2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Informativity for identification for $2D$ state-representable autonomous systems, with applications to data-driven simulation

Paolo Rapisarda[1] and Debasattam Pal[2]

Abstract— **We define persistency of excitation and informativity for system identification for the class of $2D$ state-representable autonomous systems. We characterize informativity for system identification in terms of properties of a matrix constructed from the restrictions of a system trajectory on successive consecutive lines. We state a procedure to compute arbitrary trajectories from a "sufficiently rich" one.**

I. Introduction

Data-driven approaches to control of dynamical systems are increasingly complementing traditional approaches based on modelling or system identification [1], [2], [3]. Recently, *informativity* of the data has been proposed as foundation for a new approach to data-driven control (see [4]). A dataset is informative with respect to a property $\mathcal{P}$ if it is possible to determine from it whether all systems that may have generated it possess $\mathcal{P}$. For $1D$ systems, that is, dynamical systems with 1 independent variable (time), informativity with respect to properties such as controllability, stabilizability, dissipativity, etc. has been recently analyzed in [4], [5]. *Informativity for system identification* is related to the notion of persistency of excitation (see [6], [7]). In this paper, we address the fundamental question of persistency of excitation for a class of $2D$ systems, i.e. systems with 2 independent variables (e.g. space and time).

A data-driven approach to nD systems, that is, systems with $n>1$ number of independent variables, has not yet been investigated. In general, the extension of results from $1D$ systems to the nD case is nontrivial. For example, while "past" and "future" are intuitive concepts in the $1D$ case, they are not when $n \geqslant 2$: a total ordering of $\mathbb{Z}^n$ must be *postulated* on the basis of physical insight, or *assumed* a priori. Furthermore, unlike $1D$ systems, a typical nD system does not in general admit a an input-state-output representation (see [8]). Moreover, the "lag" associated with a general nD system (a concept instrumental in $1D$ data-driven control) can turn out to be infinite ([9]).

In this paper, we begin to develop a data-driven approach to $2D$ systems. We restrict ourselves to systems with no inputs, i.e., *autonomous* systems (see Th. 3 p. 133 of [8] for the definition of autonomous $2D$ systems and its characterization). We define persistency of excitation ("informativity for identification") for $2D$ state-representable autonomous systems. Using the results of [8] we derive a characterization of such property. We exploit such characterization to devise an algorithm to generate admissible system trajectories *directly* from a given, "informative" one, *without identifying* the system beforehand. Following [10], [11], [12] we call such procedure *data-driven simulation*.

The paper is organized as follows: in Section II we summarize some basic background concepts about $2D$-systems. In Section III we define informativity for identification for $2D$-systems, and we introduce two matrices that play a crucial role in our data-driven simulation procedure. In Section IV we characterize informativity for identification for state-representable autonomous $2D$-systems, in terms of linear-algebraic properties of those matrices. In Section V we state a data-driven simulation procedure, and in Section VI we discuss some of the research being currently pursued.

Notation

We denote by $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$ respectively the set of natural, integer, real and complex numbers, and by $\mathbb{Z}^2 := \mathbb{Z}\times\mathbb{Z}$. $\mathbb{R}^n$, respectively $\mathbb{C}^n$, denote the space of n-dimensional vectors with real, respectively complex, entries. $\mathbb{R}^{n\times m}$ denotes the set of $n\times m$ matrices with real entries; and $\mathbb{R}^{n\times\infty}$ the set of real matrices with n rows and an infinite number of columns. The transpose of a matrix M is denoted by $M^\top$ and its pseudoinverse by $M^\dagger$. The image of M is denoted by $\operatorname{im} M$.

$\mathbb{R}[s_1,s_2]$ is the ring of polynomials with real coefficients in the indeterminates s_i, $i=1,\ldots,2$, and $\mathbb{R}[s_1,s_2,s_3,s_4]$ is the ring of polynomials with real coefficients in the indeterminates s_i, $i=1,\ldots,4$. Given a (finite or infinite) subset S of $\mathbb{R}[s_1,s_2]$, we denote by $\langle S\rangle$ the module of $\mathbb{R}[s_1,s_2]$ generated by the elements of S. A similar notation is used for the module generated by a subset of $\mathbb{R}[s_1,s_2,s_3,s_4]$.

II. Background results

We define $\left(\mathbb{R}^q\right)^{\mathbb{Z}^2} := \left\{w:\mathbb{Z}^2\to\mathbb{R}^q\right\}$, and we denote by σ_i, $i=1,2$, the shifts on $\left(\mathbb{R}^q\right)^{\mathbb{Z}^2}$, by

$$\sigma_1 w(i,j) := w(i+1,j)\ ,\ (i,j)\in\mathbb{Z}^2\ ,$$

and analogously for σ_2. We define σ_i^{-1}, $i=1,2$ by

$$\sigma_1^{-1}w(i,j) := w(i-1,j)\ \text{ and } \sigma_2^{-1}w(i,j) := w(i,j-1)\ .$$

We also define $\sigma := \sigma_1\sigma_2^{-1}$, the composition of σ_1 and σ_2^{-1}. We associate to a polynomial matrix $R \in \mathbb{R}^{p\times q}[s_1,s_2,s_1^{-1},s_2^{-1}]$ a polynomial operator in the shifts

$$R(\sigma_1,\sigma_2,\sigma_1^{-1},\sigma_2^{-1}) : \left(\mathbb{R}^q\right)^{\mathbb{Z}^2} \to \left(\mathbb{R}^p\right)^{\mathbb{Z}^2}\ ,$$

and we define $s := s_1 s_2^{-1}$.

[1]Paolo Rapisarda is with the School of Electronics and Computer Science at the University of Southampton, UK. pr3@ecs.soton.ac.uk

[2]Debasattam Pal is with the Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai, India. debasattam@ee.iitb.ac.in

A $2D$ system Σ is a triple $(\mathbb{Z}^2, \mathbb{R}^q, \mathfrak{B})$, where the *behavior* $\mathfrak{B}$ of the system is the subset of all possible $\mathbb{R}^q$-valued $\mathbb{Z}^2$-indexed sequences that satisfy the system equations. We use the following representations of $\Sigma = (\mathbb{Z}^2, \mathbb{R}^q, \mathfrak{B})$:

- *Square invertible:* $\mathfrak{B} := \ker R(\sigma_1, \sigma_2, \sigma_1^{-1}, \sigma_2^{-1})$, with
$$R(s_1, s_2, s_1^{-1}, s_2^{-1}) = I_q s_1^{\ell+1} + \ldots + A_0(s) , \quad (1)$$
$\deg(A_i(s)) \leq \ell + 1 - i$ or $A_i(s) = 0$, $i = 0, \ldots, \ell$.
- *State-driving variable*:
$\mathfrak{B} := \{w \mid \exists (x, v) : \mathbb{Z}^2 \to \mathbb{R}^n \times \mathbb{R}^d$ satisfying (2)$\}$
$$\sigma_1 x = A(\sigma)x + B(\sigma)v , \; w = Cx + Dv , \quad (2)$$
where $A(s) = A_0 + A_1 s \in \mathbb{R}^{n\times n}[s]$, $B(s) = B_0 + B_1 s \in \mathbb{R}^{n\times m}[s]$, $C \in \mathbb{R}^{q\times n}$ and $D \in \mathbb{R}^{q\times d}$.
- *State*: $\mathfrak{B} := \{w \mid \exists\, x$ satisfying (3)$\}$
$$\sigma_1 x = A(\sigma)x , \; w = Cx , \quad (3)$$
where $A(s) = A_0 + A_1 s \in \mathbb{R}^{n\times n}[s]$.

Remark 1: In (2) and (3) the value of x at $(i, j) \in \mathbb{Z}^2$ is computed from those of x, v at $(i - 1, j)$ and $(i, j - 1)$. ■

We make use of the following characterization, a combination of Theorem 6 p. 145 and Theorem 7 p. 153 of [8].

Theorem 1: Let $\Sigma = (\mathbb{Z}^2, \mathbb{R}^q, \mathfrak{B})$ be an autonomous system. The following statements are equivalent:

1) Σ admits a state representation (3);
2) Σ admits a square representation (1);
3) Σ admits a state-driving variable representation (2).

If any of the statements $(1) - (3)$ above holds, we call
$$\ell(\mathfrak{B}) := \min\{\ell \in \mathbb{N} \mid (1) \text{ holds}\} , \quad (4)$$
the *lag* of $\mathfrak{B}$. Given $\Sigma = (\mathbb{Z}^2, \mathbb{R}^q, \mathfrak{B})$, we denote by $\mathcal{N}(\mathfrak{B})$ the module of its *annihilators*, defined by
$$\mathcal{N}(\mathfrak{B}) := \left\{\eta \in \mathbb{R}^{1\times q}\left[s_1, s_2, s_1^{-1}, s_2^{-1}\right] \mid \eta(\sigma_1, \sigma_2, \sigma_1^{-1}, \sigma_2^{-1})w = 0 \;\forall\, w \in \mathfrak{B}\right\} . \quad (5)$$

Remark 2: $\mathfrak{B}$ and the module $\mathcal{N}(\mathfrak{B})$ are categorically dual (see [13]). This implies that the module of annihilators uniquely identifies the behavior. ■

Example 1: Consider the system with 2 state variables and 1 external variable ($n = 2$ and $q = 1$ in (3)) described by
$$\begin{aligned} \sigma_1 x &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \sigma x + \begin{bmatrix} \frac{1}{2} & 1 \\ 0 & \frac{1}{3} \end{bmatrix} x \\ w &= \begin{bmatrix} 1 & 1 \end{bmatrix} x . \end{aligned} \quad (6)$$
Eliminating the latent variable x (see section IV.C of [14]), one obtains the kernel representation
$$R(s_1, s_2) = s_1^2 + s_1\left(-\frac{5}{6} - 2s\right) + s_1^0\left(s^2 + \frac{5}{6}s + \frac{1}{6}\right) , \quad (7)$$
that satisfies the degree conditions for representations (1). The lag of this system $\ell(\mathfrak{B}) = 1$. ■

III. Informativity and the data matrix

We study systems representable as (1), (2) and (3); recalling Remark 1, our data consist of the values of system trajectories on a finite number of consecutive diagonal *lines*
$$\mathcal{L}_k := \{(i, j) \in \mathbb{Z}^2 \mid i + j = k\} ,$$
$k = 0, \ldots, N$. We define $\mathcal{L}_{0:N}$ by
$$\mathcal{L}_{0:N} := \{(i, j) \in \mathbb{Z}^2 \mid 0 \leq i + j \leq N\} .$$
We denote by $w_{|\mathcal{L}_k}$ the restriction of $w : \mathbb{Z}^2 \to \mathbb{R}^q$ to $\mathcal{L}_k$, and we denote by $w_{|\mathcal{L}_{0:N}}$ the restriction of w to $\mathcal{L}_{0:N}$.

Remark 3: We assume that w is not affected by noise. This assumption is often unrealistic, but it simplifies our analysis and allows us to gain some insight into the general problem of analyzing data in more realistic situations. ■

Given $\widehat{w} \in \mathfrak{B}$, we define the *data set* $\mathcal{D}_{0:N}(\widehat{w})$ by
$$\mathcal{D}_{0:N}(\widehat{w}) := \{\widehat{w}(i, j) \mid (i, j) \in \mathcal{L}_{0:N}\} = \widehat{w}_{|\mathcal{L}_{0:N}} . \quad (8)$$
We denote by $\mathcal{N}(\mathcal{D}_{0:N}(\widehat{w})) \subset \mathbb{R}^{1\times q}\left[s_1, s_2, s_1^{-1}, s_2^{-1}\right]$ the set of all annihilators of $\mathcal{D}_{0:N}(\widehat{w})$:
$$\mathcal{N}(\mathcal{D}_{0:N}(\widehat{w})) := \left\{\eta \in \mathbb{R}^{1\times q}\left[s_1, s_2, s_1^{-1}, s_2^{-1}\right] \mid \eta(\sigma_1, \sigma_2, \sigma_1^{-1}, \sigma_2^{-1})\widehat{w}_{|\mathcal{L}_{0:N}} = 0 \right\} , \quad (9)$$
and by $\langle\mathcal{N}(\mathcal{D}_{0:N}(\widehat{w}))\rangle$ the module generated by the elements of $\mathcal{N}(\mathcal{D}_{0:N}(\widehat{w}))$.

Definition 1: The data $\mathcal{D}_{0:N}(\widehat{w})$ are *informative for identification* if $\langle\mathcal{N}(\mathcal{D}_{0:N}(\widehat{w}))\rangle = \mathcal{N}(\mathfrak{B})$.

Given the duality between behavior and the module of its annihilators (Remark 2), data are informative for identification iff they *uniquely identify* the system producing them.

With reference to Definition 1, note that the inclusion $\langle\mathcal{N}(\mathcal{D}_{0:N}(\widehat{w}))\rangle \supseteq \mathcal{N}(\mathfrak{B})$ *always* holds: any annihilator of *all* trajectories of $\mathfrak{B}$ annihilates every particular $\widehat{w} \in \mathfrak{B}$. The data $\mathcal{D}_{0:N}(\widehat{w})$ are informative if also the *converse* inclusion holds, i.e. if the set of all difference equations satisfied by *all* system trajectories coincides with the set of *all* difference equations satisfied by the *particular* (sufficiently informative) trajectory $\widehat{w}$.

We introduce a matrix that plays a crucial role in characterizing this property. Let m be a nonnegative integer; define from $\widehat{w}_{|\mathcal{L}_k}$ the *Hankel matrix of depth* $m + 1$ by
$$\mathcal{H}_m(\widehat{w}_{|\mathcal{L}_k}) := \begin{bmatrix} \cdots & \widehat{w}_{k-1,1} & \widehat{w}_{k,0} & \cdots \\ \cdots & \widehat{w}_{k,0} & \widehat{w}_{k+1,-1} & \cdots \\ \cdots & \widehat{w}_{k+1,-1} & \widehat{w}_{k+2,-2} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & \widehat{w}_{k+m-1,-m+1} & \widehat{w}_{k+m,-m} & \cdots \end{bmatrix} . \quad (10)$$
$\mathcal{H}_m(\widehat{w}_{|\mathcal{L}_k})$ has $(m + 1)q$ rows and an infinite number of columns. From (8) we define the *data matrix*:
$$\mathbb{D}_N(\widehat{w}) := \begin{bmatrix} \mathcal{H}_{N+1}(\widehat{w}_{|\mathcal{L}_0}) \\ \mathcal{H}_N(\widehat{w}_{|\mathcal{L}_1}) \\ \vdots \\ \mathcal{H}_0(\widehat{w}_{|\mathcal{L}_N}) \end{bmatrix} . \quad (11)$$

Note that $\mathbb{D}_N(\widehat{w})$ has $\sum_{i=0}^{N}(i+1)q = \frac{(N+2)(N+3)}{2}q$ rows, and an infinite number of columns.

IV. A FUNDAMENTAL RESULT

Assume that N (the number of consecutive lines in $\mathbb{Z}^2$ on which the data are defined) is larger than or equal to $\ell(\mathfrak{B})$ (the lag of $\mathfrak{B}$, see (4)). Consider a representation (1) with $\ell = \ell(\mathfrak{B})$; it is straightforward to check that

$$s^k s_1^j R(s,s_1) = s^k s_1^j \left(I_q s_1^{\ell+1} + A_\ell(s)s_1^\ell + \ldots + A_0(s)\right)$$

also annihilates $\mathfrak{B}$ for every $k \in \mathbb{N}$ and for $j = 0,\ldots,N-\ell$. Now write the coefficient $A_i(s)$ of s_1^i in (1) as

$$A_i(s) = A_{i,0} + A_{i,1}s + \ldots + A_{i,\ell+1-i}s^{\ell+1-i}\ , \qquad (12)$$

$i = 0,\ldots,\ell$, and define $A_{\ell+1} := I_q$. Associate to (12) the $q \times (N+1-i)q$ matrix

$$\widetilde{A}_i := \begin{bmatrix} A_{i,0} & \ldots & A_{i,\ell+1-i} & 0_{q\times q} & \ldots & 0_{q\times q}\end{bmatrix}\ , \qquad (13)$$

$i = 0,\ldots,\ell$, and define

$$\widetilde{A} := \begin{bmatrix}\widetilde{A}_0 & \ldots & \widetilde{A}_N\end{bmatrix}\ , \qquad (14)$$

where we set $A_j(s) = 0$, $\widetilde{A}_j = 0$ if $j > \ell$. The association $A \leftrightarrow \widetilde{A}$ defines a bijective map between elements of $\mathbb{R}^{q\times q}[s,s_1]$ whose highest power of s is N, and elements of $\mathbb{R}^{q\times(N+1)(N+2)q}$.

Moreover, define the shift $\widetilde{\sigma}_1$ on the i-th matrix (13), $i = 0,\ldots,\ell$ by

$$\widetilde{\sigma}_1\widetilde{A}_i := \begin{bmatrix}0 & A_{i,0} & \ldots & A_{i,\ell+1-i} & 0_{q\times q} & \ldots\end{bmatrix} \in \mathbb{R}^{q\times(N+1-i)q}\ ,$$

and define its iterate $\widetilde{\sigma}_1^j$ in the natural way, $j = 0,\ldots,N-\ell$.

The following result is a straightforward consequence of these definitions.

Theorem 2: Assume that $\mathfrak{B}$ admits a representation (1), and that $N \geq \ell(\mathfrak{B})$. Define the matrices $\mathcal{H}_m(\widehat{w}_{|\mathcal{L}_k})$ and $\mathbb{D}$ by (10) and (11). The following statements are equivalent:

1) $\widehat{w}$ is informative for system identification;
2) $\mathcal{N}(\mathfrak{B})$ equals the module generated by left $\ker \mathbb{D}_N(\widehat{w})$;
3) The equality

$$\text{rank}(\mathbb{D}_N(\widehat{w})) = \frac{(N+2)(N+3)}{2}q - \frac{(N-\ell(\mathfrak{B})+1)(N-\ell(\mathfrak{B})+2)}{2}q$$

holds.

Define $d := \text{rank}(\mathbb{D}_N(\widehat{w}))$; assume that any of the statements $(1)-(3)$ holds. Let $V \in \mathbb{R}^{\frac{(N+2)(N+3)}{2}q\times d}$ be a basis matrix for $\text{im}\,\mathbb{D}_N(\widehat{w})$. For every $w \in \mathfrak{B}$ and $k \in \mathbb{N}$ there exists $L_k(w) \in \mathbb{R}^d$ such that the k-th column of $\mathbb{D}(w)$ equals

$$\begin{bmatrix} w_{k,-k} \\ \vdots \\ w_{k+N+1,-k-N-1} \\ w_{k+1,-k} \\ \vdots \\ w_{k+N+1,-k-N} \\ \vdots \\ w_{k+N+1,-k}\end{bmatrix} = VL_k(w)\ . \qquad (15)$$

Proof: The equivalence (1) $\iff$ (2) follows from the definition of informativity.

To prove (2) $\implies$ (3), recall that $N \geq \ell(\mathfrak{B})$, and write $N = \ell(\mathfrak{B}) + r$, with $r \geq 0$. Choose a basis for $\mathcal{N}(\mathfrak{B})$ consisting of the rows of a square-invertible representation (1). Associate to the elements of such basis the corresponding block-matrix $\widetilde{A}$ and its shifts $\sigma^k\widetilde{A}$, $k = 1,\ldots,N-\ell(\mathfrak{B})$. Observe that since the coefficient of $s_1^{\ell(\mathfrak{B})+1+k}$ in $s_1^k R(s,s_1)$ is the identity matrix, the shifts $\sigma^k\widetilde{A}$, $k = 0,\ldots,N-\ell(\mathfrak{B})$ are linearly independent.

To prove the claim, assume first that $r = 0$; then $\mathbb{D}_{\ell(\mathfrak{B})}(\widehat{w}) = \begin{bmatrix}\mathcal{H}_{\ell(\mathfrak{B})+1}(\widehat{w}_{|\mathcal{L}_0}) \\ \vdots \\ \mathcal{H}_0(\widehat{w}_{|\mathcal{L}_N})\end{bmatrix}$. The left-annihilators of $\mathbb{D}_{\ell(\mathfrak{B})}(\widehat{w})$ are generated by the q constant vectors associated with the rows of $I_q s_1^{\ell+1} + A_\ell(s)s_1^\ell + \ldots + A_0(s)$ through (12), (13) and (14); it follows that $\text{rank}(\mathbb{D}_{\ell(\mathfrak{B})}(\widehat{w})) = \frac{(N+2)(N+3)}{2}q - q$.

When $r = 1$, $\mathbb{D}_{\ell(\mathfrak{B})+1}(\widehat{w}) = \begin{bmatrix}\mathcal{H}_{\ell(\mathfrak{B})+2}(\widehat{w}_{|\mathcal{L}_0}) \\ \mathcal{H}_{\ell(\mathfrak{B})+1}(\widehat{w}_{|\mathcal{L}_0}) \\ \vdots \\ \mathcal{H}_1(\widehat{w}_{|\mathcal{L}_N}) \\ \mathcal{H}_0(\widehat{w}_{|\mathcal{L}_N})\end{bmatrix}$. Now $\mathbb{D}_{\ell(\mathfrak{B})}(\widehat{w})$ is a submatrix of $\mathbb{D}_{\ell(\mathfrak{B})+1}(\widehat{w})$; it follows that q left-annihilators of $\mathbb{D}_{\ell(\mathfrak{B})+1}(\widehat{w})$ are generated by the q vectors corresponding to $I_q s_1^{\ell+1} + A_\ell(s)s_1^\ell + \ldots + A_0(s)$. Moreover, $\begin{bmatrix}\mathcal{H}_{\ell(\mathfrak{B})+2}(\widehat{w}_{|\mathcal{L}_0}) \\ \mathcal{H}_{\ell(\mathfrak{B})+1}(\widehat{w}_{|\mathcal{L}_0}) \\ \vdots \\ \mathcal{H}_1(\widehat{w}_{|\mathcal{L}_N})\end{bmatrix}$ is also a submatrix of $\mathbb{D}_{\ell(\mathfrak{B})+1}(\widehat{w})$; this generates another $2q$ left-annihilators of $\mathbb{D}_{\ell(\mathfrak{B})+1}(\widehat{w})$ (since $\mathcal{H}_1(\widehat{w}_{|\mathcal{L}_N})$ has 2 block rows). It follows that $\text{rank}(\mathbb{D}_{\ell(\mathfrak{B})+1}(\widehat{w})) = \frac{(N+2)(N+3)}{2}q - (q+2q)$. Applying the same argument for $r > 1$ we conclude that $\text{rank}(\mathbb{D}_N(\widehat{w})) = \frac{(N+2)(N+3)}{2}q - \sum_{i=1}^{N-\ell(\mathfrak{B})+1} iq$; statement 3 of the Theorem is proved.

To prove the implication (3) $\implies$ (2), let (1) be a square-invertible representation of $\mathfrak{B}$; its rows generate the module $\mathcal{N}(\mathfrak{B})$. Associate to the rows of the polynomial matrix $R(s_1,s)$ constant vectors through (12), (13) and (14). Such constant vectors left-annihilate the data matrix of every trajectory in $\mathfrak{B}$, and consequently also $\mathbb{D}_N(\widehat{w})$. Such left-annihilators and their shifts are linearly independent of each other, since their highest coefficient is I_q. It is a matter of straightforward verification to check that for a fixed $N \geq \ell(\mathfrak{B})$ there are exactly $\sum_{i=1}^{N-\ell(\mathfrak{B})+1} iq$ such independent left annihilators. Now use statement 3 of the Theorem to conclude that since their dimensions are equal, the space of left-annihilators of $\mathbb{D}_N(\widehat{w})$ coincides with the space generated by the vectors associated to the rows of R. Since the rows of such square-invertible representation generate $\mathcal{N}(\mathfrak{B})$, statement 2 of the Theorem is proved.

We prove the last part of the claim. Let $w \in \mathfrak{B}$ be an arbitrary trajectory. Construct from its values on $\mathcal{L}_{|0:N}$ the matrix $\mathbb{D}_N(w)$ analogously to (11). Since the module generated by the polynomial vectors associated with elements of left $\ker \mathbb{D}(\widehat{w})$ equals $\mathcal{N}(\mathfrak{B})$, conclude that left $\ker \mathbb{D}_N(w) \supseteq$ left $\ker \mathbb{D}_N(\widehat{w})$, and consequently that $\operatorname{im} \mathbb{D}_N(w) \subseteq \operatorname{im} \mathbb{D}(\widehat{w})$. Since $\mathbb{D}_N(\widehat{w})$ has rank d (see statement 3 of the Theorem) one can compute a basis matrix V with the given dimensions. The last claim follows straightforwardly from the fact that V is a basis matrix for $\operatorname{im} \mathbb{D}_N(w)$. ∎

Example 2: We verify the statements of Theorem 2 on a numerical example. We use the state representation (6) of the system in Example 1 to compute data starting from a random sequence of two-dimensional state vectors on the line $\mathcal{L}_0$.

The Hankel matrices (10) and the data matrix $\mathbb{D}_N(\widehat{w})$ defined by (11) have an infinite number of columns; however, finite submatrices can be computed from them using a finite number of data samples. For the purposes of this example, we use 40 samples for every line $\mathcal{L}_k$.

For $N = 1$, $\mathbb{D}_1(\widehat{w})$ has 6 rows. Its singular values are $6.6577 \cdot 10$, 7.2037, 3.3494, $8.0790 \cdot 10^{-1}$, $3.5552 \cdot 10^{-1}$, and $3.0859 \cdot 10^{-15}$. It follows that the numerical rank of the data matrix is 5; the given trajectory is informative for identification (see statement 3 of Theorem 2). The singular value decomposition of $\mathbb{D}_1(\widehat{w})$ yields a single left annihilator that is proportional to

$$\begin{bmatrix} \frac{1}{6} & \frac{5}{6} & 1 & -\frac{5}{6} & -2 & 1 \end{bmatrix}, \tag{16}$$

the left annihilator corresponding to the coefficients of (7). Thus the module generated by the left annihilators of $\mathbb{D}_1(\widehat{w})$ equals the module of annihilators of $\mathfrak{B}$, confirming statement 2 in Theorem 2.

If we choose $N = 2$, the matrix $\mathbb{D}_2(\widehat{w})$ has 10 rows. Its last 3 singular values are $2.8010 \cdot 10^{-15}$, $1.4840 \cdot 10^{-15}$ and $6.8624 \cdot 10^{-16}$, respectively. The seventh singular value is $5.5213 \cdot 10^{-1}$, indicating that the numerical rank of $\mathbb{D}_2(\widehat{w})$ in this case equals 7. It can be verified that the left annihilators of $\mathbb{D}_2(\widehat{w})$ are generated by the rows of the matrix

$$\begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{5}{6} & 1 & -\frac{5}{6} & -2 & 1 \\ 0 & \frac{1}{6} & \frac{5}{6} & 1 & 0 & -\frac{5}{6} & -2 & 0 & 1 & 0 \\ \frac{1}{6} & \frac{5}{6} & 1 & 0 & -\frac{5}{6} & -2 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

The trajectory is informative for system identification and statement 2 in Theorem 2 is confirmed also for $N = 2$.

If we choose $N = 3$, $\mathbb{D}_3(\widehat{w})$ has 15 rows. Its last 6 singular values are below machine precision and the ninth is $4.7304 \cdot 10^{-1}$, indicating a numerical rank equal to 9. The subspace of left annihilators of $\mathbb{D}_3(\widehat{w})$ is generated by 6 shifts of the annihilator (16). Statements 2 and 3 of Theorem 2 are verified also in this case. ∎

Remark 4: The last claim of theorem 2 is analogous to the $1D$ discrete-time result in [6] known as the *fundamental lemma*, and to the analogous continuous-time results in [15], [12]. Namely, a matrix computed from sufficiently informative data can be used to compute *all* trajectories of the system (or restrictions thereof). ∎

V. A DATA-DRIVEN SIMULATION PROCEDURE

We use the last part of Theorem 2 to set up a data-driven recursive scheme to compute trajectories of $\mathfrak{B}$ from an informative one. For a given $w : \mathbb{Z}^2 \to \mathbb{R}^q$, we define the $\frac{(N+2)(N+3)}{2}q$-dimensional vectors $f_k(w)$ and $u_k(w)$ by

$$f_k(w) \quad := \quad \begin{bmatrix} w_{k-1,-k+1} \\ \vdots \\ w_{k+N,-k-N} \\ w_{k,-k+1} \\ \vdots \\ w_{k+N,-k-N+1} \\ \vdots \\ w_{k+N,-k} \end{bmatrix} \quad \text{and}$$

$$u_k(w) \quad := \quad \begin{bmatrix} 0 \\ \vdots \\ w_{k+N+1,-k-N-1} \\ 0 \\ \vdots \\ w_{k+N+1,-k-N} \\ \vdots \\ w_{k+N+1,-k+1} \end{bmatrix}. \tag{17}$$

In the following we call $f_k(w)$ the *unfolding at k* of the $2D$-sequence w over $\mathcal{L}_{0:N}$. We adopt such term since to construct f_k we "unfold" the values of w on an equilateral triangle of $\mathbb{Z}^2$ with apex at $(k+N, -k)$ and side length $N+1$.

Remark 5: The unfolding $f_k(w)$ defined by the first equation in (17) is a column of the data matrix $\mathbb{D}_N(w)$ computed from samples of w. Indeed its first $N+2$ block-entries constitute the column of $\mathcal{H}_{N+1}(w_{|\mathcal{L}_0})$ that starts with the value $w_{k-1,-k+1}$ (see (10)); the next $N+1$ block-entries are the samples of w on the line $\mathcal{L}_1$ starting from $w_{k,-k+1}$; and so forth until $w_{k+N,-k}$ on $\mathcal{L}_N$. ∎

Remark 6: The concept of "unfolding" is analogous to the principle underlying the state construction procedure in the proof of implication 3) $\Longrightarrow$ 1) in Th. 6 on p. 145 of [8]. ∎

We define the q-block shift matrix J_m by $J_0 = 0_{q \times q}$ and

$$J_m := \begin{bmatrix} 0 & I_q & 0 & \dots & 0 \\ 0 & 0 & I_q & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & I_q \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{(m+1)q \times (m+1)q},$$

for $m \geq 1$; and the *shift matrix* J by

$$J := \text{block diag}(J_k)_{k=N+1,\ldots,0} \; .$$

It is a matter of straightforward verification to check that for every $w : \mathbb{Z}^2 \to \mathbb{R}^q$ the equation $f_{k+1}(w) = Jf_k(w) + u_k(w)$ holds. Let V and L_k be defined as in Theorem 2. If $w \in \mathfrak{B}$, then there exists a sequence of vectors $\{L_k(w)\}$ such that $f_k(w) = VL_k(w)$ for every $k \in \mathbb{N}$; consequently

$$\begin{aligned} VL_{k+1}(w) = f_{k+1}(w) &= Jf_k(w) + u_k(w) \\ &= JVL_k(w) + u_k(w). \end{aligned}$$

Recall that V has full column rank, and denote by $V^\dagger$ a left inverse of V. Multiplying both sides of the last equation by $V^\dagger$ we obtain

$$L_{k+1}(w) = V^\dagger JVL_k(w) + V^\dagger u_k(w) \; . \tag{18}$$

We call $w \in (\mathbb{R}^q)^{\mathbb{Z}^2}$ an *admissible trajectory on* $\mathcal{L}_{0:N}$ if w satisfies the dynamical laws of the system on $\mathcal{L}_{0:N}$. We use (18) to generate values of admissible trajectories on $\mathcal{L}_{0:N}$ as follows. Let v be an initial unfolding at k; there exists a unique $L_k(w) \in \mathbb{R}^d$ such that $v = VL_k(w)$, since the columns of V are linearly independent from each other. Choose arbitrary values for the nonzero entries of u_k defined in (17), and compute $L_{k+1}(w)$ from (18). The new values of an admissible trajectory can now be read in the last q components of every $(m+1)q$-dimensional block of $VL_{k+1}(w)$, $m = N+1, \ldots, 0$. The process can be repeated and the values of an admissible trajectory on $\mathcal{L}_{0:N}$ can be iteratively computed.

We state a procedure formalizing these conceptual steps. We assume that an informative $\widehat{w}$ has been measured; that $\ell(\mathfrak{B})$ is known; that $N \geq \ell(\mathfrak{B})$ is fixed; and that $\mathbb{D}(\widehat{w})$ defined by (10) and (11) has been computed from $\widehat{w}_{\mathcal{L}_{0:N}}$. As in Theorem 2 we denote by d the dimension of $\operatorname{im} \mathbb{D}_N(\widehat{w})$.

Algorithm
Data-driven simulation of a $2D$ autonomous system $\mathfrak{B}$ representable by (1).

Input: $v = f_k(w)$, an unfolding (17) at k of some $w \in (\mathbb{R}^q)^{\mathbb{Z}^2}$; the data matrix $\operatorname{im} \mathbb{D}(\widehat{w})$ of some informative trajectory $\widehat{w} \in \mathfrak{B}$;

Output: An admissible trajectory $w_{\mathcal{L}_{0:N}}$

Step 1: Compute a basis matrix $V \in \mathbb{R}^{\frac{(N+2)(N+3)}{2}q \times d}$ for $\operatorname{im} \mathbb{D}_N(\widehat{w})$

Step 2: Compute a pseudoinverse $V^\dagger$ of V

Step 3: Solve for L_k the equation $v = VL_k$

Step 4: Choose values $p_{k+N+1,-k-N+i} \in \mathbb{R}^q$, $i = -1, 0, 1$, such that

$$Jf_k(w) + \underbrace{\begin{bmatrix} 0 \\ \vdots \\ p_{k+N+1,-k-N-1} \\ 0 \\ \vdots \\ p_{k+N+1,-k-N} \\ \vdots \\ p_{k+N+1,-k-N+1} \end{bmatrix}}_{u_k :=} \in \operatorname{im} V \; .$$

Step 4: Compute L_{k+1} by $L_{k+1} = V^\dagger JVL_k + V^\dagger u_k$.

Step 5: Define $w_{k+N+1,-k-N+i+1}$, $i = 0, 1, \ldots, N$, to be the last q components of every $(m+1)q$-dimensional block of VL_{k+1}, $m = N+1, \ldots, 0$

Step 6: Set $k = k+1$;

Step 7: Set $v = f_k(w)$.

Step 8: Go to Step 3.

Remark 7: In the Algorithm it is not stipulated that the initial unfolding is *legitimate*, i.e. that there exists $\overline{w} \in \mathfrak{B}$ and $k \in \mathbb{Z}$ such that $f_k(\overline{w}) = v$. Consequently, there may not exist a "past" of some system trajectory that is compatible with the given "initial conditions" v. Similarly, the update in Step 4 of the Algorithm uses *arbitrary* values $p_{i,j}$ satisfying the condition $Jf_k(w) + u_k \in \operatorname{im} V$.

It follows that the output of the algorithm is a trajectory that satisfies the dynamical laws of the system *only on the lines* $\mathcal{L}_k$ $0 \leq k \leq N$, but may not be the restriction of a *bona fide* system trajectory to $\mathcal{L}_{0:N}$.

How to construct legitimate unfoldings and "input" vectors u_k are open questions for further research. The first issue is analogous to that studied in [16], namely characterizing "feasible initial conditions" for Fornasini-Marchesini models (see Definition 3.1 therein). ■

Example 3: We use the same system and data of Examples 1 and 2, and we choose $N = 1 = \ell(\mathfrak{B})$. For simplicity of exposition we compute the basis matrix V for $\operatorname{im} \mathbb{D}(\widehat{w})$ from the left-annihilator (16) of $\mathbb{D}(\widehat{w})$:

$$V = \begin{bmatrix} -6 & 12 & 5 & -6 & -5 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

A pseudoinverse of V is

$$V^\dagger = \begin{bmatrix} -\frac{2}{89} & -\frac{10}{89} & -\frac{12}{89} & \frac{10}{89} & \frac{24}{89} & \frac{77}{89} \\ \frac{4}{89} & \frac{20}{89} & \frac{24}{89} & -\frac{20}{89} & \frac{41}{89} & \frac{24}{89} \\ \frac{5}{267} & \frac{25}{267} & \frac{10}{89} & \frac{242}{267} & -\frac{20}{89} & \frac{10}{89} \\ -\frac{2}{89} & -\frac{10}{89} & \frac{77}{89} & \frac{10}{89} & \frac{24}{89} & -\frac{12}{89} \\ -\frac{5}{267} & \frac{242}{267} & -\frac{10}{89} & \frac{25}{267} & \frac{20}{89} & -\frac{10}{89} \end{bmatrix} .$$

To generate a legitimate unfolding, we simulate the data-generating system (6) with a random boundary condition, different from that used in Example 2, generating a new system trajectory w. We choose the fifth column of the data matrix of the trajectory w as initial unfolding v (see Remark 5).

Choosing as u_0 in Step 4 the vector

$$u_0^\top = \begin{bmatrix} 0 & 0 & 0 & 0 & -1.2473 & 0 \end{bmatrix}$$

yields

$$f_1(w)^\top = \begin{bmatrix} 4.8380 & 5.1636 & 0 & 9.1246 & -1.2473 & 0 \end{bmatrix} ;$$

choosing

$$u_1^\top = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.9500 & 0 \end{bmatrix}$$

yields

$$f_2(w)^\top = \begin{bmatrix} 5.1636 & 0 & 0 & -1.2473 & 0.9500 & 0 \end{bmatrix} .$$

Finally, choosing

$$u_2^\top = \begin{bmatrix} 0 & 0 & 0 & 0 & -0.3958 & 0 \end{bmatrix} ,$$

yields

$$v_3^\top = \begin{bmatrix} 0 & 0 & 0 & 0.9500 & -0.3958 & 0 \end{bmatrix} .$$

These computations result in the matrix

$$D_1(w) = \begin{bmatrix} 7.1186 & 4.8380 & 5.1636 & 0 \\ 4.8380 & 5.1636 & 0 & 0 \\ 5.1636 & 0 & 0 & 0 \\ 10.0037 & 9.1246 & -1.2473 & 0.9500 \\ 9.1246 & -1.2473 & 0.9500 & -0.3958 \\ 16.2039 & 0 & 0 & 0 \end{bmatrix} .$$

$D_1(w)$ has the block-Hankel structure expected from a data matrix. Moreover, it can be verified that the vector defined in (16) is a left annihilator of this matrix.

VI. Concluding remarks

We studied the class of discrete autonomous $2D$ systems that admit a state representation where the state at every point $(i, j) \in \mathbb{Z}^2$ is computed from its values at two neighbouring points $(i-1, j)$ and $(i, j-1)$ on the "previous" diagonal line in $\mathbb{Z}^2$. Such systems admit a kernel representation with a square polynomial matrix (see equation (1) and the paper [8]). We defined the property of informativity for identification for such systems (Definition 1). We introduced the concept of *data matrix* (see formulas (10) and (11)), and we stated necessary and sufficient conditions for a given trajectory to be informative in terms of properties of the rank of such matrix (see Theorem 2). Finally, using the "unfolding" technique (see formula (17)), we showed how every trajectory in the system can be computed from a given informative one without explicitly identifying a model (see the Algorithm in Section V).

Given their limitations, the results presented in this paper can only be considered to be preliminary to a thorough investigation of data-driven simulation of $2D$-systems. We restricted our analysis to the class of autonomous $2D$ systems that admit a special representation (1), (2) and (3). It is well known that even among the class of autonomous $2D$ systems there are many that do not admit such a representation. Our current research aims at extending the results presented here to more general classes of nD systems; we expect the representation results from [9], [17] to be useful in such effort. Moreover, pressing research questions are how to generalize our results to the case of noisy data (see Remark 3); and the characterization of "feasible initial conditions" and legitimate updates (see Remark 7).

Acknowledgments

The authors thank the Reviewers for their perspicuous and insightful comments that considerably helped in improving the original manuscript.

References

[1] C. D. 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, 2020.

[2] H. van Waarde, M. Çamlibel, and M. Mesbahi, "From noisy data to feedback controllers: non-conservative design via a matrix S-lemma," *IEEE Trans. Aut. Contr.*, vol. 67, no. 1, pp. 162–175, 2022.

[3] H. van Waarde, M. Çamlibel, J. Eising, and H. Trentelman. (2023) Quadratic matrix inequalities with application to data-based control. [Online]. Available: https://arxiv.org/pdf/2203.12959.pdf

[4] H. J. van Waarde, J. Eising, H. L. Trentelman, and M. K. Çamlibel, "Data informativity: A new perspective on data-driven analysis and control," *IEEE Transactions on Automatic Control*, vol. 65, no. 11, pp. 4753–4768, 2020.

[5] P. Rapisarda, M. K. Çamlibel, and H. J. van Waarde, "Orthogonal polynomial bases for data-driven analysis and control of continuous-time systems," *IEEE Transactions on Automatic Control,* accepted for publication, 2023.

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

[7] P. Rapisarda, M. K. Çamlibel, and H. J. van Waarde, "A persistency of excitation condition for continuous-time systems," *IEEE Control Systems Letters*, vol. 7, pp. 589–594, 2023.

[8] I. Brás and P. Rocha, "State/driving-variable representation of 2D systems," *Multidimensional Systems and Signal Processing*, vol. 13, pp. 129–156, 2002.

[9] D. Pal and H. K. Pillai, "Representation formulae for discrete 2D autonomous systems," *SIAM Journal on Control and Optimization*, vol. 51, no. 3, pp. 2406–2441, 2013.

[10] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," *International Journal of Control*, vol. 81, no. 12, pp. 1946–1959, 2008.

[11] I. Markovsky, J. C. Willems, P. Rapisarda, and B. L. M. De Moor, "Data driven simulation with applications to system identification," *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 970–975, 2005.

[12] P. Rapisarda, H. van Waarde, and M. Çamlibel, "A "fundamental lemma" for continuous-time systems, with applications to data-driven simulation," *Systems & Control Letters,* in press, 2023.

[13] U. Oberst, "Multidimensional constant linear systems," *Acta Applicandae Mathematica*, vol. 20, no. 1, pp. 1–175, 1990.

[14] P. Rocha and J. Willems, "Controllability of 2-D systems," *IEEE Transactions on Automatic Control*, vol. 36, no. 4, pp. 413–423, 1991.

[15] V. G. Lopez and M. A. Müller, "On a continuous-time version of Willems' lemma," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, 2022, pp. 2759–2764.

[16] R. Pereira and P. Rocha, "Feasible initial conditions for 2D discrete state-space systems," *IFAC-PapersOnLine*, vol. 55, no. 30, pp. 127–131, 2022.

[17] M. Mukherjee and D. Pal, "On minimality of initial data required to uniquely characterize every trajectory in a discrete n-D system," *SIAM Journal on Control and Optimization*, vol. 59, no. 2, pp. 1520–1554, 2021.