

# A quantitative and constructive proof of Willems' Fundamental Lemma and its implications

Julian Berberich<sup>1</sup>, Andrea Iannelli<sup>2</sup>, Alberto Padoan<sup>2</sup>, Jeremy Coulson<sup>2</sup>, Florian Dörfler<sup>2</sup>, and Frank Allgöwer<sup>1</sup>

**Abstract**—Willems' Fundamental Lemma provides a powerful data-driven parametrization of all trajectories of a controllable linear time-invariant system based on one trajectory with persistently exciting (PE) input. In this paper, we present a novel proof of this result which is inspired by the classical adaptive control literature and differs from existing proofs in multiple aspects. The proof involves a quantitative and directional PE notion, allowing to characterize robust PE properties via singular value bounds, as opposed to binary rank-based PE conditions. Further, the proof is constructive, i.e., we derive an explicit PE lower bound for the generated data. As a contribution of independent interest, we generalize existing PE results from the adaptive control literature and reveal a crucial role of the system's zeros.

## I. INTRODUCTION

The Fundamental Lemma by Willems et al. [1] allows all finite-length trajectories of a controllable linear time-invariant (LTI) system to be parametrized based on one input-output trajectory with persistently exciting (PE) input. Starting from this data-driven system parametrization, a plethora of system analysis and control methods have been developed, analyzed, and successfully applied in practice, see [2] for a recent survey. Further, different generalizations of the Fundamental Lemma have been developed in recent years, e.g., to certain classes of nonlinear systems [3]–[5] or by relaxing some of its assumptions [6]–[8]. Whereas the original result by [1] was formulated and proven in the behavioral framework, the recent paper [9] contains a proof in the state-space framework. An alternative proof in the behavioral framework for the single-input case is provided in [10]. In this paper, we present a quantitative and constructive proof of the Fundamental Lemma based on arguments from the classical literature on adaptive control [11]–[13].

PE properties and their implications on parameter convergence have been at the center of the field of adaptive control for several decades [11]–[13]. These classical PE notions are typically based on squares and summation, e.g., for some signal  $\{u_k\}_{k=0}^{N-1}$ , the sum  $\sum_{k=0}^{N-1} u_k u_k^\top$  is positive

F. Allgöwer is thankful that his work was funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2075 - 390740016 and under grant 468094890. F. Allgöwer acknowledges the support by the Stuttgart Center for Simulation Science (SimTech).

<sup>1</sup>University of Stuttgart, Institute for Systems Theory and Automatic Control, 70550 Stuttgart, Germany (email: {julian.berberich, frank.allgower}@ist.uni-stuttgart.de)

<sup>2</sup>Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland (email: {iannelli,apadoan,jcoulson,doerfler}@control.ee.ethz.ch)

definite. On the contrary, the Fundamental Lemma [1] and its applications in data-driven control [2] mainly rely on a rank-based PE notion. That is, the Hankel matrix of depth  $L$  corresponding to the sequence  $\{u_k\}_{k=0}^{N-1}$  with each  $u_k \in \mathbb{R}^m$ ,

$$H_L(u) = \begin{bmatrix} u_0 & u_1 & \dots & u_{N-L} \\ u_1 & u_2 & \dots & u_{N-L+1} \\ \vdots & \vdots & \ddots & \vdots \\ u_{L-1} & u_L & \dots & u_{N-1} \end{bmatrix} \in \mathbb{R}^{mL \times (N-L+1)},$$

needs to have full row rank. While this rank condition is equivalent to positive definiteness of  $H_L(u)H_L(u)^\top$ , i.e., to a squares-and-summation-based PE notion, only the latter allows to precisely quantify the richness of the data in terms of lower bounds on the eigenvalues of  $H_L(u)H_L(u)^\top$ . This quantification has several important implications which we describe in more detail later in the paper and which motivated [14] to derive a *robust* Fundamental Lemma. In this paper, we seek to further reconcile the behavioral approach to data-driven control with the classical adaptive control literature. To this end, we show that the Fundamental Lemma [1] can be proved using as starting point the main result of [12]. In addition to providing a different viewpoint on the Fundamental Lemma, the presented proof has the following advantages: 1) it involves a quantitative (similar to [14]) and directional description of PE properties, which leads to more robust conditions; 2) it is constructive, i.e., we derive an explicit lower bound on the richness of the data matrices. Further, open research questions for methods relying on the Fundamental Lemma concern noisy data and nonlinearities in the data-generating system. Owing to the constructive nature of the proof and its more direct links with PE properties, we believe that the results presented here can provide the basis for further generalizations, which constitute an interesting direction for future research.

Before providing a new proof of the Fundamental Lemma (Section III), we first extend the results of [12] in Section II, where we derive input design conditions to guarantee a PE output signal for arbitrary initial conditions and arbitrary output reachable systems. In contrast to [12], our results employ a directional PE notion, leading to less conservative results and illustrating the role of the Markov parameters in a transparent fashion. Moreover, we show that the PE requirements can be relaxed under certain conditions related to the number of zeros of the system.

*Notation:* We write  $\mathbb{I}_{\geq 0}$  ( $\mathbb{I}_{> 0}$ ) for the set of nonnegative (positive) integers and  $\mathbb{I}_{[a,b]}$  for all integers in the interval  $[a,b]$ . An identity matrix of dimension  $n$  is denoted by  $I_n$ . We denote the image of a matrix  $A$  by  $\text{im}(A)$ . The minimum eigenvalue and singular value of  $A = A^\top$  are defined as  $\lambda_{\min}(A)$  and  $\sigma_{\min}(A)$ , respectively. We write  $A \succ 0$  ( $A \succeq 0$ ) or  $A \prec 0$  ( $A \preceq 0$ ) if  $A$  is positive or negative (semi-)definite, respectively. The Kronecker product of two matrices  $A$  and  $B$  is denoted by  $A \otimes B$ . Further, we denote a window of a multivariate sequence  $\{x_k\}_{k=0}^{N-1}$  by  $x_{[a,b]} = [x_a^\top \ \dots \ x_b^\top]^\top$  and we write  $x = x_{[0,N-1]}$  for the full stacked sequence. We write  $\Sigma_{n,m,p}^{\text{or}}$  for the set of output reachable systems (see Definition II.3) of order  $n$  with  $m$  inputs and  $p$  outputs, i.e.,  $\Sigma_{n,m,p}^{\text{or}} = \{(A, B, C, D) \mid A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}, C \in \mathbb{R}^{p \times n}, D \in \mathbb{R}^{p \times m}, (A, B, C, D) \text{ output reachable}\}$ , and similarly  $\Sigma_{n,m,p}^{\text{c}}$  for the set of controllable systems. Finally, we define  $\{y_k(\Sigma, x_0, u)\}_{k=0}^{N-1}$  as the output sequence generated by applying the input  $\{u_k\}_{k=0}^{N-1}$  to the system  $\Sigma$  with initial condition  $x_0$ .

## II. REVISITING PERSISTENCE OF EXCITATION

After introducing the considered PE notion in Section II-A, we characterize under what conditions on the input signal the output of an LTI system is PE in Section II-B. This result will be instrumental in the new proof of the Fundamental Lemma in Section III, and it generalizes the main result of [12] by using a more general quantitative PE notion. In Section II-C, we show that the input PE condition can be relaxed under an assumption on the Markov parameters.

### A. Quantitative notion of PE

Consider the discrete-time LTI system

$$\Sigma_0 : \begin{cases} x_{k+1} &= Ax_k + Bu_k, \\ y_k &= Cx_k + Du_k \end{cases} \quad (1)$$

with state  $x_k \in \mathbb{R}^n$ , input  $u_k \in \mathbb{R}^m$ , and output  $y_k \in \mathbb{R}^p$ . Collect the first  $n+1$  Markov parameters of (1) in the matrix

$$\Gamma = \begin{bmatrix} \Gamma_0 & \Gamma_1 & \dots & \Gamma_n \\ D & CB & CAB & \dots & CA^{n-1}B \end{bmatrix}. \quad (2)$$

The following is a classical definition of persistence of excitation [1].

**Definition II.1.** [1] *The sequence  $\{u_k\}_{k=0}^{N-1}$  with  $u_k \in \mathbb{R}^m$  is persistently exciting (PE) of order  $L$  if  $\text{rank}(H_L(u)) = mL$ .*

This concept of persistence of excitation has the shortcoming that it is merely *qualitative* in the sense that, for a fixed order  $L$ , it is not possible to compare the PE levels of different signals. This motivates the following quantitative version of persistence of excitation as recently proposed by [14] in the context of data-driven control and originally proposed in a similar form in the adaptive control literature [11]–[13].

**Definition II.2.** *The sequence  $\{u_k\}_{k=0}^{N-1}$  is  $K$ -PE of order  $L$  for some matrix  $K \succ 0$  if*

$$H_L(u)H_L(u)^\top = \sum_{k=0}^{N-L} u_{[k,k+L-1]} u_{[k,k+L-1]}^\top \succeq K. \quad (3)$$

Note that  $\{u_k\}_{k=0}^{N-1}$  is  $K$ -PE of order  $L$  for some  $K \succ 0$  if and only if  $\{u_k\}_{k=0}^{N-1}$  is PE of order  $L$ . However, the flexibility gained via  $K$  in Definition II.2 allows for a more precise quantification of PE properties than Definition II.1. In particular, any choice  $K$  in Definition II.2 always reduces to the same rank condition in Definition II.1.

Persistence of excitation properties of system components, e.g., states, outputs, or regressors, play a crucial role in the classical literature on parameter estimation and adaptive control [11]–[13], where they can be used to guarantee convergence, as well as in the recent literature on data-driven control, see [2] for an overview. In contrast to the classical works [11]–[13], which typically involve infinitely many, shifted time windows or infinite sums, Definition II.2 imposes a condition over one finite time interval. As a second difference to these results and to the recently suggested PE notion by [14], we propose a not necessarily diagonal lower bound  $K$ , which allows to capture directional PE properties.

### B. When does a PE input imply a PE output?

By the Cayley-Hamilton theorem, there exists a polynomial  $d(z) = \sum_{j=0}^n d_j z^{n-j}$  with

$$d = [d_n \ \dots \ d_1 \ d_0]^\top \in \mathbb{R}^{n+1}, \quad d_0 \neq 0, \quad (4)$$

such that  $d(A) = 0$ . Without loss of generality, we assume  $\|d\|_2 = 1$ . Further, we define

$$M = [M_n \ \dots \ M_1 \ M_0] \in \mathbb{R}^{p \times m(n+1)} \quad (5)$$

with  $M_j = \sum_{q=0}^j d_{j-q} \Gamma_q$ . The following input-output representation of system (1) will play a crucial role, cf. [12].

**Lemma II.1.** [12] *Any trajectory  $\{u_k, y_k\}_{k=0}^\infty$  of (1) satisfies, for any  $k \geq n$ ,*

$$Mu_{[k-n,k]} = [y_{k-n} \ \dots \ y_{k-1} \ y_k] d. \quad (6)$$

Note that, for a single-input single-output (SISO) system,  $M$  and  $d$  simply contain the coefficients of the transfer function. The following alternative representations of  $M$  will be useful later in the paper:

$$M = (d^\top \otimes I_p) \bar{\Gamma} = \Gamma (\bar{D} \otimes I_m), \quad (7)$$

where

$$\bar{D} = \begin{bmatrix} d_n & d_{n-1} & \dots & d_0 \\ d_{n-1} & d_{n-2} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ d_0 & 0 & \dots & 0 \end{bmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}, \quad (8)$$

$$\bar{\Gamma} = \begin{bmatrix} \Gamma_0 & 0 & \dots & 0 \\ \Gamma_1 & \Gamma_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \Gamma_n & \dots & \Gamma_1 & \Gamma_0 \end{bmatrix} \in \mathbb{R}^{p(n+1) \times m(n+1)}. \quad (9)$$

The following result characterizes PE properties of the output of (1) in terms of the input for arbitrary initial conditions.

**Theorem II.1.** *If the signal  $\{Mu_{[k-n,k]}\}_{k=n}^{N-1}$  is PE of order 1, then, for any  $x_0 \in \mathbb{R}^n$ , the sequence  $\{y_k(\Sigma_0, x_0, u)\}_{k=0}^{N-1}$  is PE of order 1. Further, if  $\{Mu_{[k-n,k]}\}_{k=n}^{N-1}$  is  $K_u$ -PE of order 1, then, for any  $x_0 \in \mathbb{R}^n$ ,  $\{y_k(\Sigma_0, x_0, u)\}_{k=0}^{N-1}$  is  $K_y$ -PE of order 1 with*

$$K_y = \frac{1}{n+1} K_u. \quad (10)$$

*Proof.* Let us abbreviate  $y(\Sigma_0, x_0, u)$  by  $y$ . Note that

$$\begin{aligned} K_u &\preceq \sum_{k=n}^{N-1} Mu_{[k-n,k]} u_{[k-n,k]}^\top M^\top \\ &\stackrel{(6)}{=} \sum_{k=n}^{N-1} [y_{k-n} \ \dots \ y_k] dd^\top [y_{k-n} \ \dots \ y_k]^\top \\ &\preceq \|d\|_2^2 \sum_{k=n}^{N-1} \sum_{j=0}^n y_{k-j} y_{k-j}^\top \preceq (n+1) \sum_{k=0}^{N-1} y_k y_k^\top, \end{aligned} \quad (11)$$

using  $\|d\|_2^2 = 1$  in the last step. This proves both statements.  $\square$

Theorem II.1 extends [12, Theorem 2.1] by providing a matrix-valued lower bound (10) instead of a diagonal one, i.e.,  $K_u = k_u I$  for some  $k_u > 0$ . The result yields a sufficient condition for designing an input which guarantees a PE output, i.e.,  $\{Mu_{[k-n,k]}\}_{k=n}^{N-1}$  needs to be PE of order 1. In [12], it is claimed that this condition is not only sufficient but also necessary for achieving PE of the output for arbitrary initial conditions. This is, however, not true. In the appendix, we give a counterexample and discuss where the proof of [12] fails.

Note that the input design condition in Theorem II.1 involves the matrix  $M$ , which is typically unknown in estimation and data-driven control setups. In order to develop conditions which only involve the input, the concept of *output reachability* is essential, compare [12].

**Definition II.3.** *The system  $\Sigma$  is output reachable if, for any  $\hat{y} \in \mathbb{R}^p$ , there exist  $T \in \mathbb{I}_{\geq 0}$  and an input sequence  $\{u_k\}_{k=0}^T$  such that the resulting output  $\{y_k(\Sigma, 0, u)\}_{k=0}^T$  satisfies  $y_T(\Sigma, 0, u) = \hat{y}$ .*

By writing down the explicit output evolution, one observes that system (1) is output reachable if and only if  $\Gamma$  has full row rank, that is, if and only if  $M$  has full row rank [12, Lemma 2.1]. Note that, if the signal  $\{Mu_{[k-n,k]}\}_{k=n}^{N-1}$  is PE of order 1, then  $M$  has full row rank [12, Lemma 2.2]. Thus, output reachability is necessary for ensuring a PE output via Theorem II.1.

The following corollary of Theorem II.1, which extends [12, Corollary 2.1], provides a design condition for the input signal to ensure a PE output for an arbitrary output reachable system and for arbitrary initial conditions.

**Corollary II.1.** *If the signal  $\{u_k\}_{k=0}^{N-1}$  is PE of order  $n+1$ , then, for any  $\Sigma \in \Sigma_{n,m,p}^{\text{or}}$  and any  $x_0 \in \mathbb{R}^n$ , the sequence*

*$\{y_k(\Sigma, x_0, u)\}_{k=0}^{N-1}$  is PE of order 1. Further, if  $\{u_k\}_{k=0}^{N-1}$  is  $K_u$ -PE of order  $n+1$ , then, for any  $\Sigma \in \Sigma_{n,m,p}^{\text{or}}$  and any  $x_0 \in \mathbb{R}^n$ ,  $\{y_k(\Sigma, x_0, u)\}_{k=0}^{N-1}$  is  $K_y$ -PE of order 1 with*

$$K_y = \frac{1}{n+1} M K_u M^\top. \quad (12)$$

*Proof.* The first claim follows from Theorem II.1 since  $\{Mu_{[k-n,k]}\}_{k=n}^{N-1}$  is PE of order 1 for arbitrary  $M$  with full row rank if  $\{u_{[k-n,k]}\}_{k=n}^{N-1}$  is PE of order 1, which, in turn, holds if  $\{u_k\}_{k=0}^{N-1}$  is PE of order  $n+1$ . For the second claim, if  $\{u_k\}_{k=0}^{N-1}$  is  $K_u$ -PE of order  $n+1$  then  $M K_u M^\top \preceq \sum_{k=n}^{N-1} M u_{[k-n,k]} u_{[k-n,k]}^\top M^\top$ , from which (12) follows by the same steps as in the proof of Theorem II.1.  $\square$

Equation (12) provides a quantitative input design condition to ensure a desired output PE level  $K_y \succ 0$ , i.e., it is sufficient to choose the input as  $K_u$ -PE of order  $n+1$  for some  $K_u \succ 0$  satisfying  $\frac{1}{n+1} M K_u M^\top \succeq K_y$ . Corollary II.1 is conceptually also similar to the classical result [13, Theorem 2.9], which describes PE properties of filtered signals. The result remains true when replacing 1 by any integer  $L \in \mathbb{I}_{>0}$ , i.e., it can be shown that the output is PE of order  $L$  if the input is PE of order  $L+n$ . This extension is omitted for space reasons.

Corollary II.1 extends [12, Corollary 2.1] by allowing for a matrix-valued PE lower bound, yielding insights on *directional persistence of excitation*. To be precise, the matrix in (12) does not only bound the minimum singular value of  $\sum_{k=0}^{N-1} y_k y_k^\top$ , but is described via a (not necessarily diagonal) matrix depending on the system parameters via  $M$ . In fact, inserting  $M = \Gamma(\bar{D} \otimes I_m)$  from (7) in (12), we have

$$K_y \succeq \frac{\lambda_{\min}(K_u)}{n+1} \sigma_{\min}(\bar{D})^2 \Gamma \Gamma^\top. \quad (13)$$

Thus, the Markov parameters  $\Gamma$  directly influence (a lower bound on) the matrix  $K_y$  and, in particular, directions that are easy to reach or observe are more PE. This insight is conceptually similar to [15], which shows that the difficulty of system identification is influenced by the controllability properties of the underlying system.

Using  $M = (d^\top \otimes I_p) \bar{\Gamma}$  by (7) as well as  $\|d\|_2 = 1$ , we can also derive the following alternative bound

$$K_y \succeq \frac{1}{n+1} \lambda_{\min}(\bar{\Gamma} K_u \bar{\Gamma}^\top) I_p. \quad (14)$$

Since the lower bound (14) is diagonal, it does not provide any directional information on the richness of the output. However, in contrast to (13), it illustrates the role of the PE directions of the input. Specifically, the bound depends on  $\lambda_{\min}(\bar{\Gamma} K_u \bar{\Gamma}^\top)$ , i.e., to be more effective, the input should be tailored to the row space of the Toeplitz matrix  $\bar{\Gamma}$  containing the Markov parameters.

To summarize, Corollary II.1 implies the bounds (13) and (14), which allow for an intuitive interpretation in terms of directional persistence of excitation of the output and input, respectively. While these results cannot be used to construct explicit PE bounds if  $\Gamma$  and  $d$  are unknown, they reveal interesting insights in the interplay of the input and output PE properties and the system's Markov parameters.

### C. Relaxing PE requirements

Corollary II.1 provides an input design condition for guaranteeing a PE output for arbitrary output reachable systems. In this section, we investigate whether this condition can be relaxed when restricting the system class by using additional prior knowledge on the system. To this end, we consider the condition that the first  $r$  Markov parameters vanish, i.e.,

$$\Gamma_i = 0 \text{ for } i = 0, \dots, r-1, \text{ and } \Gamma_r \neq 0. \quad (15)$$

**Corollary II.2.** *If  $\{u_k\}_{k=0}^{N-r-1}$  is PE of order  $n+1-r$ , then, for any  $\Sigma \in \Sigma_{n,m,p}^{or}$  satisfying (15) and any  $x_0 \in \mathbb{R}^n$ , the sequence  $\{y_k(\Sigma, x_0, u)\}_{k=0}^{N-1}$  is PE of order 1.*

*Proof.* Note that (7) and (15) imply

$$Mu_{[k-n,k]} = \bar{M}_r u_{[k-n,k-r]} \quad (16)$$

with  $\bar{M}_r := (\bar{d}_r^\top \otimes I_p) \bar{\Gamma}_r$ ,  $\bar{d}_r := [d_{n-r} \ \dots \ d_0]^\top$ , and

$$\bar{\Gamma}_r := \begin{bmatrix} \Gamma_r & 0 & \dots & 0 \\ \Gamma_{r+1} & \Gamma_r & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \Gamma_n & \dots & \Gamma_{r+1} & \Gamma_r \end{bmatrix}.$$

If  $\{u_k\}_{k=0}^{N-r-1}$  is PE of order  $n+1-r$ , then  $\{\bar{M}_r u_{[k-n,k-r]}\}_{k=n}^{N-1}$  is PE of order 1 for arbitrary matrices  $\bar{M}_r$  with full row rank. Together with (16), this implies that  $\{Mu_{[k-n,k]}\}_{k=n}^{N-1}$  is PE of order 1 for arbitrary output reachable systems satisfying (15), i.e., by Theorem II.1, the output is PE of order 1.  $\square$

According to Corollary II.2, the PE requirements of Corollary II.1 can be relaxed depending on how many Markov parameters vanish. For SISO systems, (15) is equivalent to system (1) having relative degree  $r$  and  $n-r$  zeros. In this case, Corollary II.2 means that a sufficient condition for the output being PE is that the input is PE of order  $1 + \#\text{zeros}$ , i.e., additional zeros require a richer input signal to achieve the same PE level of the output. Verifying condition (15) only from data is an interesting issue for future research.

Note that the lower bound (14) is zero if  $\Gamma_0 = D$  does not have full row rank. In this case, a non-trivial bound can be derived by following the same strategy as in Corollary II.2.

Finally, the persistence of excitation condition for the input in Corollary II.2 only involves the data up to time step  $N-r-1$ . This is due to the fact that the time steps  $k = N-r, \dots, N-1$  have no influence on the output trajectory.

### III. A QUANTITATIVE AND CONSTRUCTIVE PROOF OF THE FUNDAMENTAL LEMMA

In this section, we present our main result: a novel proof of (a generalization of) the Fundamental Lemma based on [12], i.e., based on the results in Section II. Let us recall the Fundamental Lemma [1, Theorem 1] in the context of state-space systems, adjusted to our notation.

**Theorem III.1.** [1] *If  $\{u_k\}_{k=0}^{N-1}$  is PE of order  $L+n$ , then, for any  $\Sigma \in \Sigma_{n,m,p}^c$  and any  $x_0 \in \mathbb{R}^n$ ,*

$$\begin{aligned} & \text{im} \begin{bmatrix} H_L(u) \\ H_L(y(\Sigma, x_0, u)) \end{bmatrix} \\ &= \left\{ \begin{bmatrix} \bar{u} \\ \bar{y}(\Sigma, \bar{x}_0, \bar{u}) \end{bmatrix} \mid \bar{u} \in \mathbb{R}^{mL}, \bar{x}_0 \in \mathbb{R}^n \right\}. \end{aligned} \quad (17)$$

Theorem III.1 provides a parametrization of all system trajectories based on one input-output trajectory, and the result has found wide use in designing and analyzing data-driven control algorithms [2]. In the original paper [1], the Fundamental Lemma has been proven in the behavioral framework, and a proof for state-space systems has been given in [9]. In the following, we provide a novel, quantitative and constructive proof of Theorem III.1, which exploits the system description (6). To this end, we define the extended output  $\phi_k = \begin{bmatrix} x_k \\ u_k \end{bmatrix}$  which gives rise to the system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k, \\ \phi_k &= \tilde{C}x_k + \tilde{D}u_k, \end{aligned} \quad (18)$$

where  $\tilde{C} = \begin{bmatrix} I_n \\ 0 \end{bmatrix}$ ,  $\tilde{D} = \begin{bmatrix} 0 \\ I_m \end{bmatrix}$ . Throughout this section, whenever we apply the results from Section II, the considered ‘‘output’’ is  $\phi_k$  and the corresponding ‘‘input-output representation’’ (6) is associated with (18). Note that, for system (18), the Markov parameters  $\Gamma$  take the form  $\Gamma = \begin{bmatrix} 0 & \Gamma_{xu} \\ I & 0 \end{bmatrix}$  with the controllability matrix  $\Gamma_{xu} = [B \ AB \ \dots \ A^{n-1}B]$ . Thus, system (18) is output reachable if and only if  $(A, B)$  is controllable.

To prove the Fundamental Lemma, we are only concerned with the input-state system (18) and consider the matrix

$$\begin{bmatrix} H_1(x_{[0,N-L]}) \\ H_L(u) \end{bmatrix} = \begin{bmatrix} x_0 & x_1 & \dots & x_{N-L} \\ u_0 & u_1 & \dots & u_{N-L} \\ u_1 & u_2 & \dots & u_{N-L+1} \\ \vdots & \vdots & \ddots & \vdots \\ u_{L-1} & u_L & \dots & u_{N-1} \end{bmatrix}. \quad (19)$$

Specifically, we study under which conditions on the input the rank condition

$$\text{rank} \left( \begin{bmatrix} H_1(x_{[0,N-L]}) \\ H_L(u) \end{bmatrix} \right) = mL + n, \quad (20)$$

holds, compare [1, Corollary 2 (iii)]. In the language of Definition II.2, this means that  $\left\{ \begin{bmatrix} x_k \\ u_{[k,k+L-1]} \end{bmatrix} \right\}_{k=0}^{N-L}$  needs to be PE of order 1. Establishing (20) is the hard part of proving the Fundamental Lemma and, once shown, directly implies the statement of Theorem III.1, cf. [16, Lemma 2]. We summarize this fact in the following statement.

**Proposition III.1.** (17) *holds if (20) holds.*

In the following, we derive a sufficient condition for (20) to hold for arbitrary initial conditions and arbitrary controllable systems, which, according to Proposition III.1,

provides a sufficient condition for (17). To this end, with  $M$  corresponding to system (18), we define

$$Z = \begin{bmatrix} M & 0_{(m+n) \times m(L-1)} \\ 0_{m(L-1) \times m} & T \otimes I_m \end{bmatrix} \quad (21)$$

$$= \begin{bmatrix} M_n & M_{n-1} & \dots & M_1 & M_0 & 0 & \dots & 0 \\ 0 & d_n I & d_{n-1} I & \dots & d_1 I & d_0 I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_n I & d_{n-1} I & \dots & d_1 I & d_0 I \end{bmatrix},$$

where  $T$  is the Toeplitz matrix

$$T = \begin{bmatrix} d_n & \dots & d_0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & d_n & \dots & d_0 \end{bmatrix} \in \mathbb{R}^{(L-1) \times (L+n-1)}. \quad (22)$$

**Theorem III.2.** *If  $\{u_k\}_{k=0}^{N-1}$  is PE of order  $L+n$ , then, for any  $\Sigma \in \Sigma_{n,m,p}^c$  and any  $x_0 \in \mathbb{R}^n$ , (17) holds. Further, if  $\{u_k\}_{k=0}^{N-1}$  is  $K_u$ -PE of order  $L+n$ , then*

$$\begin{bmatrix} H_1(x_{[0,N-L]}) \\ H_L(u) \end{bmatrix} \begin{bmatrix} H_1(x_{[0,N-L]}) \\ H_L(u) \end{bmatrix}^\top \succeq \frac{1}{n+1} Z K_u Z^\top. \quad (23)$$

*Proof.* Define

$$\Phi_k := [\phi_{k-n} \quad \dots \quad \phi_k] = \begin{bmatrix} x_{k-n} & \dots & x_k \\ u_{k-n} & \dots & u_k \end{bmatrix},$$

$$U_k := \begin{bmatrix} u_{k-n+1} & u_{k-n+2} & \dots & u_{k+1} \\ u_{k-n+2} & u_{k-n+3} & \dots & u_{k+2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{k-n+L-1} & u_{k-n+L} & \dots & u_{k+L-1} \end{bmatrix}.$$

First, note that

$$U_k d = \begin{bmatrix} d_n u_{k-n+1} + \dots + d_0 u_{k+1} \\ \vdots \\ d_n u_{k-n+L-1} + \dots + d_0 u_{k+L-1} \end{bmatrix}$$

$$= (T \otimes I_m) u_{[k-n+1, k+L-1]}$$

with  $T$  from (22). Hence, using (6), we have

$$\begin{bmatrix} \Phi_k \\ U_k \end{bmatrix} d = \begin{bmatrix} M & 0 \\ 0 & T \otimes I_m \end{bmatrix} \begin{bmatrix} u_{[k-n, k]} \\ u_{[k-n+1, k+L-1]} \end{bmatrix} \quad (24)$$

$$= Z u_{[k-n, k+L-1]}.$$

Using that  $\{u_k\}_{k=0}^{N-1}$  is  $K_u$ -PE of order  $L+n$ , we obtain

$$Z K_u Z^\top \succeq \sum_{k=n}^{N-L} Z u_{[k-n, k+L-1]} u_{[k-n, k+L-1]}^\top Z^\top \quad (25)$$

$$\stackrel{(24)}{=} \sum_{k=n}^{N-L} \begin{bmatrix} \Phi_k \\ U_k \end{bmatrix} d d^\top \begin{bmatrix} \Phi_k \\ U_k \end{bmatrix}^\top \preceq \|d\|_2^2 \sum_{k=n}^{N-L} \begin{bmatrix} \Phi_k \\ U_k \end{bmatrix} \begin{bmatrix} \Phi_k \\ U_k \end{bmatrix}^\top$$

$$\preceq (n+1) \sum_{k=0}^{N-L} \begin{bmatrix} x_k \\ u_{[k, k+L-1]} \end{bmatrix} \begin{bmatrix} x_k \\ u_{[k, k+L-1]} \end{bmatrix}^\top,$$

which shows (23). Note that  $Z$  has full row rank since  $M$  has full row rank (by controllability) and  $d_0 \neq 0$ . Together with Proposition III.1, this proves the statement.  $\square$

Theorem III.2 shows that (a variation of) the main result from Section II can be used to prove the Fundamental Lemma from [1]. In fact, for  $L=1$ , Theorem III.2 follows from Corollary II.1 without any modification, i.e., the one-step version of the Fundamental Lemma [1, Corollary 2 (ii)] is actually a direct consequence of [12, Corollary 2.1]. The longer proof above is required for the general case  $L > 1$ .

In addition to providing a novel proof strategy, Theorem III.2 in fact generalizes the original Fundamental Lemma from [1] by providing a *quantitative and constructive* statement. Specifically, (23) reveals a direct connection between the PE level of the input and that of the input-state matrix  $\begin{bmatrix} H_1(x_{[0,N-L]}) \\ H_L(u) \end{bmatrix}$ . While positive definiteness of the left-hand side in (23) (i.e., the existence of *some* lower bound) is also guaranteed in existing proofs of the Fundamental Lemma [1], [9], we construct an explicit lower bound on the PE level. Improved PE bounds, in turn, imply better robustness with respect to noise, e.g., in system identification [14] or in data-driven predictive control [17, Theorem 3].

These insights are of practical use even in the realistic scenario that the system parameters in  $Z$  are only known approximately or if they are fully unknown. To be precise, suppose only an estimate  $\hat{Z}$  of  $Z$  in (21) is available with error bound  $\|\hat{Z} - Z\|_2 < \varepsilon$  for some  $\varepsilon > 0$ , and we want to use this knowledge to guarantee a certain PE bound in (23). If we let  $\{u_k\}_{k=0}^{N-1}$  be  $K_u$ -PE of order  $L+n$ , then the lower bound in (23) implies

$$\begin{bmatrix} H_1(x_{[0,N-L]}) \\ H_L(u) \end{bmatrix} \begin{bmatrix} H_1(x_{[0,N-L]}) \\ H_L(u) \end{bmatrix}^\top \succeq \frac{1}{n+1} Z K_u Z^\top$$

$$\succeq \frac{1}{2(n+1)} \hat{Z} K_u \hat{Z}^\top - \frac{\varepsilon^2}{n+1} \sigma_{\max}(K_u).$$

In this sense, our PE characterization is *robust* with respect to model inaccuracy. Similar bounds can be derived in the presence of process or measurement noise. Working out these insights in more detail and employing them for robust data-driven control is an interesting issue for future research.

We note that related results (Fundamental Lemma with quantitative PE bounds) are provided in [14] using a proof strategy similar to [9], however, only diagonal PE bounds and  $L=1$  are addressed.

#### IV. CONCLUSION

We presented a novel, quantitative and constructive proof of Willems' Fundamental Lemma, which has received significant attention in the context of data-driven control. Additionally, we extended classical results on adaptive control, which inspired our proof. We believe that the proposed new proof technique paves the way for further generalizations of the Fundamental Lemma and, since it enables robust PE guarantees for noisy/uncertain setups, it opens the door to new robust data-driven control results.

#### REFERENCES

- [1] J. C. Willems, P. Rapisarda, I. Markovsky, and B. De Moor, "A note on persistency of excitation," *Syst. Contr. Lett.*, vol. 54, pp. 325–329, 2005.

- [2] 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.
- [3] J. Berberich and F. Allgöwer, "A trajectory-based framework for data-driven system analysis and control," in *Proc. European Control Conf. (ECC)*, 2020, pp. 1365–1370.
- [4] J. G. Rueda-Escobedo and J. Schiffer, "Data-driven internal model control of second-order discrete Volterra systems," in *Proc. 59th IEEE Conf. Decision and Control (CDC)*, 2020, pp. 4572–4579.
- [5] C. Verhoeck, R. Tóth, S. Haesart, and A. Koch, "Fundamental lemma for data-driven analysis of linear parameter-varying systems," in *Proc. 60th IEEE Conf. Decision and Control (CDC)*, 2021, pp. 5040–5046.
- [6] Y. Yu, S. Talebi, H. J. van Waarde, U. Topcu, M. Mesbahi, and B. Açikmeşe, "On controllability and persistency of excitation in data-driven control: extensions of Willems' fundamental lemma," in *Proc. 60th IEEE Conf. Decision and Control (CDC)*, 2021, pp. 6485–6490.
- [7] V. K. Mishra, I. Markovsky, and B. Grossmann, "Data-driven tests for controllability," *IEEE Control Systems Lett.*, vol. 5, no. 2, pp. 517–522, 2021.
- [8] I. Markovsky and F. Dörfler, "Identifiability in the behavioral setting," *IEEE Trans. Automat. Control*, 2022, doi: 10.1109/TAC.2022.3209954.
- [9] 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 Lett.*, vol. 4, no. 3, pp. 602–607, 2020.
- [10] I. Markovsky, E. Prieto-Araujo, and F. Dörfler, "On the persistency of excitation," *Automatica*, vol. 147, p. 110657, 2023.
- [11] J. B. Moore, "Persistence of excitation in extended least squares," *IEEE Trans. Automat. Control*, vol. 28, no. 1, pp. 60–68, 1983.
- [12] M. Green and J. Moore, "Persistence of excitation in linear systems," *Syst. Contr. Lett.*, vol. 7, pp. 351–360, 1986.
- [13] K. J. Åström and B. Wittenmark, *Adaptive Control*. Reading, MA, USA: Addison-Wesley, 2008, second edition.
- [14] J. Coulson, H. van Waarde, and F. Dörfler, "Robust fundamental lemma for data-driven control," *arXiv:2205.06636*, 2022.
- [15] A. Tsiamias and G. J. Pappas, "Linear systems can be hard to learn," in *Proc. 60th Conf. Decision and Control (CDC)*, 2021, pp. 2903–2910.
- [16] C. De Persis and P. Tesi, "Formulas for data-driven control: stabilization, optimality and robustness," *IEEE Trans. Automat. Control*, vol. 65, no. 3, pp. 909–924, 2020.
- [17] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Trans. Automat. Control*, vol. 66, no. 4, pp. 1702–1717, 2021.

#### APPENDIX: ON NECESSITY IN THEOREM II.1

[12, Theorem 2.1] contains the following claim.

**Claim A.1.** *If  $\{y_k(\Sigma_0, x_0, u)\}_{k=0}^{N-1}$  is PE of order 1 for any  $x_0 \in \mathbb{R}^n$ , then  $\{Mu_{[k-n,k]}\}_{k=n}^{N-1}$  is PE of order 1.*

Claim A.1 means that PE of order 1 of  $\{Mu_{[k-n,k]}\}_{k=n}^{N-1}$  is not only sufficient but also *necessary* for  $\{y_k(\Sigma_0, x_0, u)\}_{k=0}^{N-1}$  being PE of order 1 for arbitrary initial conditions. However, as we show in the following, the claim is incorrect in general.

Suppose the system  $\Sigma_0$  takes the form

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, C = I_2, D = 0. \quad (26)$$

With the minimal polynomial  $A^2 = 0$ , we have  $d = [0 \ 0 \ 1]^\top$ ,  $M = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ . Suppose now the input is chosen as  $u_{[0,N-1]} = [1 \ 0 \ \dots \ 0]^\top$ . For a fixed initial condition  $x_0 = \begin{bmatrix} x_0(1) \\ x_0(2) \end{bmatrix}$ , the resulting output is

$$H_1(y) = \begin{bmatrix} y_0 & y_1 & y_2 & y_3 & \dots & y_{N-1} \end{bmatrix} \quad (27)$$

$$= \begin{bmatrix} x_0(1) & 1 & 0 & 0 & \dots & 0 \\ x_0(2) & x_0(1) & 1 & 0 & \dots & 0 \end{bmatrix}.$$

Note that  $H_1(y)$  has full row rank for any initial condition  $x_0 \in \mathbb{R}^2$ . On the other hand,  $\{Mu_{[k-n,k]}\}_{k=n}^{N-1}$  is not PE of order 1 since  $Mu_{[0,2]} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  but  $Mu_{[k-n,k]} = 0$  for all  $k = n+1, \dots, N-1$ . This invalidates Claim A.1.

Let us take a closer look at the proof of necessity in [12, Theorem 2.1], i.e., an attempt of proof for Claim A.1. To this end, suppose  $\{Mu_{[k-n,k]}\}_{k=n}^{N-1}$  is not PE of order 1, i.e., there exists  $0 \neq a \in \mathbb{R}^p$  such that  $a^\top Mu_{[k-n,k]} = 0$  for  $k = n, \dots, N-1$ . By (6), this implies for  $k = n, \dots, N-1$

$$\sum_{j=0}^n d_j a^\top y_{k-j} = 0. \quad (28)$$

The main proof idea from [12] is now to choose initial conditions such that  $a^\top y_k = 0$  for  $k = 0, \dots, n-1$ . If this were possible, then one could repeatedly apply (28) to conclude that  $a^\top y_k = 0$  for all  $k = n, \dots, N-1$ , thus proving that the output is not PE of order 1, i.e., proving Claim A.1. In [12], it is stated that such initial conditions can be chosen based on an observable realization of the transfer function from  $u$  to  $a^\top y$ . Indeed, suppose  $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$  is such a realization of order  $\tilde{n}$ , and denote the corresponding state by  $\tilde{x}_k \in \mathbb{R}^{\tilde{n}}$  and the output by  $\tilde{y}_k = a^\top y_k \in \mathbb{R}$ . The system dynamics imply

$$\tilde{y}_{[0,\tilde{n}-1]} = \Phi_{\tilde{n}} \tilde{x}_0 + \mathcal{T}_{\tilde{n}} u_{[0,\tilde{n}-1]} \quad (29)$$

for the observability matrix  $\Phi_{\tilde{n}}$  and a suitable matrix  $\mathcal{T}_{\tilde{n}}$ . Since  $\Phi_{\tilde{n}}$  has full column rank (by observability) and is square (since  $\tilde{y}_k \in \mathbb{R}$ ), it is invertible. Hence, choosing the initial condition  $\tilde{x}_0 = -\Phi_{\tilde{n}}^{-1} \mathcal{T}_{\tilde{n}} u_{[0,\tilde{n}-1]}$  leads to  $\tilde{y}_k = a^\top y_k = 0$  for  $k = 0, \dots, \tilde{n}-1$ .

The problem arises from the fact that  $\tilde{n}$  need not be equal to  $n$ . In particular, if  $\tilde{n} < n$ , there may not exist initial conditions which ensure that  $a^\top y_k = 0$  holds for  $k = \tilde{n}, \dots, n-1$ . Indeed, for the system (26), we have  $a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

and the system  $(A, B, a^\top C, a^\top D)$  is not observable. The state of an observable realization necessarily has dimension  $\tilde{n} = 1 < 2 = n$ . This means that we can only choose initial conditions to ensure  $a^\top y_0 = 0$ , but not  $a^\top y_1 = 0$ , cf. (27).

In view of the above argument, the proof of necessity from [12] applies and Claim A.1 holds if  $\tilde{n} = n$ , i.e.,  $(A, a^\top C)$  is observable. This gives rise to the following open questions: Are there relevant system classes for which observability of  $(A, a^\top C)$  can be guaranteed for any  $a$  satisfying  $a^\top Mu_{[k-n,k]}$  for  $k = n, \dots, N-1$ , independent of the choice of input? Conversely, for which systems is  $(A, a^\top C)$  not observable, i.e.,  $\tilde{n} < n$ , such that smaller levels of PE for the input may suffice to achieve a certain PE level for the output? What are the implications for necessity of PE in Corollary II.1 and in the Fundamental Lemma (Theorem III.2)? We leave these questions for future research.