Left-invertibility under sparse assumption: the linear case

Jean-Pierre Barbot, Krishna Busawon, Christopher Edwards

Left-invertibility under sparse assumption: the linear case

J-P. Barbot[*], K Busawon[†] and C. Edwards[‡]

July 7, 2021

Abstract

This paper deals with the issue of dynamical left-invertibility for linear continuous-time dynamical systems. More precisely, we provide sufficient conditions in order to estimate unknown inputs, using known outputs, under sparse input assumptions for linear continuous-time dynamical systems that are not necessarily square. In fact, there exists an algorithm in the literature that allows verification of left-invertibility for linear square systems; that is, systems with p known outputs and p unknown inputs. However, a similar algorithm does not exist for the rectangular case, where the number of inputs is much larger than the number of outputs. In this paper, we first use the square case algorithm as a stepping stone in order to propose a new algorithm for the rectangular case. However, it shown that even if the proposed algorithm converges successfully, it is not sufficient to estimate the unknown inputs. Consequently, it has been deemed necessary to include a sparse input assumption and to verify the well-known Restrictive Isometric Property (RIP) conditions of a specific matrix. Finally, an academic example is given in order to highlight the feasibility of the proposed approach.

[*]Jean-Pierre Barbot is in secondment at LS2N UMR 6004 CNRS, École Centrale de Nantes and with Quartz EA 7393, ENSEA France.

[†]Krishna Busawon is with the Faculty of Engineering and Environment, Northumbria University, UK.

[‡]Christopher Edward is with the University of Exeter, Exeter, UK.

1 Introduction

In the context of control systems, the problem of Dynamical Left-Invertibility (DLI) refers to the estimation of unknown inputs from known outputs (measurements) for dynamical systems. In this work, we investigate the DLI problem for multi-input multi-output linear continuous-time dynamical systems whereby the number of unknown inputs is much larger than the number of outputs, with the additional property that the number of non-zero unknown inputs are s-sparse (i.e. only s unknown inputs are non-zero). One of the motivations for considering such classes of system is that some diagnosis or cyber attack detection problems can be treated as a left-invertible problem under sparse assumptions [8]. Moreover, in the context of over-actuated systems or networks, left-invertibilty under sparsity assumptions can generate a control design with a minimum of active actuators.

The main aim of this paper is to provide sufficient conditions for the estimation of the unknown inputs, under sparsity assumption on the latter. For this, we start by extending the left-invertibility algorithm presented in [11] for square linear systems (see also for an other approach [4] and [2] for the nonlinear case) to rectangular linear systems. Prior to that, some essential definitions are introduced; namely, uniform observability, s-sparse left-invertibility and relative degree with respect to the unknown inputs to be estimated, which provide the necessary ingredients to develop the proposed algorithm. After that, from the seminal works of Tao, Candes, Donoho [6, 8, 10] and other authors from the signal processing community (see references herein), a specific exact input recovery method is employed in order to solve the problem of left-invertibility for linear dynamical systems under sparse assumptions. Finally, some conclusions are drawn together with some scope for future works.

Notation: For any time function $z(t)$ that is at least of class $\mathcal{C}^j$, $\dot{z}$ and $\ddot{z}$ denotes the first and second derivative with respect to time t and $z^{(j)}$ denotes the j^{th} derivative of z with respect to t.
Uppercase letters are employed to denote vectors or matrices and lowercase letters to denote scalars.
For any two vectors U and V, we shall denote by $(U, V) = (U \quad V)$ and by $(U; V) = \begin{pmatrix} U \\ V \end{pmatrix}$.

2 Preliminary discussions and problem statement

Dynamical left-invertibility (DLI) is encountered in many engineering fields and various results on this topic exist in the literature. Hereafter, we shall present the DLI problem for linear systems with more unknown inputs that outputs under sparse assumptions.

We consider linear continuous-time systems of the following form:

$$(\Sigma_M :) \left\{ \begin{array}{rcl} \dot{X} & = & AX + BU + D^1 W_x \\ Y & = & CX + D^2 W_y + N \end{array} \right. \tag{1}$$

where $X \in \mathbf{R}^n$ is the state, $Y \in \mathbf{R}^p$ is the output, $U \in \mathbf{R}^m$ is the known input, $W_x \in \mathbf{R}^{q_x}$ is the unknown input on the state dynamics, $W_y \in \mathbf{R}^{q_y}$ is the unknown input with respect to the output, and finally $N \in \mathbf{R}^p$ is the measurement noise vector.
Throughout the paper, we assume that the vector N is composed of p uncorrelated white noise, $q := q_x + q_y >> p$.

Moreover, the following assumptions are made:

Assumption 1 *$Rank\{C\} = p$ and without loss of generality Y is considered to be equal to $(x_1; x_2; ...; x_p)$.*

Assumption 2 *The unknown input vector W_y with respect to the output is assumed to be constant or at least slowly varying i.e. $W_{q_y}^{(j)} \simeq 0$ for all $j \in \{1, ..., n-1\}$.*

Assumption 3 *$W = (W_x; W_y)$ is an s-sparse vector (i.e. s is the maximum number of non-zero component of W)*

Now, it is necessary to make some definitions with respect to observability and left-invertibility.

Definition 2.1 *The system (1) is said to be uniformly observable with respect to W_x, if it is possible to obtain $x(t)$ with any desired convergence time from the knowledge of its input U and its output Y at any positive time when $W_y = 0$.*

Definition 2.2 *The system (1) is said to be s-sparse left-invertible with respect to W_x, if it is possible to obtain $x(t)$ and W_x with any desired convergence time from the knowledge of its input U and its output Y at any positive time when $W_y = 0$.*

And finally,

Definition 2.3 *The system (1) is said to be s-sparse left-invertible with respect to W, if it s-sparse left-invertible with respect to W_x and for $W_y \neq 0$ it is possible, to obtain W with any desired convergence time.*

From the previous definitions, it is obvious that the s-sparse left-invertibility with respect to W implies s-sparse left-invertibility with respect to W_x. Similarly, left-invertibility with respect to W_x implies the uniform observability with respect to W_x and, finally, uniform observability with respect to W_x implies the classical observability.

Now, it is necessary to introduce and adapt some computational tools in the context of dynamical left-invertibility based on the above definitions.

Definition 2.4 *The relative degree of y_i, $i \in \{1, ..., p\}$ with respect to W_x for the system (Σ_M) (1) is the smallest integer $\rho_i > 0$ such that $\exists j \in \{1, ..., q_x\}$ such that $C_i A^{\rho_i - 1} D_j^1 \neq 0$.*

It is obvious that, from the previous definition, it is possible to define the following matrix:

$$\Gamma_{q_x} = \begin{pmatrix} C_1 A^{\rho_1 - 1} D_1^1 & C_1 A^{\rho_1 - 1} D_2^1 & ... & C_1 A^{\rho_1 - 1} D_{q_x}^1 \\ C_2 A^{\rho_2 - 1} D_1^1 & C_2 A^{\rho_2 - 1} D_2^1 & ... & C_2 A^{\rho_2 - 1} D_{q_x}^1 \\ \vdots & \vdots & \vdots & \vdots \\ C_p A^{\rho_p - 1} D_1^1 & C_p A^{\rho_p - 1} D_2^1 & ... & C_p A^{\rho_p - 1} D_{q_x}^1 \end{pmatrix} \tag{2}$$

The following preliminary proposition provides an aid for the adaptation of the left-invertibility algorithm presented in [11] to the above rectangular case.

Proposition 2.1 *The system (1) is uniformly observable with respect to W_x, if $\rho := \sum_{i=1}^{p} \rho_i = n$ and rank of Γ_{q_x} is equal to p.*

Proof : As $\rho := \sum_{i=1}^{p} \rho_i = n$ the above state transformation matrix T is composed of the following n independent co-vectors :

$$T = \begin{pmatrix} C_1 \\ C_1 A \\ \vdots \\ C_1 A^{\rho_1 - 1} \\ \vdots \\ C_p \\ C_p A \\ \vdots \\ C_p A^{\rho_p - 1} \end{pmatrix} \tag{3}$$

As a result, the system in the $Z := TX$ coordinates becomes:

$$\begin{cases} \begin{pmatrix} \dot{z}_{1,1} \\ \vdots \\ \dot{z}_{\rho_1 - 1,1} \\ \dot{z}_{\rho_1,1} \end{pmatrix} = \begin{pmatrix} z_{2,1} + C_1 BU \\ \vdots \\ z_{\rho_1,1} + C_1 A^{\rho_1 - 2} BU \\ C_1 A^{\rho_1 - 1}(AT^{-1}Z + Bu + DW_x) \end{pmatrix} \\ \cdots \qquad \cdots \qquad \cdots \\ \begin{pmatrix} \dot{z}_{1,p} \\ \vdots \\ \dot{z}_{\rho_p - 1,p} \\ \dot{z}_{\rho_p,p} \end{pmatrix} = \begin{pmatrix} z_{2,p} + C_p BU \\ \vdots \\ z_{\rho_p,p} + C_p A^{\rho_p - 2} BU \\ C_p A^{\rho_p - 1}(AT^{-1}Z + Bu + D^2 W_x) \end{pmatrix} \end{cases} \tag{4}$$

which may be rewritten as follows:

$$\begin{pmatrix} \dot{z}_{1,1} \\ \vdots \\ \dot{z}_{\rho_1 - 1,1} \\ \vdots \\ \dot{z}_{1,p} \\ \vdots \\ \dot{z}_{\rho_p - 1,1} \end{pmatrix} = \begin{pmatrix} z_{2,1} + C_1 BU \\ \vdots \\ z_{\rho_1,1} + C_1 A^{\rho_1 - 2} BU \\ \vdots \\ z_{2,p} + C_p BU \\ \vdots \\ z_{\rho_p,p} + C_p A^{\rho_p - 2} BU \end{pmatrix} \tag{5}$$

$$\begin{pmatrix} \dot{z}_{\rho_1,1} \\ \vdots \\ \dot{z}_{\rho_p,p} \end{pmatrix} \qquad \begin{pmatrix} C_1 A^{\rho_1 - 1}(AT^{-1}Z + BU) \\ \vdots \\ C_p A^{\rho_p - 1}(AT^{-1}Z + BU) \end{pmatrix} + \Gamma_{q_x} W_x$$

As U is known and W_x is bounded it is possible to design an observer (for example sliding mode [17, 18, 12] or a high gain [3, 5]) in order to estimate, the state Z. □

The result of Proposition 2.1 is a particular case, in fact some output dynamics can be independent of the unknown input W_x and some states cannot be estimated directly. The next algorithm considers all these cases.

3 Uniform observability with respect to W_x

In this section, a modified version of the algorithm introduced in [11] is proposed to determine whether or not the system (1) is uniformly observable with respect to W_x.

Algorithm I

Step 1: Consider the output vector $y^{\underline{1}} = y = CX$, and $p^1 = p$.

a. Without loss of generality, one can reorder the components of $y^{\underline{1}}$ as follows

$$y^{\underline{1}} = \left[\; C_1^T; \;\; \cdots ; \;\; C_{\eta^1}^T; \;\; C_{\eta^1+1}^T; \;\; \cdots ; \;\; C_{p^1}^T \; \right]^T X \tag{6}$$

so that for $1 \leq j \leq \eta^1$:

$$C_j A^k D^1 = 0, \text{for all } k \in \mathbb{N} \tag{7}$$

and for $1 \leq j \leq p^1 - \eta^1$, there exists an integer ρ_j^1 (i.e. the relative degree associated to the j^{th} output; see Definition 2.4) such that:

$$\begin{cases} C_{\eta^1+j} A^k D^1 = 0, & \forall k < \rho_j^1 - 1 \\ C_{\eta^1+j} A^{\rho_j^1 - 1} D^1 \neq 0, & \forall k < \rho_j^1 - 1 \end{cases} \tag{8}$$

Note that the outputs $y_j^{\underline{1}} = C_j X$, $j \leq \eta^1$, are not affected by the unknown inputs.

b. Compute the set of covectors

$$\Phi^1 = \text{span} \left\{ C_1; ...; C_1 A^{n-1}; ...; C_{\eta^1}; ...; C_{\eta^1} A^{n-1} \right\} \tag{9}$$

and denote the dimension of Φ^1 by φ^1. Define the integers φ_j^1, $1 \leq j \leq \eta^1$ such that

$$I^1 = \left\{ C_1; ...; C_1 A^{\varphi_1^1 - 1} ...; C_{\eta^1}; ...; C_{\eta^1} A^{\varphi_{\eta^1}^1 - 1} \right\} \tag{10}$$

is a basis of Φ^1. A direct consequence of the previous definition is that $\varphi^1 = \varphi_1^1 + \ldots + \varphi_{\eta^1}^1$. Note that the choice of the φ_i^j is not unique [16]. **If** $\varphi^1 = n$, it is possible to obtain all the state variables of system (1) using a classical observer (Luenberger [20], Kalman and Bucy [15], ...), **then the algorithm stops and the system is uniformly observable with respect to** W_x. Note that in this case $D^1 = 0$.

c. **If** $\varphi^1 < n$, let us define the matrix

$$\Gamma^1 = \begin{bmatrix} C_{\eta^1+1} A^{\rho_1^1 - 1} D^1 \\ \vdots \\ C_p A^{\rho_{p-\eta^1}^1 - 1} D^1 \end{bmatrix} \tag{11}$$

and the rank of note Γ^1 is denoted by γ^1. Compute the covectors space

$$\begin{aligned} \Upsilon^1 \quad = \quad & span \left\{ C_{\eta^1+1}; ...; C_{\eta^1+1} A^{\rho_1^1 - 1}; \right. \\ & \left. ...; C_p; ...; C_p A^{\rho_{p-\eta^1}^1 - 1} \right\} \end{aligned} \tag{12}$$

and the dimension of the direct sum of Φ^1 with Υ^1 (i.e. dimension $\{\Phi^1 \oplus \Upsilon^1\}$ is equal to $\varphi^1 + \delta^1$, where δ^1 is equal to the dimension of Υ^1. Define the integers $(\delta_1^1, ..., \delta_{p-\eta^1}^1)$ such that (after a possible reordering of the C_i, $\eta^1 + 1 \leq i \leq p$), the set $\{I^1; J^1\}$, where

$$J^1 = \left\{ C_{\eta^1+1}; ...; C_{\eta^1+1} A^{\delta_1^1 - 1}; , ...; C_p; ...; C_p A^{\delta_{p-\eta^1}^1 - 1} \right\} \tag{13}$$

is a basis of $\Phi^1 \oplus \Upsilon^1$.
From (13) δ^1 is equal to $\sum_{l=1}^{p-\eta_1} \delta_l^1$.
If $\varphi^1 + \delta^1 = n$**, the system is uniformly observable with respect to** W_x, since the observability matching condition is satisfied and it is possible to estimate all the state variables using, for example, high gain observer [14, 13] or sliding mode observer [26, 23].

d. **If $\varphi^1 + d^1 < n$, the algorithm continues**. Then, define the matrix $\Lambda^1 \in \mathbb{R}^{p_2 \times (p-\eta_1)}$ of rank p^2, where $p^2 = p^1 - \eta^1 - \gamma^1$, such that

$$\Lambda^1 \Gamma^1 = 0_{p_2 \times (p^1 - \eta^1)} \tag{14}$$

Define also the auxiliary variable (dummy output)

$$y^{\underline{2}} = \Lambda^1 \begin{bmatrix} C_{\eta^1+1} A^{\rho_1^1} \\ \vdots \\ C_p A^{\rho_{p-\eta_1}^1} \end{bmatrix} x := C^2 x,\ C^2 := \begin{bmatrix} C_1^2 \\ \vdots \\ C_{p^2}^2 \end{bmatrix}$$

and compute the rank of $\{I^1; J^1; C^2 A\}$ if this rank is equal to the Rank of $\{I^1; J^1\}$, this means that their are not new direction generated by the dummy output then
if rank$\{I^1; J^1; C^2 A\}$ = rank$\{I^1; J^1\}$ the system (1) is not uniformly observable with respect to W_x
else go to step 2.

Step 2: Bearing in mind that the covector space Φ^1 and Υ^1 generated by C and A, the procedure employed for $y^{\underline{1}}$ is now applied to $y^{\underline{2}} \in \mathbb{R}^{p_2}$.

a. After possible reordering of the components of $y^{\underline{2}}$, by analogy with Step 1.a), one can define the integers η^2 and ρ_j^2, $1 \leq j \leq p_2 - \eta^2$ so that the η^2 first components of $y^{\underline{2}}$ are not influenced by the unknown inputs and the $p^2 - \eta^2$ components of $y^{\underline{2}}$ have relative degree equal to ρ_j^2.

b. The covector space Φ^2 is defined as follows:

$$\Phi^2 = \text{span}\left\{C_1^2; ...; C_1^2 A^{n-1}; C_2^2; ; ...; C_{\eta^2}^2 A^{n-1}\right\}. \tag{15}$$

Similarly to b) of step 1, φ^2 is defined such that $\dim(\Phi^1 \oplus \Upsilon^1 \oplus \Phi^2) = \varphi^1 + \delta^1 + \varphi^2$. Then, define the integers φ_j^2, $1 \leq j \leq \eta^2$ and the related set I^2 such that the set $\{I^1; J^1; I^2\}$ is a basis of $\Phi^1 \oplus \Upsilon^1 \oplus \Phi^2$.

If $\varphi^1 + \delta^1 + \varphi^2 = n$ the system (1) is uniformly observable with respect to W_x

c. **If** $\varphi^1 + \delta^1 + \varphi^2 < n$, then define the matrix

$$\Gamma^2 = \begin{bmatrix} \Gamma^1 \\ C^2_{\eta^2+1} A^{\rho^2_1 - 1} D^1 \\ \vdots \\ C^2_{p^2} A^{\rho^2_{p_2-\eta^2}{}^{-1}} D^1 \end{bmatrix} \tag{16}$$

and denote $\gamma^2 = \text{rank}\, \Gamma^2$. By analogy with the c) of step 1, it is defined the covector space

$$\begin{aligned} \Upsilon^2 \;\; = \;\; & span\{C^2_{\eta^2+1}; ...; C^2_{\eta^2+1} A^{\rho^2_1 - 1}; \\ & ...; C^2_{p^2}; ...; C^2_{p^2} A^{\rho^2_{p^2-\eta^2}{}^{-1}}\} \end{aligned} \tag{17}$$

and

$$J^2 = \{C^2_{\eta^2+1}; ...; C^2_{\eta^2+1} A^{\delta^2_1}; ...; C^2_{p^2}; ...; C^2_{p^2} A^{\delta^2_{p^2-\eta^2}}\} \tag{18}$$

such that $\{I^1; J^1; I^2; J^2\}$ is a basis of $\Phi^1 \oplus \Upsilon^1 \oplus \Phi^2 \oplus \Upsilon^2$ and δ^2 is the rank of J^2 (i.e. $\delta^2 = \delta^2_1 + ... + \delta^2_{p^2-\eta^2}$).
If $\varphi^1 + d^1 + \varphi^2 + d^2 = n$ **the system (1) is uniformly observable with respect to** W_x**.**

d. **If** $\varphi^1 + \delta^1 + \varphi^2 + \delta^2 < n$ **the algorithm continues** and the matrix $\Lambda^2 \in \mathbb{R}^{p^3 \times (\left(p^1 - \eta^1\right) + \left(p^2 - \eta^2\right))}$ is defined, where $p^3 = (p^2 - \eta^2) - \gamma^2$, such that

$$\Lambda_2 \Gamma_2 = 0 \tag{19}$$

where Λ^2 is composed of p^3 independent co-vectors. Then a new auxiliary variable (dummy output) is introduced

$$Y^{\underline{3}} = \Lambda^2 \begin{bmatrix} C_{\eta^1+1} A^{\rho^1_1} \\ \vdots \\ C_{p^1} A^{\rho^1_{p^1-\eta^1}} \\ C^2_{\eta^2+1} A^{\rho^2_1} \\ \vdots \\ C^2_{p^2} A^{\rho^2_{p^2-\eta^2}} \end{bmatrix} X := \; C^3 X. \tag{20}$$

If rank$\{I^1, J^1, I^2, C^3 A\}$ = **rank**$\{I^1, J^1, I^2\}$**, the system (1) is not uniformly observable with respect to** W_x**,**

else the algorithm continues and go to step 3.

Repeating the above procedure, one has:
Step k: The fictitious output $y^{\underline{k}} \in \mathbb{R}^{p^k}$, that has been defined in Step $k-1$, is considered.

a. Determine integers η^k and ρ_j^k $1 \leq j \leq p^k - \eta^k$.

b. Compute the co-vectors space

$$\begin{aligned} \Phi^k &= span\left\{C_1^k; ...; C_1^k A^n, C_2^k; \right. \\ &\quad \left. ...; C_2^k A^n; ...; C_{\eta^k}^k; ...; C_{\eta^k}^k A^n\right\} \end{aligned} \tag{21}$$

and define I^k as the following set of co-vectors

$$I^k = \left\{C_1^k; ...; C_1^k A^{\varphi_1^k - 1}; ...; C_{\eta^k}^k; ...; C_{\eta^k}^k A^{\varphi_{\eta^k}^k - 1}\right\}, \tag{22}$$

such that $\left(\bigcup_{i=1}^{k-1} I^i \cup J^i\right) \cup I^k$ is a basis of $\left(\bigcup_{i=1}^{k-1} \Phi^i \cup \Upsilon^i\right) \cup \Phi^k$. So, the relation

$$\text{rank}\left(\left(\bigcup_{i=1}^{k-1} \Phi^i \cup \Upsilon^i\right) \cup \Phi^k\right) = \sum_{i=1}^{k-1} \left(\varphi^i + \rho^i\right) + \varphi^k \tag{23}$$

is verified with $\varphi^k = \varphi_1^k + \ldots + \varphi_{\eta_k}^k$.

c. Define

$$\Gamma^k = \begin{bmatrix} \Gamma^1 \\ \vdots \\ \Gamma^{k-1} \\ C_{\eta^k+1}^k A^{\rho_1^k - 1} \\ \vdots \\ C_{p^k}^k A^{\rho_{p^k - \eta^k}^k - 1} \end{bmatrix} \tag{24}$$

and denote by γ^k the rank of Γ^k. By analogy with c) in all the previous steps, it is defined the co-vector space ;

$$\begin{aligned} \Upsilon^k &= span\left\{C_{\eta^k+1}^k; ..., C_{\eta^k+1}^k A^{\rho_1^k - 1}; \right. \\ &\quad \left. ...; C_{p^k}^k; ...; C_{p^k}^k A^{\rho_{p^k - \eta^k}^k - 1}\right\} \end{aligned} \tag{25}$$

and

$$J^k = \{C^k_{\eta^k+1}; ...; C^k_{\eta^k+1}A^{\delta^k_1-1}; ...; C^k_{p^k}; ...; C^k_{p^k}A^{\delta^k_{p^k-\eta^k}-1}\} \tag{26}$$

such that $\{I^1; J^1; I^2; J^2; ..; I^K; J^k\}$ is a basis of $\Phi^1 \oplus \Upsilon^1 \oplus \Phi^2 \oplus \Upsilon^2 \oplus ... \oplus \Phi^k \oplus \Upsilon^k$ and δ^k is the rank of J^k (i.e $\delta^k = \delta^k_1 + ... + \delta^k_{p^k-\eta^k}$).
If $\varphi^1 + \delta^1 + \varphi^2 + \delta_2 + ... + \varphi^k + \delta^k = n$ the system is uniformly observable with respect to W_x.

d. **If $\varphi^1 + \delta^1 + \varphi^2 + \delta^2 + ... + \varphi^k + \delta^k < n$ the algorithm continues** and a matrix $\Lambda^k \in \mathbb{R}^{p^3 \times (\left(p^1-\eta^1\right)+\left(p^k-\eta^k\right))}$ is defined, where $p^{k+1} = \left(p^k - \eta^k\right) - \gamma^k$, such that

$$\Lambda_k \Gamma_k = 0 \tag{27}$$

where Λ^k is composed of p^k independent co-vectors, which has at least one of this last term $> \eta^k$ non null, (or equivalently the rank of Λ^k is equal to p^k). Then a new auxiliary variable (dummy output) is introduced

$$Y^{\underline{k+1}} = \Lambda^k \begin{bmatrix} C_{\eta^1+1}A^{\rho^1_1} \\ \vdots \\ C_{p^1}A^{\rho^1_{p^1-\eta^1}} \\ \vdots \\ C^k_{\eta^k+1}A^{\rho^k_1} \\ \vdots \\ C^k_{p^k}A^{\rho^k_{p^k-\eta^k}} \end{bmatrix} X := \ C^{k+1}X. \tag{28}$$

if $rank\{I^1, J^1.., I^k, C^{k+1}A\} = rank\{I^1, J^1.., I^k\}$, the system (1) is not uniformly observable with respect to W_x.
else go to step $k+1$

It is important to note that, as at d) for each step, the dummy output must add at least one new direction, then **the algorithm stops before step** $n - p + 1$. Moreover, in the algorithm p^l can be equal to zero. In this case no new dummy output is generated and consequently $\text{rank}\{I^1, J^1, ..., I^l, C^{l+1}A\} = \text{rank}\{I^1, J^1, ..., I^l\}$ and the algorithm concludes that the system is not uniformly observable with respect to W_x.

4 Structural form for left-invertibility with respect to W

In this section, a structural form is given in order to define the left-invertibility problem with respect to W as a "classical" optimization problem.
From the previous algorithm, first of all we reorder the bases $\{I^1; J^1; ...; I^k; J^K\}$ as follows $\{I^1; ..., I^k, J^1; ...; J^K\}$ where J^k is empty if the algorithm stops in b) of step k. After that, it is possible to introduce the following change of coordinates $Z = TX$:

$$T = \begin{pmatrix} C_1 \\ \vdots \\ C_1 A^{\varphi_1 - 2} \\ \vdots \\ C_{\eta_t} \\ \vdots \\ C_{\eta_t} A^{\varphi_{\eta_t} - 2} \\ C_{\eta_t + 1} \\ \vdots \\ C_{\eta_t + 1} A^{\delta_{\eta_t + 1} - 2} \\ \vdots \\ C_{p'} \\ \vdots \\ C_{p'} A^{\delta_{p'} - 2} \\ C_1 A^{\varphi_1 - 1} \\ \vdots \\ C_{\eta_t} A^{\varphi_{\eta_t} - 1} \\ C_{\eta_t + 1} A^{\delta_{\eta_t + 1} - 1} \\ \vdots \\ C_{p'} A^{\delta_{p'} - 1} \end{pmatrix} \tag{29}$$

with $\eta_t = \sum_{i=1}^{k} \eta_i$ and $p' = \eta_t + Car_\delta$ where Car_δ is the cardinality of all δ_i^j. Considering that the $z_{i,j}$ come directly from the sensors and, hence, are

affected by the unknown inputs W_y, the change of coordinate yelds:

$$\begin{pmatrix} \dot{z}_{1,1} \\ \vdots \\ \dot{z}_{\varphi_1-1,1} \\ \vdots \\ \dot{z}_{1,\eta_t} \\ \vdots \\ \dot{z}_{\varphi_{\eta_t}-1,\eta_t} \\ \dot{z}_{1,\eta_t+1} \\ \vdots \\ \dot{z}_{\delta_{\eta_t+1}-1,\eta_t+1} \\ \vdots \\ \dot{z}_{1,p'} \\ \vdots \\ \dot{z}_{\delta_{p'-1},p'} \end{pmatrix} = \begin{pmatrix} z_{2,1} + C_1 BU \\ \vdots \\ z_{\varphi_1,1} + C_1 A^{\varphi_1-2} BU \\ \vdots \\ z_{2,\eta_t} + C_{\eta_t} BU \\ \vdots \\ z_{\varphi_t,\eta_t} + C_{\eta_t} A^{\varphi_t-2} BU \\ z_{2,\eta_t+1} + C_{\eta_t+1} BU \\ \vdots \\ z_{\delta_{\eta_t+1},\eta_t+1} + C_{\eta_t+1} A^{\delta_{\eta_t+1}-2} BU \\ \vdots \\ z_{2,p'} + C_{p'} BU \\ \vdots \\ z_{\delta_{p'},p'} + C_p A^{\delta_{p'}-2} BU \end{pmatrix} \tag{30}$$

$$\begin{pmatrix} \dot{z}_{\varphi_1,1} \\ \vdots \\ \dot{z}_{\varphi_{\eta_t},\eta_t} \\ \dot{z}_{\delta_{\eta_t}+1,\eta_t+1} \\ \vdots \\ \dot{z}_{\delta_{p'},p'} \end{pmatrix} = \begin{pmatrix} C_1 A^{\varphi_1-1}(AT^{-1}Z + BU) \\ \vdots \\ C_{\eta_t} A^{\varphi_{\eta_t}-1}(AT^{-1}Z + BU) \\ C_{\eta+1} A^{\delta_{\eta+1}-1}(AT^{-1}Z + BU) \\ \vdots \\ C_{p'} A^{\delta_{p'}-1}(AT^{-1}Z + BU) \end{pmatrix} + \Gamma W \tag{31}$$

where Γ is equal to:

$$\Gamma = \begin{pmatrix} -\Gamma_{1,1} & 0 \\ -\Gamma_{2,1} & \Gamma_{2,2} \end{pmatrix}$$

with

$$\Gamma_{1,1} = \begin{pmatrix} C_1 A^{\varphi_1} D^{2,ex} \\ \vdots \\ C_{\eta_t} A^{\varphi_{\eta_t}} D^{2,ex} \end{pmatrix}$$

$$\Gamma_{2,1} = \begin{pmatrix} C_{\eta_t+1} A^{\delta_{\eta_t+1}} D^{2,ex} \\ \vdots \\ C_{p'} A^{\delta_{p'}} D^{2,ex} \end{pmatrix}$$

$$\Gamma_{2,2} = \begin{pmatrix} C_{\eta_t+1} A^{\delta_{\eta_t+1}} D^1 \\ \vdots \\ C_{p'} A^{\delta_{p'}} D^1 \end{pmatrix}$$

and $D^{2,ex} \in \mathbf{R}^{n \times q_y}$ is defined as follows:

$$D^{2,ex} := \begin{pmatrix} D^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{32}$$

Then, equation (31) gives

$$\bar{Y} = \Gamma W \tag{33}$$

where $\bar{Y}$ is composed of the estimated state, the derivative of the state and the known input

$$\bar{Y} = \begin{pmatrix} \dot{z}_{\varphi_1,1} \\ \vdots \\ \dot{z}_{\varphi_{\eta_t},\eta_t} \\ \dot{z}_{\delta_{\eta_t}+1,\eta_t+1} \\ \vdots \\ \dot{z}_{\delta_{p'},p'} \end{pmatrix} - \begin{pmatrix} C_1 A^{\varphi_1 - 1}(AT^{-1}Z + BU) \\ \vdots \\ C_{\eta_t} A^{\varphi_{\eta_t} - 1}(AT^{-1}Z + BU) \\ C_{\eta+1} A^{\delta_{\eta+1} - 1}(AT^{-1}Z + BU) \\ \vdots \\ C_{p'} A^{\delta_{p'} - 1}(AT^{-1}Z + BU) \end{pmatrix} \tag{34}$$

$\bar{Y}$ is available for example from sliding mode observer [22, 19]. Now, from (33), it is possible to determine if Γ verifies the RIP condition (defines in the next section), for the prefixed s-sparsity.

5 Exact recovery for dynamical system

The objective of this method is to reconstruct the most parsimonious solution of the huge unknown input vector $W(t)$ from the small measurements vector $\bar{Y}$ and the matrix Γ. This leads to the following optimization problem:

$$\min_{W \in \mathbb{R}^q} \|W\|_0, \quad \text{under the constraint} \quad \bar{Y} = \Gamma W, \tag{35}$$

where $\|W\|_0$ is a pseudo-norm and corresponds to the number of non-zero values of W.

The problem in (35) can be rewritten as follows

$$\min_{W \in \mathbf{R}^q} \{\frac{1}{2}\|\bar{Y} - \Gamma W\|_2^2 + \lambda \|W\|_0\}. \tag{36}$$

where λ is a balancing parameter (see, for example, [7]):

In order to solve this problem a preliminary assumptions is made:

Assumption 4 *$\bar{Y} \in \mathbb{R}^{p'}$ verifies:*

$$2s + 1 \leq p'. \tag{37}$$

Nevertheless, the optimization problem using the pseudo-norm zero in (36) remains a very difficult one and it is considered as an N.P-complete problem. However, from the seminal works [8, 6, 25, 10] if the matrix Γ satisfies the following property, the so-called Restricted Isometry Property (RIP), problem (36) can be transformed into an equivalent one where the pseudo-norm $\|.\|_0$ is replaced by the norm $\|.\|_1$.

A matrix Γ is said to satisfy the s-order RIP condition, if for any s-sparse signal W, the following condition is verified:

$$(1 - \delta_s)\|W\|_2^2 \leq \|\Gamma W\|_2^2 \leq (1 + \delta_s)\|W\|_2^2, \tag{38}$$

where $\delta_s \in (0, 1)$ is a constant parameter. This property will be called the s-order RIP condition.

There exist some conditions in order to check if a matrix Γ satisfies the RIP condition [8, 27], but this is out of the scope of this paper. If Γ of relation (33) does not verify the RIP condition, it is possible to reduce the number s or/and to cluster some unknown inputs (see for example [28]), but this is also out of the scope of the paper. Consequently, the following assumption is made:

Assumption 5 *The matrix Γ satisfies the s-sparse RIP.*

[7] gives

Proposition 5.1 *Under Assumption 5, the parsimony problem (36) becomes:*

$$\min_{W \in \mathbf{R}^q} \{\frac{1}{2}\|\bar{Y} - \Gamma W\|_2^2 + \lambda \|W\|_1\}, \tag{39}$$

and its solution exists and is unique.

From this proposition it is possible to conclude as follow:

Theorem 5.1 *The system (8) is s-parse left invertible with respect to W if the system is informally observable with respect to W_x (algorithm 1) and the assumption 5 is verified.*

From this Theorem, it is possible to use a continuous optimisation algorithm in order to obtain W in real time. Hereafter, in order to solve (39), the following dynamical algorithm, based on sliding mode techniques [24], proposed by [29] (see also [1] for exponential convergence instead of finite time convergence) is used:

$$\begin{cases} \tau \dot{u} = -\lceil u + (\Gamma^T \Gamma - I_d)\hat{W} - \Gamma^T \bar{Y} \rfloor^\alpha \\ \hat{W} = \varphi_\lambda(u(t)), \end{cases}, \tag{40}$$

where $\lceil . \rfloor^\alpha = |.|^\alpha sign(.)$, $u(t) \in \mathbb{R}^q$ is the internal state vector, $\varphi_\lambda(u(t)) = \max(|u| - \lambda, 0) sign(u)$ is the continuous soft thresholding function, $\lambda \in \mathbb{R}^q$ is a vector of constant positive parameters that has to be suitably chosen based on the noise and the minimum possible absolute values of the disturbances, the vector $\hat{W}$ represents the estimation of the sparse signal W, τ is a $(q \times q)$ diagonal matrix with constant parameters, τ_i, with $i \in \{1, ..., q\}$, determined by the physical properties of the implementing system, I_d is the $(q \times q)$ identity matrix and $0 \leq \alpha \leq 1$ represents the exponential coefficient.

In many recent applications this algorithm was successfully applied for diagnostics [9], Cyber attacks [21], and in signal processing [28].

6 An academic example

Let us consider the following academic system:

$$\begin{aligned} \dot{x}_1 &= -x_1 + B_{1,1} u_1 \\ \dot{x}_2 &= x_3 + W_{x,1} + 0.5 W_{x,2} + W_{x,3} - W_{x,4} \\ \dot{x}_3 &= -x_2 + B_{3,2} u_2 + W_{x,1} + W_{x,2} + 0.5 W_{x,3} + W_{x,4} \\ \dot{x}_4 &= x_5 \\ \dot{x}_5 &= -3 x_4 + x_5 + B_{5,3} u_3 + W_{x,1} \\ &+ 1.5 W_{x,2} + 0.75 W_{x,3} + 0.5 W_{x,4} \end{aligned} \tag{41}$$

with outputs (sensors)

$$\begin{aligned} y_1 &= x_1 + W_{y,1} \\ y_2 &= x_2 + W_{y,2} \\ y_3 &= x_3 + W_{y,3} \\ y_4 &= x_4 + W_{y,4} \end{aligned} \tag{42}$$

where the $W_{x,i}$ for $i \in \{1,..,4\}$ are the unknown inputs affecting the state dynamics, and the $W_{y,i}$ for $i \in \{1,..,4\}$ are constant (or slowly varying or piece-wise constant).
The step 1 of the algorithm I section 3 gives:

$$\begin{aligned} I^1 &= \{(1,0,0,0,0)\} \\ J^1 &= \left\{ \begin{array}{ccccc} (0 & 1 & 0 & 0 & 0) \\ (0 & 0 & 1 & 0 & 0) \\ (0 & 0 & 0 & 1 & 0) \\ (0 & 0 & 0 & 0 & 1) \end{array} \right\} \end{aligned} \tag{43}$$

This implies that $\varphi^1 + \delta^1 = 5 = n$, then the system (41) is uniform observable with respect to W_x. Consequently, it is possible to obtain the equation (33), with

$$\bar{Y} = \begin{pmatrix} \dot{y}_1 + y_1 - B_{1,1}u_1 \\ \dot{y}_2 - y_3 \\ \dot{y}_3 + y_2 - B_{3,2}u_2 \\ \ddot{y}_4 - \dot{y}_4 + 3y_4 - B5,3u_3 \end{pmatrix} \tag{44}$$

and

$$\Gamma = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0.5 & 1 & -1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0.5 & 1 \\ 0 & 0 & 0 & 3 & 1 & 1.5 & 0.75 & 0.5 \end{pmatrix} \tag{45}$$

The specificity of this matrix shows that the unknown input $W_{y,1}$ can be estimate independently from the other ones and consequently the first row of Γ can be studied separately. So the RIP of Γ for $s = 1$ (i.e. W is one sparse) is reduced to the sparsity of Γ_r

$$\Gamma_r = \begin{pmatrix} 0 & -1 & 0 & 1 & 0.5 & 1 & -1 \\ 1 & 0 & 0 & 1 & 1 & 0.5 & 1 \\ 0 & 0 & 3 & 1 & 1.5 & 0.75 & 0.5 \end{pmatrix} \tag{46}$$

and Γ_r verifies the RIP for $s = 1$, so from (40), it is possible to estimate the vector $(W_{y,2}; W_{y,3}; W_{y,4}; W_{x,1}; W_{x,2}; W_{x,3}; W_{x,4})$ if it is one sparse. Moreover, $W_{y,1}$ is given by:

$$W_{y,1} = \dot{y}_1 + y_1 - B_{1,1}u_1$$

7 conclusion

In this paper, we have presented an algorithm that allows to determine whether a rectangular linear dynamical system with more unknown inputs than outputs is potentially left-invertible under the assumption that the unknown inputs are s-sparse. To achieve this, it is shown that it is necessary to determine if the matrix Γ, given by the algorithm, satisfies the well-known RIP condition. A direct extension of this paper is to consider the nonlinear case.

References

[1] A. Balavoine, J. Romberg, and C. J. Rozell. Convergence and rateanalysis of neural networks for sparse approximation. IEEE Transactions on Neural Networks and Learning Systems,, 23(9):1377–1389, 2012.

[2] J-P Barbot, D. Boutat, and T. Floquet. An observation algorithm for nonlinear systems with unknown inputs. Automatica, 45(8):1970–1974, 2009.

[3] V. Bijalwan, P. Kumari, J. Pascual, and V. B. Semwal. Machine learning approach for text and document mining. arXiv preprint arXiv:1406.1580, 2014.

[4] M. Bonilla Estrada, M. Figueroa Garcia, M. Malabre, and J.C. MartÃnez GarcÃa. Left invertibility and duality for linear systems. Linear Algebra and its Applications, 425(2):345–373, 2007. Special Issue in honor of Paul Fuhrmann.

[5] G. Bornard and H. Hammouri. A high gain observer for a class of uniformly observable systems. 1991.

[6] E. J. Candes, J Romberg, and T Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on information theory, 52(2):489–509, 2006.

[7] E. J. Candes and T. Tao. Decoding by linear programming. IEEE Transactions on Information Theory, 51(12):4203–4215, 2005.

[8] E. J. Candes and M. B. Walkin. An introduction to compressive sampling. IEEE Signal Processing Magazine, 21(2):21–30, 2008.

[9] S Derbel, F. Nicolau, N Feki, M Abbes JP Barbot and, and M. Haddar. Detection of gear faults via a dynamical sparse recovery method. In internal report, https://hal.archives-ouvertes.fr/hal-02389942/document, 2018.

[10] D. L. Donoho. Compressed sensing. IEEE Trans. Inf. Theory, 52(4):1289–1306, 2006.

[11] T. Floquet and J.P. Barbot. State and unknown input estimation for linear discrete-time systems. Automatica, 42(11):1883–1889, 2006.

[12] L. Fridman, Y. Shtessel, C. Edwards, and X.G. Yan. Higher-order sliding-mode observer for state estimation and input reconstruction in nonlinear systems. International Journal of Robust and Nonlinear Control, 18(10):399–412, 2008.

[13] J.-P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems with applications to bioreactors. IEEE Transactions on Automatic Control, 37(6):875–880, 1992.

[14] L Praly H Khalil. Highâgain observers in nonlinear feedback control. International Journal of Robust and Nonlinear Control, 24(6):993–1015, 2014.

[15] R. Kalman and R. Bucy. New results in linear filtering and prediction theory. Journal of Basic Engineering, 83:95–108, 1961.

[16] A. Krener and W. Respondek. Nonlinear observer with linearizable error dynamics. SIAM J. Control and Optimization, pages 197–216, 1985.

[17] A. Levant. Robust exact differentiation via sliding mode technique. Automatica, 34(3):379–384, 1998.

[18] A. Levant. Higher-order sliding modes, differentiation and output-feedback control. International Journal on automatic, 76(9/10):924–941, 2003.

[19] A. Levant. Finite differences in homogeneous discontinuous control. IEEE TAC, 52:1208–1217, 2007.

[20] D. Luenberger. Observer for multivariable systems. IEEE Transactions on Automatic Control, 11(2):190–197, 1966.

[21] S Nateghi, Y Shtessel, JP Barbot, and C Edwards. Cyber attack reconstruction of nonlinear systems via higher-order sliding-mode observer and sparse recovery algorithm. In 2018 IEEE Conference on Decision and Control (CDC), December 1018.

[22] W. Perruquetti, T. Floquet, and E. Moulay. Finite time observers: application to secure communication. IEEE Transactions on Automatic Control, 53(1):356–360, 2008.

[23] Y Shtessel, C. Edwards, L. Fridman, and A. Levant. Sliding Mode Control and Observation. BirkhÃuser, New York, NY, 2014.

[24] V.I. Utkin. Sliding Modes in Control and Optimization. Springer-Verlag, 1992.

[25] M. Wakin, B. Sanandaji, and T. Vincent. On the observability of linear systems from ramdon, compressive mesurements. IEEE CDC, pages 4447–4454, 2010.

[26] Y. Xiong and M. Saif. Sliding mode observer for nonlinear uncertain systems. IEEE Transactions on Automatic Control, 46(12):2012–2017, 2001.

[27] L Yu, JP Barbot, G Zheng, and H. Sun. Compressive sensing with chaotic sequence. IEEE Signal Processing Letters, 17(8):731–734, 2010.

[28] L. Yu, H. Sun, J.P. Barbot, and G. Zheng. Bayesian compressive sensing for cluster structured sparse signals. Signal Processing, 92(1):259–269, 2012.

[29] L. Yu, G Zheng, and JP Barbot. Dynamical sparse recovery with finite-time convergence. IEEE Transactions on Signal Processing,, 65(23):6146–6157, 2017.