j,j+L-1]} \end{bmatrix} = \mathcal{VU} \begin{bmatrix} \bar{z}_1(j) \\ \bar{\mathbf{u}}_{[j,j+L+s-2]} \end{bmatrix}.$$

Hence, we have

$$\begin{bmatrix} H_L(\bar{\mathbf{z}}_{1\,[0,T-s]}) \\ H_L(\bar{\mathbf{z}}_{2\,[0,T-s]}) \\ H_L(\bar{\mathbf{u}}_{[0,T-s]}) \\ H_L(\bar{\mathbf{y}}_{[0,T-s]}) \end{bmatrix} = \mathcal{VUH}.$$

Consequently, $(z,u,y) \in \mathfrak{B}'_{\mathrm{f}}[0,L-1]$ if and only if

$$\begin{bmatrix} \mathbf{z}_{1\,[0,L-1]} \\ \mathbf{z}_{2\,[0,L-1]} \\ \mathbf{u}_{[0,L-1]} \\ \mathbf{y}_{[0,L-1]} \end{bmatrix} \in \mathrm{im}(\mathcal{VU}) = \mathrm{im}(\mathcal{VUH}) = \mathrm{im} \begin{bmatrix} H_L(\bar{\mathbf{z}}_{1\,[0,T-s]}) \\ H_L(\bar{\mathbf{z}}_{2\,[0,T-s]}) \\ H_L(\bar{\mathbf{u}}_{[0,T-s]}) \\ H_L(\bar{\mathbf{y}}_{[0,T-s]}) \end{bmatrix}.$$

The assertion follows from the definition of the manifest behavior (2) and $\mathfrak{B}'_{\mathrm{m}}[0,L-1] = \mathfrak{B}_{\mathrm{m}}[0,L-1]$. ■

Remark 6 (Upper-bounding the data demand): In general, the index s of the nilpotent matrix N and the dimension q of A_1 in the quasi-Weierstraß system (5) are unknown. However,

$$\begin{bmatrix} H_L(\bar{\mathbf{u}}_{[0,T-1]}) \\ H_L(\bar{\mathbf{y}}_{[0,T-1]}) \end{bmatrix} \alpha \in \mathfrak{B}_{\mathrm{m}}[0,L-1],$$

holds, provided that $\bar{u}$ is persistently exciting of order $L+k$, where $k \geq q+s-1$. An upper bound on k is given by the state dimension n of the original system (1). □

Remark 7 (Descriptor systems can work with less data): In the case the matrix E is invertible, i.e. $q=n$, $r=0$, and $s=1$, Lemma 5 coincides with results for LTI systems, see for instance in [27]. However, it deserves to be noted that in case of a singular matrix E the input-output trajectories of length L can be reconstructed by the Hankel matrix in (11) which contains only values of the trajectory $(\bar{u},\bar{y}) \in \mathfrak{B}_{\mathrm{m}}[0,T-1]$ up to the time $T-s$, while in the LTI case all values of $(\bar{u},\bar{y})$ are needed. This might be exploited for system whose physical interpretation gives rise to insights on s and q. □

Moreover, we conjecture that recent results which allow further reduction of the data demand in the Hankel matrix [28] carry over to the descriptor setting without major issues. The details are, however, beyond the scope of the present paper.

IV. Data-driven control for descriptor systems

In this section, we demonstrate the ramifications of Lemma 5 for optimal and predictive control. Suppose that system (1) is R-controllable and R-observable.

A. Descriptor systems: data-driven optimal control

The control objective is to steer the system to the origin in finite time, i.e., until the end of the optimization horizon. Moreover, the input-output trajectory is chosen such that a quadratic cost function is minimized. In the successor subsection, we embed this Optimal Control Problem (OCP) into a predictive control methodology.

Given an observed trajectory $(u,y) \in \mathfrak{B}_{\mathrm{m}}[t-q-s+1,t-1]$, we consider the OCP

$$\underset{(\hat{u},\hat{y})}{\mathrm{minimize}} \sum_{k=0}^{L-1} \|\hat{y}(t+k)\|_Q^2 + \|\hat{u}(t+k)\|_R^2 \tag{14a}$$

subject to $(\hat{u},\hat{y}) \in \mathfrak{B}_{\mathrm{m}}[t-q-s+1,t+L-1]$ and

$$\begin{bmatrix} \hat{\mathbf{u}}_{[t-q-s+1,t-1]} \\ \hat{\mathbf{y}}_{[t-q-s+1,t-1]} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_{[t-q-s+1,t-1]} \\ \mathbf{y}_{[t-q-s+1,t-1]} \end{bmatrix}, \tag{14b}$$

$$\begin{bmatrix} \hat{\mathbf{u}}_{[t+L-q-s+1,t+L-1]} \\ \hat{\mathbf{y}}_{[t+L-q-s+1,t+L-1]} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{14c}$$

with symmetric positive-definte matrices $Q \in \mathbb{R}^{p\times p}$ and $R \in \mathbb{R}^{m\times m}$ in the quadratic stage cost. Clearly, the terminal equality constraint (14c) can be replaced by a terminal inequality constraint on the control $\hat{u}$ and the output $\hat{y}$ or even dropped. In the same way one can formulate an OCP targeting a setpoint $(u^{\mathrm{s}},y^{\mathrm{s}})$. We say $(u^{\mathrm{s}},y^{\mathrm{s}}) \in \mathbb{R}^m \times \mathbb{R}^p$ is a stationary setpoint if there is $(u,y) \in: \mathfrak{B}_{\mathrm{m}}$ with $u(t)=u^{\mathrm{s}}$ and $y(t)=y^{\mathrm{s}}$ for all $t \in \mathbb{N}_0$. In this setting the stage cost function penalizes the distance to $(u^{\mathrm{s}},y^{\mathrm{s}})$ and the terminal constraint is adapted to $(u^{\mathrm{s}},y^{\mathrm{s}})$.

The *consistency condition* (14b) ensures that the latent internal states of the true and the predicted trajectory are aligned up to time $t-1$, cf. Lemma 3. In particular, the internal state at time $t-1$ imposes further restrictions on the predicted input signal up to the time $t+s-2$.

Remark 8 (Relaxing the consistency condition): According to (10) in the proof of Lemma 3, the consistency condition (14b), which ensures consistency of the latent internal state with input and output, can be relaxed to

$$\begin{bmatrix} \hat{\mathbf{u}}_{[t-\vartheta-s+1,t-1]} \\ \hat{\mathbf{y}}_{[t-\vartheta-s+1,t-1]} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_{[t-\vartheta-s+1,t-1]} \\ \mathbf{y}_{[t-\vartheta-s+1,t-1]} \end{bmatrix},$$

if the rank condition

$$\mathrm{rk}\left(\begin{bmatrix} C_1 \\ C_1A_1 \\ \vdots \\ C_1A_1^{\vartheta-1} \end{bmatrix}\right) = q$$

holds with $\vartheta < q$ for the quasi-Weierstraß form (5). □

Lemma 5 implies that all trajectories contained in the manifest behavior $\mathfrak{B}_{\mathrm{m}}[t-q-s+1,t-1]$ can be parameterised by a Hankel matrix. Hence, assuming that there is an input-output trajectory $(\bar{u},\bar{y}) \in \mathfrak{B}_{\mathrm{m}}[0,T-1]$ such that $\bar{u}$ is persistently exciting of order $L+2(q+s-1)$, OCP (14) is equivalent to

$$\underset{\substack{(\hat{u},\hat{y}):[t-q-s+1:t+L-1]\to\mathbb{R}^m\times\mathbb{R}^p\\ \alpha(t)\in\mathbb{R}^{T-L-2s-q+3}}}{\text{minimize}} \sum_{k=0}^{L-1}\|\hat{y}(t+k)\|_Q^2+\|\hat{u}(t+k)\|_R^2 \tag{15a}$$

subject to

$$\begin{bmatrix}\hat{\mathbf{u}}_{[t-q-s+1,t+L-1]}\\ \hat{\mathbf{y}}_{[t-q-s+1,t+L-1]}\end{bmatrix} = \begin{bmatrix}H_{L+q+s-1}(\bar{\mathbf{u}}_{[0,T-s]})\\ H_{L+q+s-1}(\bar{\mathbf{y}}_{[0,T-s]})\end{bmatrix}\alpha(t), \tag{15b}$$

$$\begin{bmatrix}\hat{\mathbf{u}}_{[t-q-s+1,t-1]}\\ \hat{\mathbf{y}}_{[t-q-s+1,t-1]}\end{bmatrix} = \begin{bmatrix}\mathbf{u}_{[t-q-s+1,t-1]}\\ \mathbf{y}_{[t-q-s+1,t-1]}\end{bmatrix}, \tag{15c}$$

$$\begin{bmatrix}\hat{\mathbf{u}}_{[t+L-q-s+1,t+L-1]}\\ \hat{\mathbf{y}}_{[t+L-q-s+1,t+L-1]}\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}. \tag{15d}$$

We summarize our findings in the following proposition.

Proposition 9 (Equivalence of the OCPs): The OCPs (14) and (15) are equivalent, i.e.,

(a) OCP (14) is feasible if and only if OCP (15) is feasible,
(b) for every optimal solution $(u^\star, y^\star) \in \mathfrak{B}_{\mathrm{m}}[t-q-s+1, t+L-1]$ of the OCP (14), there exists $\alpha^\star(t) \in \mathbb{R}^{T-L-2s-q+3}$ such that $(u^\star, y^\star, \alpha^\star(t))$ is an optimal solution of OCP (15),
(c) for every optimal solution $(u^\star, y^\star, \alpha^\star(t))$ of OCP (15), $(u^\star, y^\star)$ is contained in the manifest behavior $\mathfrak{B}_{\mathrm{m}}[t-q-s+1, t+L-1]$ and optimal for OCP (14). □

Observe that the comments made in Remark 6 on the knowledge of the nilpotency index s, on the dimension q of A_1 in the quasi-Weierstraß form (5), as well as the principal need for less data (Remark 7) remain valid in the context of OCP (15).

B. Descriptor systems: data-driven predictive control

In predictive control OCP (15) is solved at each time step t and, for the solution $(u^\star, y^\star, \alpha^\star(t))$, the value $u^\star(t)$ is applied as new input $u(t)$ to the system (1). For the descriptor system (1) we propose the predictive control scheme based on the OCP (15) as summarized in Algorithm 1.

Here, we emphasize that, due to the absence of input constraints and due to R-controllability, the optimization problem with convex objective function and affine constraints has a feasible (and, thus, also an optimal solution) for all consistent initial values if the optimization horizon is sufficiently long, i.e., $L \geq \tilde{L}+q+s-2$, where $\tilde{L} = 2s+q$. Roughly speaking the first $s-1$ time steps of the prediction serve to satisfy the noncausal restrictions established by the consistency condition (14b) (see (6b)), followed by $q+s$ steps to steer the latent state into the origin (see (6) and (9a)). The terminal constraint (14c) guarantees that the latent state is zero on $[t+L-q, t+L-1]$, see Lemma 3. This ensures that every (initially) feasible and, in particular, every optimal solution can be extended, recursively feasible, i.e., feasibility of OCP (15) at the successor time instant $t+1$, cf. [29]. Analogously, one may conclude asymptotic stability of the origin—or of an arbitrary controlled equilibrium $(y^{\mathrm{s}}, u^{\mathrm{s}})$ if the stage cost is suitably adapted, i.e., $\|u-u^{\mathrm{s}}\|_R^2 + \|y-y^{\mathrm{s}}\|_Q^2$—w.r.t. the predictive control closed loop resulting from Algorithm 1. Moreover, note that the terminal equality constraint may be replaced by suitably constructed terminal inequality constraints, see, e.g., [17]–[19].

Proposition 10 (Recursive feasibility and stability): Let system (1) be R-controllable and R-observable and suppose that Q and R are symmetric positive definite. Let the prediction horizon $L \geq 2q+3s-2$. Assuming initial feasibility, i.e., feasibility of the OCP (15) at time $t=0$, feasibility is ensured for all $t \in \mathbb{N}$. Moreover, the origin is globally asymptotically stable w.r.t. the predictive control closed loop, whereby the domain of attraction is implicitly characterized by the set of all feasible consistent initial values. □

The proof follows the usual arguments [29] and is hence omitted. We remark that initial feasibility is guaranteed for consistent initial values at time $t=0$ if the optimization horizon is sufficiently long in view of the assumed R-controllability and R-observability as pointed out in the previous subsection. Moreover, we emphasize that the assertions of Proposition 10 remain valid if control constraints and output constraints are imposed. However, the assumed initial feasibility can then not be simply covered by choosing the prediction horizon L sufficient long despite the assumed R-controllability and R-observability.

Algorithm 1 : Data-driven predictive control
Input: horizon L, (pers. exciting) input/output data $(\bar{u}, \bar{y})$

1: Set $t = 0$
2: Measure $(u, y) \in \mathfrak{B}_{\mathrm{m}}[t-q-s+1, t-1]$
3: Compute $(u^\star, y^\star, \alpha^\star(t))$ to (15)
4: Apply $u(t) = u^\star(t)$
5: $t \leftarrow t+1$ and goto Step 2

C. Numerical example

We consider system (1) with

$$E = \begin{bmatrix}0&0&1&0\\1&2&0&2\\2&3&1&3\\1&2&0&2\end{bmatrix}, A = \begin{bmatrix}1&1&0&2\\0&2&1&1\\1&4&2&3\\-1&1&1&0\end{bmatrix}, B = \begin{bmatrix}-1\\2\\2\\3\end{bmatrix},$$

and the output matrices

$$C = \begin{bmatrix}1&2&1&2\\0&1&0&1\\1&2&1&1\\2&2&1&2\end{bmatrix}, \quad D = 0_{4\times 1}.$$

Via the matrices

$$P = \begin{bmatrix}0&-1&0&1\\-1&0&1&1\\1&0&0&-1\\1&1&-1&-1\end{bmatrix}, \quad S = \begin{bmatrix}0&-1&1&0\\1&2&-1&0\\-1&-1&1&0\\0&1&0&-1\end{bmatrix}$$

the system can be transformed into quasi-Weierstraß form ($s = q = 2$), which allows easily to verify the R-controllability as well as the R-observability via (9a) and (9b), cf. Remark 2.

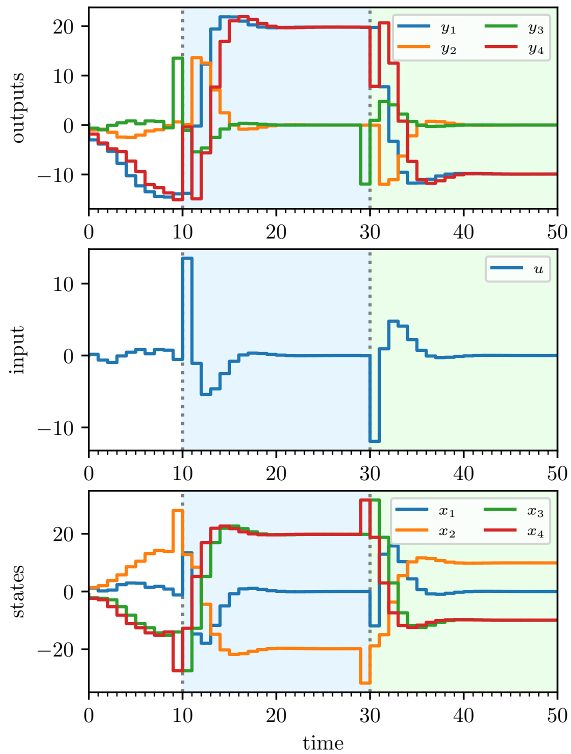

Fig. 1. A trajectory emerging from the predictive control scheme in Algorithm 1. At time $t = 10$ the transition from a random input signal to the optimal control input (blue shaded), which steers the system in to the controlled equilibrium $(u^{\mathrm{s},1}, y^{\mathrm{s},1})$, can be seen. At time $t = 30$ (green shaded) the desired controlled equilibrium is changed to $(u^{\mathrm{s},2}, y^{\mathrm{s},2})$. The depicted state x was calculated after the optimization for the sake of illustration.

We apply the predictive control Algorithm 1 with prediction horizon $L = 20$. For the input-output trajectory $(\bar{u}, \bar{y}) \in \mathfrak{B}_{\mathrm{m}}[0, T-1]$ with $T = 30$, the values of $\bar{u}$ are drawn independently from a uniform distribution over the interval $[-1, 1]$ such that $\bar{u}$ is persistently exciting of order $L + 2(q + s - 1) = 26$. Further, we assume that $R = I_m$ and $Q = I_p$. We want to steer the (1) to the setpoints

$$(u^{\mathrm{s},1}, y^{\mathrm{s},1}) = \left(0, \begin{bmatrix} 20 & 0 & 0 & 20 \end{bmatrix}^\top\right),$$
$$(u^{\mathrm{s},2}, y^{\mathrm{s},2}) = \left(0, \begin{bmatrix} -10 & 0 & 0 & -10 \end{bmatrix}^\top\right)$$

one by one. A closed-loop predictive control trajectory generated by Algorithm 1 is shown in Figure 1.

V. Conclusions

This paper has investigated data-driven control for linear discrete-time descriptor systems. We have shown that—compared to the usual LTI case—in the descriptor setting the data demand for the non-parametric system description via Hankel matrices is reduced. We leveraged Willems' fundamental lemma tailored to descriptor system to propose a data-driven predictive control scheme. We presented sufficient stability conditions and illustrated the findings with a numerical example. Interestingly, in the data-driven predictive control setting, and under the considered assumptions, the differences between usual LTI systems and their descriptor counterparts are marginal. This underpins the usefulness of Willems' fundamental and the prospect of data-driven predictive control.

References

[1] 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.

[2] 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.

[3] C. Verhoek, R. Tóth, S. Haesaert, and A. Koch, "Fundamental lemma for data-driven analysis of linear parameter-varying systems," *arXiv preprint arXiv:2103.16171*, 2021.

[4] G. Pan, R. Ou, and T. Faulwasser, "On a stochastic fundamental lemma and its use for data-driven MPC," *arXiv:2111.13636*, 2021.

[5] R. Strässer, J. Berberich, and F. Allgöwer, "Data-driven stabilization of nonlinear systems with rational dynamics," *arXiv preprint arXiv:2011.11355*, 2020.

[6] M. Alsalti, J. Berberich, V. G. Lopez, F. Allgöwer, and M. A. Müller, "Data-based system analysis and control of flat nonlinear systems," *arXiv preprint arXiv:2103.02892*, 2021.

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

[8] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 909–924, 2019.

[9] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1702–1717, 2020.

[10] E. O'Dwyer, E. C. Kerrigan, P. Falugi, M. A. Zagorowska, and N. Shah, "Data-driven predictive control with reduced computational effort and improved performance using segmented trajectories," *arXiv preprint arXiv:2108.10753*, 2021.

[11] S. Campbell, A. Ilchmann, V. Mehrmann, T. Reis *et al.*, *Applications of differential-algebraic equations: examples and benchmarks*. Springer, 2019.

[12] S. Trenn, "Solution concepts for linear daes: a survey," in *Surveys in Differential-Algebraic Equations I*. Springer, 2013, pp. 137–172.

[13] L. Dai, *Singular control systems*. Berlin etc.: Springer-Verlag, 1989, vol. 118.

[14] P. Kunkel and V. Mehrmann, *Differential-Algebraic Equations Analysis and Numerical Solution*. European Mathematical Society Publishing House, Zurich, CH, 2006.

[15] T. Reis and M. Voigt, "Linear-quadratic optimal control of differential-algebraic systems: the infinite time horizon problem with zero terminal state," *SIAM Journal on Control and Optimization*, vol. 57, no. 3, pp. 1567–1596, 2019.

[16] R. Findeisen and F. Allgöwer, "Nonlinear model predictive control for index—one dae systems," in *Nonlinear Model Predictive Control*. Springer, 2000, pp. 145–161.

[17] A. Yonchev, R. Findeisen, C. Ebenbauer, and F. Allgöwer, "Model predictive control of linear continuous time singular systems subject to input constraints," in *2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601)*, vol. 2. IEEE, 2004, pp. 2047–2052.

[18] J. Sjöberg, R. Findeisen, and F. Allgöwer, "Model predictive control of continuous time nonlinear differential algebraic systems," *IFAC Proceedings Volumes*, vol. 40, no. 12, pp. 48–53, 2007.

[19] A. Ilchmann, J. Witschel, and K. Worthmann, "Model predictive control for linear differential-algebraic equations," *IFAC-PapersOnLine*, vol. 51, no. 20, pp. 98–103, 2018.

[20] ——, "Model predictive control for singular differential-algebraic equations," *International Journal of Control*, pp. 1–10, 2021.

[21] J. He, F. Xu, and X. Wang, "Data-driven controllability analysis and stabilization for linear descriptor systems," *arXiv preprint arXiv:2112.03665*, 2021.

[22] J. W. Polderman and J. C. Willems, *Introduction to mathematical systems theory. A behavioral approach*. New York, NY: Springer, 1997, vol. 26.

[23] T. Berger, A. Ilchmann, and S. Trenn, “The quasi-Weierstrass form for regular matrix pencils,” *Linear Algebra Appl.*, vol. 436, no. 10, pp. 4052–4069, 2012.

[24] A. Ilchmann, L. Leben, J. Witschel, and K. Worthmann, “Optimal control of differential-algebraic equations from an ordinary differential equation perspective,” *Optimal Control Applications and Methods*, vol. 40, no. 2, pp. 351–366, 2019.

[25] A. A. Belov, O. G. Andrianova, and A. P. Kurdyukov, *Control of discrete-time descriptor systems. An anisotropy-based approach*. Cham: Springer, 2018, vol. 157.

[26] T. Stykel, “Analysis and numerical solution of generalized Lyapunov equations.” dissertation, TU Berlin, 2002.

[27] H. J. van Waarde, C. De Persis, M. K. Camlibel, and P. Tesi, “Willems' fundamental lemma for state-space systems and its extension to multiple datasets,” *IEEE Control Systems Letters*, vol. 4, no. 3, pp. 602–607, 2020.

[28] Y. Yu, S. Talebi, H. J. van Waarde, U. Topcu, M. Mesbahi, and B. Açıkmeşe, “On controllability and persistency of excitation in data-driven control: Extensions of Willems' fundamental lemma,” *arXiv preprint arXiv:2102.02953*, 2021.

[29] D. Mayne, J. Rawlings, C. Rao, and P. Scokaert, “Constrained model predictive control: Stability and optimality,” *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.