Data Informativity for the Open-Loop Identification of MIMO Systems in the Prediction Error Framework

Kévin Colin, Xavier Bombois, Laurent Bako, Federico Morelli

Data Informativity for the Open-Loop Identification of MIMO Systems in the Prediction Error Framework

Kévin Colin [a], Xavier Bombois [b], Laurent Bako [a], Federico Morelli [a]

[a] *Laboratoire Ampère, UMR CNRS 5005, Ecole Centrale de Lyon, Université de Lyon, France*

[b] *Centre National de la Recherche Scientifique (CNRS), France*

Abstract

In Prediction Error identification, to obtain a consistent estimate of the true system, it is crucial that the input excitation yields informative data with respect to the chosen model structure. We consider in this paper the data informativity property for the identification of a Multiple-Input Multiple-Output system in open loop and we derive conditions to check whether a given input vector will yield informative data with respect to the chosen model structure. We do that for the classical model structures used in prediction-error identification and for the classical types of input vectors, i.e., input vectors whose elements are either multisines or filtered white noises.

Key words: System Identification; Data informativity; MIMO Systems; Prediction Error Methods; Consistency

1 Introduction

When designing the input excitation for an identification experiment, it is crucial that this input excitation yields informative data. Indeed, data informativity ensures the consistency of the prediction error estimate (if the considered model structure is identifiable at the true parameter vector) [12,14]. In this paper, we consider this important problem for the open-loop identification of Multi-Input Multi-Output (MIMO) systems.

Data informativity is guaranteed when the input excitation is sufficiently rich to guarantee that the prediction error is different for different models in the considered model structure. Data informativity has been studied extensively in the Single-Input Single-Output (SISO) case. In this particular case, the data are informative if the input signal is sufficiently rich of an order that depends on the type and the complexity of the considered model structure. That an input signal u is sufficiently rich of a certain order η is equivalent to the fact that its power spectrum $\Phi_u(\omega)$ is nonzero in at least η frequencies in the interval $]-\pi,\ \pi]$ (see e.g. [12]). In the recent years, due to the renewed interest in optimal experiment design (see e.g. [10,4]) where the covariance matrix of the identified model plays an important role, there has been a number of efforts to relate the positive definiteness property of the covariance matrix and the data informativity [9,2]. In [9] and also in [8], necessary and sufficient conditions on the signal richness of the excitation are also derived for the data informativity in both the open-loop and the closed-loop case. In particular, these papers derive the minimal order of signal richness that the excitation signal must have to ensure data informativity and this is done for all classical model structures (BJ, ARX, ...). In the closed-loop case, this minimal order is related to the complexity of the controller present in the loop during the identification. It is also shown in these papers (see also [14]) that, if the controller is sufficiently complex, data informativity can also be obtained by just using the noise excitation.

While the analysis of data informativity in the SISO case seems a mature research area, this cannot be said for the MIMO case. While there has been attention towards determining the order of the MIMO controller to ensure that an identification in closed loop without external excitation will be informative [3,13], there is to this date almost no results that allow one to verify whether a given input vector will yield informative data in the open-loop case or in the closed-loop case (when the controller is not sufficiently complex). Up to our knowledge, the only available condition for data informativity is to check that the power spectrum of the input vector is strictly positive definite at (almost) all frequencies (see [3] or [7, Propo-

Email addresses: kevin.colin@ec-lyon.fr (Kévin Colin), xavier.bombois@ec-lyon.fr (Xavier Bombois), laurent.bako@ec-lyon.fr (Laurent Bako), federico.morelli@ec-lyon.fr (Federico Morelli).

sition 3]). However, as we will see in this paper, this sufficient condition is too restrictive (e.g. a multisine input vector will never respect this condition). Due to this observation, we derive in this paper a less conservative condition that allows one to check whether a given input vector will yield informative data to identify a model in open loop in a given model structure. We will do that for the classical model structures used in prediction-error identification (ARX, FIR, BJ, OE model structures) and for both multisine input vector and for an input vector generated as $u(t) = N(z)v(t)$ with v a vector of independent white noises of arbitrary dimension and N a matrix of transfer functions. These conditions can, e.g., be of importance for MIMO optimal experiment design (see, e.g., [1,11]) since it will allow to choose the input vector parametrization in such a way that data informativity is guaranteed.

To derive these conditions, we first show that the MIMO case can be treated by analyzing the data informativity channel by channel. An analysis of the data informativity channel by channel boils down to the analysis of data informativity for Multiple-Input Single-Output (MISO) systems. We will show that the data informativity with respect to a MISO model structure can be guaranteed by the linear independence of the elements of a vector of signals that we will call *regressor vector*. This regressor vector $\phi_u(t)$ contains delayed versions of the elements of the input vector. There is of course an analogy here with the SISO case since a scalar input signal is sufficiently rich of a given order η if the elements of such a regressor vector of dimension η are linearly independent [8].

We then show that, in the MISO case, the linear independence of the elements of $\phi_u(t)$ is equivalent to the fact that a given matrix is full row rank. This matrix will be a function of the input parametrization, i.e., the amplitudes, phase shifts and frequencies of the different sinusoids in the multisine case and the coefficients in the matrix $N(z)$ in the filtered white noise case. We then analyze which conditions are necessary to make this matrix full rank. We therefore derive, for the multisine case, conditions on the number of sinusoids that are present in the multisine input vector and, for the filtered white noise case, conditions on the complexity of the matrix N and on the number of white noises in v.

This paper builds upon our previous contributions [6,5] where we restricted attention to the MISO case and to FIR model structures.

2 Notations

For a complex-valued matrix A, A^T denotes its transpose and A^* its conjugate transpose. A positive semi-definite (resp. definite) matrix A is denoted $A \succeq 0$ (resp. $A \succ 0$). We will denote A_{ik} the (i,k)-entry of the matrix A, $A_{i:}$ the i-th row of A and $A_{:k}$ the k-th column of A. The identity matrix of size $n \times n$ is denoted $\mathbf{I}_n$ and $\mathbf{0}_{n\times p}$ is the $n \times p$ matrix full of zeros. The rank of a matrix A will be denoted rank(A). For a vector $x \in \mathbb{R}^n$, the notation $||x||$ refers to the Euclidean norm, i.e., $||x|| = \sqrt{x_1^2 + \cdots + x_n^2}$.

For a discrete-time transfer function $G(z)$, z represents the forward-shift operator and its frequency response is given by $G(z = e^{j\omega})$ where ω is the normalized frequency in $[0\ \pi]$. For a polynomial $P(X)$, the coefficient linked to X^k is denoted $p^{(k)}$. Thus, $P(X) = p^{(0)} + p^{(1)}X + \cdots + p^{(k)}X^k + \cdots + p^{(n)}X^n$. The degree n of this polynomial $P(X)$ is denoted deg$(P(X))$. When $X = z^{-1}$, we say that ρ is the delay of $P(z^{-1})$ when the first non-zero coefficient is linked to $z^{-\rho}$, i.e., $P(z^{-1}) = p^{(\rho)}z^{-\rho} + p^{(\rho+1)}z^{-(\rho+1)} + \cdots + p^{(n)}X^n$ with $p^{(\rho)} \neq 0$.

Finally, for quasi-stationary signals $x(t)$, the operator $\bar{E}[x(t)]$ is defined as $\bar{E}[x(t)] = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} E[x(t)]$ (E being the expectation operator) [12].

3 Prediction Error Framework

Consider a MIMO system $\mathscr{S}$ in open-loop with an input vector $u \in \mathbb{R}^{n_u}$ and an output vector $y \in \mathbb{R}^{n_y}$, described by

$$\mathscr{S}:\ y(t) = G_0(z)u(t) + H_0(z)e(t) \tag{1}$$

where $G_0(z)$ is a stable matrix of (rational) transfer functions of dimension $n_y \times n_u$, $H_0(z)$ a stable, inversely-stable and monic matrix[1] of (rational) transfer functions of dimension $n_y \times n_y$ and e is a vector made up of n_y white noise signals such that $\bar{E}\left[e(t)e^T(t)\right] = \Sigma_0 \succ 0$. Since the system $\mathscr{S}$ is operated in open-loop, the excitation signal that will be used for identification purpose will be the input vector $u(t)$ which is assumed to be independent of $e(t)$. Moreover, we will assume two different types of quasi-stationary excitation u.

In the first type, each entry u_k of u ($k = 1, \cdots, n_u$) is a multisine made up of sinusoids at p different frequencies ω_l ($l = 1, \cdots, p$), i.e.,

$$u_k(t) = \sum_{l=1}^{p} \Lambda_{kl} \cos(\omega_l t + \Psi_{kl}) \quad k = 1, \cdots, n_u \tag{2}$$

where Λ_{kl} and Ψ_{kl}, are respectively the amplitude and phase shift of the sinusoid at the frequency ω_l. Note that Λ_{kl} can be zero for some value(s) of k ($k = 1, \cdots, n_u$), but, for each $l = 1, \cdots, p$, there exists (at least) a value of k for which $\Lambda_{kl} \neq 0$. For further reference, we will denote by p_k ($k = 1, \cdots, n_u$) the number of sinusoids for which the amplitude Λ_{kl} in the expression (2) for u_k is nonzero ($p_k \leq p\ \forall k$).

In the second type of input u, u is generated as $u(t) = N(z)v(t)$ via a stable transfer function matrix $N(z) = (N_{kq}(z))_{(k,q)\in [\![1,n_u]\!] \times [\![1,r]\!]}$ and a vector

[1] i.e. H_0 and H_0^{-1} are stable and $H_0(z = \infty) = \mathbf{I}_{n_y}$

$v = (v_1, \cdots, v_r)^T$ containing r independent white noises signals v_q $(q = 1, \cdots, r)$. In other words, each entry u_k of u is given by:

$$u_k(t) = \sum_{q=1}^{r} N_{kq}(z) v_q(t) \quad k = 1, \cdots, n_u. \tag{3}$$

Note that r is not necessarily equal to n_u and that $N_{kq}(z)$ can be identically zero for some value(s) of k $(k = 1, \cdots, n_u)$, but, for each $q = 1, \cdots, r$, there exists (at least) a value of k for which $N_{kq}(z) \neq 0$.

To identify $\mathscr{S}$, we consider a full-order model structure $\mathscr{M} = \{(G(z,\theta), H(z,\theta)) \mid \theta \in \mathscr{D}_\theta\}$ where $\theta \in \mathscr{D}_\theta \subset \mathbb{R}^n$ is a parameter vector. This model structure is full-order if $\exists \theta_0 \in \mathscr{D}_\theta \subset \mathbb{R}^n$, called the true parameter vector, such that $(G(z,\theta_0), H(z,\theta_0)) = (G_0(z), H_0(z))$. We will further suppose that this model structure is globally identifiable at the true parameter vector θ_0, i.e., $(G(z,\theta), H(z,\theta)) = (G(z,\theta_0), H(z,\theta_0)) \implies \theta = \theta_0$. The set $\mathscr{D}_\theta$ restricts the parameter vector θ to those values for which $G(z,\theta)$ is stable and $H(z,\theta)$ is stable and inversely stable.

Consider a set of data $Z^N = \{s(t) = (u^T(t), y^T(t))^T \mid t = 1, \cdots, N\}$ collected on the true system $\mathscr{S}$. We can define, based on the model structure $\mathscr{M}$, the one-step ahead predictor $\hat{y}(t,\theta)$ for the output $y(t)$ at time t:

$$\hat{y}(t,\theta) = W_u(z,\theta) u(t) + W_y(z,\theta) y(t) = W(z,\theta) s(t) \tag{4}$$

where

$$W_u(z,\theta) = H^{-1}(z,\theta) G(z,\theta) \tag{5}$$
$$W_y(z,\theta) = \mathbf{I}_{n_y} - H^{-1}(z,\theta) \tag{6}$$
$$W(z,\theta) = (W_u(z,\theta), W_y(z,\theta)). \tag{7}$$

As we will see in the sequel, it is important that the data $s(t) = (u^T(t), y^T(t))^T$ collected on the true system $\mathscr{S}$ be informative with respect to $\mathscr{M}$.

Definition 1 (Data Informativity [3]) *Consider the framework defined above with data $s(t) = (u^T(t), y^T(t))^T$ collected by applying a quasi-stationary input vector u to the true system $\mathscr{S}$. Consider a model structure $\mathscr{M}$ yielding the predictor $\hat{y}(t,\theta) = W(z,\theta)s(t)$. Let us also define the set $\Delta_{\mathbf{W}} = \{\Delta W(z) \mid \Delta W(z) = W(z,\theta') - W(z,\theta'')$ with θ' and θ'' in $\mathscr{D}_\theta\}$. The data $s(t)$ is said to be informative w.r.t. the model structure $\mathscr{M}$ when, for all $\Delta W(z) \in \Delta_{\mathbf{W}}$, we have*

$$\bar{E}\left[||\Delta W(z) s(t)||^2 \right] = 0 \Longrightarrow \Delta W(z) \equiv 0 \tag{8}$$

where $\Delta W(z) \equiv 0$ means that $\Delta W(e^{j\omega}) = 0$ for almost all ω. □

Data informativity is an important property since it implies that the prediction error criterion [12,14] defined below yields a consistent estimate $\hat{\theta}_N$ for θ_0:

$$\hat{\theta}_N = \arg \min_{\theta \in \mathscr{D}_\theta} V_N(\theta) \tag{9}$$

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \epsilon^T(t,\theta) \Sigma_0^{-1} \epsilon(t,\theta) \tag{10}$$

where $\epsilon(t,\theta) = y(t) - \hat{y}(t,\theta)$ and where Σ_0 is supposed known for simplicity (it can however be estimated together with $\hat{\theta}_N$ (see, e.g., [12, Chapter 15]). The estimate $\hat{\theta}_N$ is consistent if it converges to θ_0 with probability 1 when $N \to +\infty$ or, equivalently, if and only if θ_0 is the unique minimum of the asymptotic criterion $\bar{E}\left[\epsilon^T(t,\theta)\Sigma_0^{-1}\epsilon(t,\theta)\right]$.

In the sequel, we will therefore derive conditions to ensure the data informativity for input vectors u of the types (2)-(3). It is important to note that the parametrizations (2)-(3) for the input vector u cover many cases where $\det(\Phi_u(\omega)) = 0 \ \forall \omega$ (e.g., when $r < n_u$ in (3) or, always, for the multisine case). Consequently, many of these input vectors u do not satisfy the sufficient condition for informativity discussed in the introduction (see, e.g., [3] or [7, Proposition 3]).

We will derive conditions for data informativity for the following classical model structures $\mathscr{M}$: FIR, ARX, OE and BJ model structures. Note that we will not consider the ARMAX model structure to simplify the presentation and since an ARMAX system can always be represented within a BJ model structure. The parametrization used in these classical model structures are reminded in the sequel. For this purpose, let us decompose θ as $\theta = (\tilde{\theta}^T, \ \eta^T)^T$ where $\tilde{\theta}$ contains the parameters that we only find in $G(z,\theta)$ and η contains the rest of the parameters (i.e. the parameters in $H(z,\theta)$ and the parameters that are common to $G(z,\theta)$ and $H(z,\theta)$). Based on this decomposition, the ARX model structure can be defined as follows

$$\begin{cases} G(z,\theta) = A(z,\eta)^{-1} B(z,\tilde{\theta}) \\ H(z,\theta) = A(z,\eta)^{-1} \end{cases} \tag{11}$$

where $A(z,\eta)$ and $B(z,\tilde{\theta})$ are respectively $n_y \times n_y$ and $n_y \times n_u$ polynomial matrices and where $A(z,\eta)$ is restricted to be monic. The parameters in the vector θ are here given by the coefficients of the polynomials in these matrices and we will suppose that all the polynomials in these matrices are independently parametrized (i.e. do not share any common parameters). The FIR model structure is a special case of the ARX model structure where $A(z,\eta) = \mathbf{I}_{n_y}$.

In the BJ model structure, there are no common parameters in $G(z,\theta)$ and $H(z,\theta)$:

$$\begin{cases} G(z,\theta) \;=\; G(z,\tilde{\theta}) \\ H(z,\theta) \;=\; H(z,\eta) \end{cases} \tag{12}$$

The parameter vector $\tilde{\theta}$ contains the parameters of the numerators and (monic) denominators of each transfer function in the matrix of transfer functions $G(z,\tilde{\theta})$. The same holds for η and $H(z,\eta)$, but in this case, the parametrization must also be such that $H(z=\infty,\eta)=\mathbf{I}_{n_y}$. We here also suppose that the individual entries in these matrices do not share any common parameters. The OE model structures is a special case of the BJ model structure where $H\big(z,\eta\big)=\mathbf{I}_{n_y}$.

When the true system is ARX/FIR, a full-order ARX/FIR model structure will always be globally identifiable at θ_0. In the BJ/OE case, the model structure will be globally identifiable at θ_0 if θ_0 does not lead to any pole-zero cancellation in the entries of $G(z,\theta)$ and $H(z,\theta)$.

4 Data informativity for MIMO systems

As mentioned above, our objective in this paper is to derive conditions to ensure the data informativity for input signals u of the types (2)-(3) and for the classical types of model structures $\mathcal{M}$. For this purpose, in the following two theorems, we will first see that the analysis of the data informativity in the MIMO case can be simplified. In particular, we will see that, by virtue of the particular parametrization of the classical model structures, the condition (8) in the definition of data informativity can be simplified (see Theorem 2) and, subsequently, that the verification of this simplified condition can be performed channel by channel (see Theorem 3).

Theorem 2 *Consider Definition 1 and the model structures $\mathcal{M}$ defined at the end of Section 3. Let us also define the set $\Delta_{\mathbf{X}}=\{\Delta X(z)\mid\Delta X(z)=X(z,\theta')-X(z,\theta'')$ with θ' and θ'' in $\mathcal{D}_\theta\}$ where $X(z,\theta)=B(z,\tilde{\theta})$ for ARX/FIR model structures and where $X(z,\theta)=G(z,\tilde{\theta})$ for BJ/OE model structures. Then, the data $s(t)=(u^T(t),y^T(t))^T$ are informative w.r.t. the model structure $\mathcal{M}$ if and only if, for all $\Delta X(z)\in\Delta_{\mathbf{X}}$, we have:*

$$\bar{E}\left[||\Delta X(z)u(t)||^2\right]=0 \Longrightarrow \Delta X(z)\equiv 0 \tag{13}$$

□

PROOF. See Appendix A.

Since $\bar{E}\left[||\Delta X(z)u(t)||^2\right]$ is equal to the trace of $(1/2\pi)\int_{-\pi}^{\pi}\Delta X(e^{j\omega})\Phi_u(\omega)\Delta X^*(e^{j\omega})\;d\omega$, independently of the model structure $\mathcal{M}$, (2) will hold for all $\Delta X\in\Delta_{\mathbf{X}}$ it is clear that (13) will hold for any model structure if $\Phi_u(\omega)\succ 0$ at almost all frequencies. This sufficient condition is the one discussed in the introduction.

Theorem 3 *Consider Definition 1 and Theorem 2 and let us define based on $\Delta_{\mathbf{X}}$ the sets $\Delta_{\mathbf{X,i}}=\{\Delta X_{i:}(z)\mid\Delta X_{i:}(z)$ is the ith row of $\Delta X(z)\in\Delta_{\mathbf{X}}\}$ $(i=1,\cdots,n_y)$. Then, the data $s(t)$ are informative w.r.t. the model structure $\mathcal{M}$ if and only if, for all $i=1,\cdots,n_y$, the following property holds for all $\Delta X_{i:}\in\Delta_{\mathbf{X,i}}$:*

$$\bar{E}\left[||\Delta X_{i:}(z)u(t)||^2\right]=0 \Longrightarrow \Delta X_{i:}(z)\equiv 0 \tag{14}$$

PROOF. See Appendix B.

Theorems 2 and 3 are important since they imply that, to check the informativity of the data $s(t)=(u^T(t),y^T(t))^T$ with respect to a MIMO model structure $\mathcal{M}$, we can equivalently check, for all $i=1,\cdots,n_y$, the informativity of the data $s_i(t)=(u^T(t),y_i(t))^T$ (Theorem 3) and this verification for s_i only involves the ith row $B_{i:}(z,\tilde{\theta})$ of the parametrization $B(z,\tilde{\theta})$ (for ARX/FIR model structures) or the ith row $G_{i:}(z,\tilde{\theta})$ of the parametrization $G(z,\tilde{\theta})$ (for BJ/OE structures). It is also clear that the data $s_i(t)=(u^T(t),y_i(t))^T$ correspond to the data of a MISO system with n_u inputs. Consequently, we can, from now on, restrict attention to MISO systems and MISO model structures.

5 Data informativity for MISO systems

5.1 Notations for a MISO model structure

Let us introduce some further notations for the MISO model structures and let us restrict attention to those parts that are relevant for data informativity (see Theorem 2). In a MISO ARX/FIR model structure, $B(z,\tilde{\theta})$ is a vector made up of n_u independently parametrized polynomials $B_k(z,\tilde{\theta}_k)$ $(k=1,\cdots,n_u)$. In the BJ/OE case, $G(z,\tilde{\theta})$ is a vector made up of n_u independently parametrized transfer functions $G_k(z,\tilde{\theta}_k)$. In both cases, the parameter vector $\tilde{\theta}\in\mathbb{R}^{\tilde{n}}$ is the concatenation of $\tilde{\theta}_k$ $(k=1,\cdots,n_u)$, i.e., $\tilde{\theta}=(\tilde{\theta}_1^T,\tilde{\theta}_2^T,\cdots,\tilde{\theta}_{n_u}^T)^T$ and where $\tilde{n}$ is the dimension of $\tilde{\theta}$. Let us also introduce the following notation for each entry of these vectors:

$$B_k(z,\tilde{\theta}_k)=z^{-\rho_k}\mathring{B}_k(z,\tilde{\theta}_k) \tag{15}$$

$$G_k(z,\tilde{\theta}_k)=z^{-\rho_k}\frac{\mathring{B}_k(z,\tilde{\theta}_k)}{\mathring{F}_k(z,\tilde{\theta}_k)} \tag{16}$$

where the delays ρ_k can be different for each k and where both $\mathring{B}_k$ and $\mathring{F}_k$ are polynomials ($\mathring{F}_k$ is moreover a monic polynomial, i.e., $\mathring{F}_k(z=\infty)=1$). In the sequel, we will denote the degree of these polynomials by $\deg(\mathring{B}_k)$ and $\deg(\mathring{F}_k)$, respectively (see Section 2). The coefficients of the polynomials $\mathring{B}_k$ and $\mathring{F}_k$ are the parameters in $\tilde{\theta}_k$. Consequently, we have that:

$$\mathring{B}_k(z,\tilde{\theta}_k) = \tilde{\theta}_{k,1} + \sum_{m=1}^{\deg(\mathring{B}_k)} \tilde{\theta}_{k,(m+1)} z^{-m} \tag{17}$$

$$\mathring{F}_k(z,\tilde{\theta}_k) = 1 + \sum_{m=1}^{\deg(\mathring{F}_k)} \tilde{\theta}_{k,(m+\deg(\mathring{B}_k)+1)} z^{-m} \tag{18}$$

where $\tilde{\theta}_{k,m}$ denotes the mth entry of $\tilde{\theta}_k$. The number of parameters to identify in $\mathring{B}_k$ and $\mathring{F}_k$ is thus equal to $\deg(\mathring{B}_k)+1$ and $\deg(\mathring{F}_k)$, respectively.

In Section 3, we have restricted the value of $\theta \in \mathbb{R}^n$ to the ones in $\mathscr{D}_\theta$. Since only $\tilde{\theta} \in \mathbb{R}^{\tilde{n}}$ is relevant for data informativity, we introduce the set $\mathscr{D}_{\tilde{\theta}}$ as $\mathscr{D}_{\tilde{\theta}} = \{\tilde{\theta} \mid B(z,\tilde{\theta}) \text{ is stable}\}$ (ARX/FIR model structures) and as $\mathscr{D}_{\tilde{\theta}} = \{\tilde{\theta} \mid G(z,\tilde{\theta}) \text{ is stable}\}$ (BJ/OE model structures). Note that $\mathscr{D}_{\tilde{\theta}}$ covers the whole parameter space $\mathbb{R}^{\tilde{n}}$ in the ARX/FIR case since $B(z,\tilde{\theta})$ is always stable.

5.2 Persistency of a vector of signals

In the sequel, we will show that the data informativity can be linked to the persistency of a vector of signals. The persistency of a (complex- or real-valued) vector of signals is defined as follows:

Definition 4 (Persistency) *A quasi-stationary real or complex-valued vector of signals ϕ is persistently exciting (PE) if and only if $\bar{E}[\phi(t)\phi^*(t)] \succ 0$.* □

We have also the following useful result regarding real-valued vector of signals.

Lemma 5 *Consider a vector ϕ containing p real-valued quasi-stationary signals ϕ_i. Then, the three following propositions are equivalent:*

- (a) *ϕ is PE.*
- (b) $\bar{E}[\phi(t)\phi^T(t)] \succ 0$.
- (c) $\forall \alpha \in \mathbb{R}^p$, $\bar{E}[\alpha^T\phi(t)]^2 = 0$ *if and only if* $\alpha = \mathbf{0}_{p\times 1}$. □

PROOF. $(a) \Leftrightarrow (b)$: this follows from a straightforward application of Definition 4 for a real-valued vector of signals.

$(b) \Leftrightarrow (c)$: $\bar{E}[\phi(t)\phi^T(t)] \succ 0$ is equivalent to the fact that, for any $\alpha \in \mathbb{R}^p$, $\alpha^T\bar{E}[\phi(t)\phi^T(t)]\alpha = 0 \Leftrightarrow \alpha = \mathbf{0}_{p\times 1}$ which is in turn equivalent to condition (c). ■

Remark 6 *The condition (c) of Lemma 5 means that the elements ϕ_i of ϕ are linearly independent signals.* □

5.3 Input regressor persistency

Using the notion of persistency defined in the previous subsection, we show in the next theorem that the data informativity in the MISO case can be guaranteed by the persistency of a particular vector of signals depending on the input $u(t)$. This vector will be denoted $\phi_u(t)$ and will be called *regressor*.

Theorem 7 (MISO informativity) *Consider the data $s(t) = (u^T(t), y(t))^T$ obtained by applying an input vector u on a MISO system. Consider a full-order model structure $\mathscr{M}$ for this MISO system and the notations introduced in Section 5.1. Then, the data $s(t)$ are informative with respect to $\mathscr{M}$ if the regressor $\phi_u(t)$ defined below is PE:*

$$\phi_u(t) = \begin{pmatrix} \phi_{u_1}(t) \\ \vdots \\ \phi_{u_{n_u}}(t) \end{pmatrix} \quad \textit{with} \quad \phi_{u_k}(t) = \begin{pmatrix} u_k(t-\rho_k) \\ \vdots \\ u_k(t-\rho_k-\eta_k) \end{pmatrix} \tag{19}$$

The dimension μ of the regressor ϕ_u in (19) is given by $\mu = n_u + \sum \eta_k$ (the dimension of ϕ_{u_k} being equal to $\eta_k + 1$ $(k = 1, \cdots, n_u)$) and the scalars[2] η_k $(k = 1, \cdots, n_u)$ in (19) are respectively given by:

- $\eta_k = deg(\mathring{B}_k)$ *for ARX and FIR model structures.*
- $\eta_k = deg(\mathring{B}_k) + deg(\mathring{F}_k) + 2\sum_{m=1, m\neq k}^{n_u} deg(\mathring{F}_m)$ *for BJ and OE model structures.*

For ARX and FIR model structures, the persistency of ϕ_u is not only a sufficient condition for data informativity but also a necessary condition. □

PROOF. See Appendix C ■

If we apply this theorem to the SISO case ($n_u = 1$), we observe that the condition for data informativity is that the regressor $\phi_u = \phi_{u_1}$ of dimension $\eta_1 + 1$ is PE. This condition is equivalent to the fact that the scalar signal $u = u_1$ is sufficiently rich of order $\eta_1 + 1$ (see [8,9]) or, equivalently, that the power spectrum $\Phi_u(\omega)$ of $u = u_1$ is nonzero in at least $\eta_1 + 1$ frequencies in $]-\pi, \pi]$. By computing η_1 when $n_u = 1$, we then retrieve the classical results for data informativity with respect to the ARX/FIR/BJ/OE models structures in the SISO case (see, e.g., [8]).

While, in the SISO case, the persistency of ϕ_{u_1} is a necessary and sufficient condition for data informativity for all model structures, the persistency of the regressor ϕ_u defined in (19) is only a necessary and sufficient condition for data informativity for the ARX/FIR model structures while it is just a sufficient condition in the BJ/OE case. In Section 9, we will show that this sufficient condition is nevertheless not overly conservative.

In the next sections, we will derive a necessary and sufficient condition on the input vector u to guarantee that

[2] The notations η_k are not related to the vector η introduced at the end of Section 3.

ϕ_u is PE. In the ARX/FIR case, this condition will then be a necessary and sufficient condition for data informativity while it will be only sufficient one in the BJ/OE case. Before deriving this necessary and sufficient condition for ϕ_u to be PE, let us formulate the following lemma that will give us a first idea on how u has to be chosen for this purpose.

Lemma 8 *Consider a quasi-stationary input vector u of dimension n_u and the corresponding regressor ϕ_u defined in* (19). *If ϕ_u is PE, then we have necessarily the following two properties:*

- (i) *each ϕ_{u_k} $(k = 1, \cdots, n_u)$ is PE*
- (ii) *for any $m \in \mathbb{N}$, the vector $\phi_{\mathcal{U}_m}$ made up of the elements of the set $\mathcal{U}_m = \{u_k(t) \mid u_k(t-m)$ is present in $\phi_u(t)\}$ is PE. In other words, the inputs u_k that appear with the same delay in ϕ_u must be linearly independent.*

Moreover, for any $k = 1, \cdots, n_u$, the vector ϕ_{u_k} of dimension $\eta_k + 1$ is PE if and only if the power spectrum $\Phi_{u_k}(\omega)$ of u_k $(k = 1, \cdots, n_u)$ is nonzero in at least $\eta_k + 1$ frequencies in $]-\pi, \pi]$. □

PROOF. The first point directly follows from condition (c) of Lemma 5 and the expression (19) for ϕ_u. Indeed, condition (c) can only hold for ϕ_u if it also holds for each ϕ_{u_k} $(k = 1, \cdots, n_u)$. The second point follows from a similar reasoning and from the fact that a quasi-stationary vector $\phi(t)$ is PE if and only if $\phi(t-m)$ is PE $\forall m \in \mathbb{N}$.

Finally, due to Definition 4, the persistency of ϕ_{u_k} is equivalent to the fact that u_k is sufficiently rich of order $\eta_k + 1$ (see also [8,9]) and this property of u_k is equivalent to the property on its spectrum given in the statement of the lemma [12]. ■

Let us first discuss the necessary condition (i). This condition will be always respected in the filtered white noise case since $\Phi_{u_k}(\omega)$ will be non-zero at (almost) all frequencies ω (u_k is indeed generated by at least one filtered white noise). In the multisine case, the condition (i) entails that, for each $k = 1, \cdots, n_u$, the number p_k of sinusoids that are effectively present in u_k (i.e. for which $\Lambda_{kl} \neq 0$) satisfies the following constraint:

$$p_k \geq \frac{\eta_k + 1}{2} \tag{20}$$

Note that this condition is independent of the choice of the amplitudes and phase shifts in (2).

Let us now consider the property (ii). In the majority of the model structures, there will be a value of the delay m for which $\mathcal{U}_m$ will contain all entries of u (e.g. this happens when the delays ρ_k are the same for all k). When this is the case, property (ii) requires that the signals u_k $(k = 1, \cdots, n_u)$ are linearly independent (the vector of signals u is PE). In the multisine case, we will see in the sequel that u is PE if and only if a certain matrix made up of the amplitude Λ_{kl} and the phase shifts Ψ_{kl} is full (row) rank. In the filtered white noise case, this entails that the rows of $N(z)$ are linearly independent. Even though choosing an input vector that is PE is a good practice, this property is only strictly necessary for ϕ_u to be PE for the model structures having the property described in the beginning of this paragraph, as shown in the following example that also illustrates the construction of ϕ_u.

Example 9 *Consider a MISO OE model structure with $n_u = 3$ and for which the predictor $\hat{y}(t, \theta) = G(z, \theta)u(t)$ is given by:*

$$\hat{y}(t, \theta) = \frac{\theta_1}{1 + \theta_2 z^{-1}} u_1(t) + \theta_3 z^{-9} u_2(t) + \frac{\theta_4 z^{-4}}{1 + \theta_5 z^{-1}} u_3(t)$$

with $\theta^T = \tilde{\theta}^T = \begin{pmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_5 \end{pmatrix}$. We observe that, in this model structure, the delays are given by $(\rho_1, \rho_2, \rho_3) = (0, 9, 4)$ and that the transfer functions G_k $(k = 1, 2, 3)$ in G have the form (16) *with*

$$\begin{array}{lll} \mathring{B}_1 = \theta_1 & \mathring{B}_2 = \theta_3 & \mathring{B}_3 = \theta_4 \\ \mathring{F}_1 = 1 + \theta_2 z^{-1} & \mathring{F}_2 = 1 & \mathring{F}_3 = 1 + \theta_5 z^{-1} \end{array}$$

Let us deduce the regression order η_k defined in Theorem 7:

- $\eta_1 = deg(\mathring{B}_1) + deg(\mathring{F}_1) + 2deg(\mathring{F}_2) + 2deg(\mathring{F}_3) = 3$
- $\eta_2 = deg(\mathring{B}_2) + deg(\mathring{F}_2) + 2deg(\mathring{F}_1) + 2deg(\mathring{F}_3) = 4$
- $\eta_3 = deg(\mathring{B}_3) + deg(\mathring{F}_3) + 2deg(\mathring{F}_1) + 2deg(\mathring{F}_2) = 3$

and let us form the input regressor ϕ_u defined in (19), *i.e., $\phi_u(t) = (\phi_{u_1}(t)^T, \phi_{u_2}(t)^T, \phi_{u_3}(t)^T)^T$ with*

$$\begin{aligned} \phi_{u_1}^T(t) &= \begin{pmatrix} u_1(t) & u_1(t-1) & u_1(t-2) & u_1(t-3) \end{pmatrix} \\ \phi_{u_2}^T(t) &= \begin{pmatrix} u_2(t-9) & u_2(t-10) & u_2(t-11) & u_2(t-12) & u_2(t-13) \end{pmatrix} \\ \phi_{u_3}^T(t) &= \begin{pmatrix} u_3(t-4) & u_3(t-5) & u_3(t-6) & u_3(t-7) \end{pmatrix} \end{aligned}$$

We can see that each set $\mathcal{U}_m$ contains at most one signal. This means that the condition (ii) of Lemma 8 simply says that the inputs $u_k(t)$ must not be identically equal to 0. It is therefore possible to get the persistency of ϕ_u even if there is a linear dependence between $u_1(t)$, $u_2(t)$ and $u_3(t)$. In particular, we can choose $u_1 = u_2 = u_3 = v$ where v is a white noise signal. In this case, ϕ_u is indeed PE since it is made up of different elements of a white noise sequence. This choice of u thus guarantees data informativity (Theorem 7). □

It is important to stress that Lemma 8 only gives necessary conditions for ϕ_u to be PE. In the next section, we will therefore continue our analysis by deriving a necessary and sufficient condition for the input regressor to be PE. This result will entail (among other aspects) that, to guarantee this property which implies the data informativity, we also require an additional condition on the total number p of sinusoids in the multisine case and on the complexity of the filter matrix $N(z)$ in the filtered white noise case.

6 Necessary and sufficient condition for the input regressor persistency

6.1 General idea

The necessary and sufficient condition will be derived by rewriting $\phi_u(t)$ (see (19)) into the form $\phi_u(t) = \mathcal{N}\varphi(t)$ where $\mathcal{N}$ is a (possibly complex) matrix whose entries are functions of the input parametrization and $\varphi(t)$ is a vector of (possibly complex-valued) quasi-stationary signal that is always PE (see Definition 4). This decomposition is important since we have then the following result:

Lemma 10 *Consider a vector of quasi-stationary signals $\phi(t)$ given by $\phi(t) = \mathcal{N}\varphi(t)$ where $\mathcal{N}$ is a complex matrix and $\varphi(t)$ is a vector of complex-valued quasi-stationary signals that has the property of being PE. Then, $\phi(t)$ is PE if and only if $\mathcal{N}$ is full row rank.* □

PROOF. Due to Definition 1, the lemma will be proven if we show that $\bar{E}[\phi(t)\phi^*(t)] = \mathcal{N}\bar{E}[\varphi(t)\varphi^*(t)]\mathcal{N}^* \succ 0$. Since $\varphi(t)$ is PE, we have $\bar{E}[\varphi(t)\varphi^*(t)] \succ 0$. Consequently, a full row rank $\mathcal{N}$ is indeed a necessary and sufficient condition for $\bar{E}[\phi(t)\phi^*(t)] \succ 0$. ■

In the next two sections, we will show that we can rewrite ϕ_u in this way for multisine and filtered white noise excitation.

6.2 Input regressor for multisine excitation

In this section, we consider that the input $u(t)$ is a multisine such that each $u_k(t)$ is given by (2). By using Euler's formula, we have that

$$u_k(t-m_k) = \frac{1}{2}\sum_{l=1}^{p}\left(\overline{\Lambda}_{kl}e^{-jm_k\omega_l}e^{j\omega_l t} + \overline{\Lambda}_{kl}^{*}e^{jm_k\omega_l}e^{-\omega_l t}\right)$$

where $\overline{\Lambda}_{kl} = \Lambda_{kl}e^{j\Psi_{kl}}$ is a phasor and $m_k \in [\![\rho_k, \rho_k+\eta_k]\!]$. Consequently,

$$\phi_{u_k}(t) = \mathcal{Q}_k\phi_{sin}(t) \tag{21}$$

where $\mathcal{Q}_k$ and $\phi_{sin}(t)$ are successively defined by

$$\mathcal{Q}_k^T = \begin{pmatrix} \overline{\Lambda}_{k1}e^{-j\rho_k\omega_1} & \dots & \overline{\Lambda}_{k1}e^{-j(\rho_k+\eta_k)\omega_1} \\ \overline{\Lambda}_{k1}^{*}e^{j\rho_k\omega_1} & \dots & \overline{\Lambda}_{k1}e^{j(\rho_k+\eta_k)\omega_1} \\ \vdots & \dots & \vdots \\ \overline{\Lambda}_{kp}e^{-j\rho_k\omega_p} & \dots & \overline{\Lambda}_{kp}e^{-j(\rho_k+\eta_k)\omega_p} \\ \overline{\Lambda}_{kp}^{*}e^{j\rho_k\omega_p} & \dots & \overline{\Lambda}_{kp}^{*}e^{j(\rho_k+\eta_k)\omega_p} \end{pmatrix} \tag{22}$$

$$\phi_{sin}(t) = \frac{1}{2}\begin{pmatrix} e^{j\omega_1 t} \\ e^{-j\omega_1 t} \\ \vdots \\ e^{j\omega_p t} \\ e^{-j\omega_p t} \end{pmatrix} \tag{23}$$

Therefore,

$$\phi_u(t) = \mathcal{Q}\phi_{sin}(t) \tag{24}$$

where $\mathcal{Q} = \left(\mathcal{Q}_1^T \ \cdots \ \mathcal{Q}_{n_u}^T\right)^T$ is a matrix of dimension $\mu \times 2p$ with μ the dimension of ϕ_u and p the number of frequencies in the multisine (each $\mathcal{Q}_k$ being of dimension $(\eta_k+1)\times 2p$).

An example of construction of the matrix $\mathcal{Q}$ will be given later in the text (see Example 18).

Remark 11 *For the sequel, it is important to note that the matrix $\mathcal{Q}$ will have the property that none of its column is identically zero since, in Section 3, the set of frequencies $(\omega_1, \cdots, \omega_p)$ is defined in such a way that, for each frequency ω_l $(l = 1, \cdots, p)$, there will always be at least one u_k for which $\Lambda_{kl} \neq 0$.*

6.3 Input regressor for filtered white noise

In this section, we consider input vectors u generated as $u(t) = N(z)v(t)$ with a stable matrix $N(z)$ of transfer functions and with a vector v made up of r independent white noises, i.e., each $u_k(t)$ is given by (3). We will rewrite $N(z)$ as $N(z) = L(z)/w(z)$ where $L(z)$ is a $n_u \times r$ matrix of polynomials and $w(z)$ the least common multiple of the denominators of $N(z)$. The entry $L_{kq}(z)$ of $L(z)$ $(k = 1, \cdots, n_u,\ q = 1, \cdots, r)$ will always be of the form

$$L_{kq}(z) = \sum_{h=d_{kq}}^{n_{kq}} l_{kq}^{(h)} z^{-h} \tag{25}$$

with $n_{kq} = \deg(L_{kq}(z))$, d_{kq} the delay of $L_{kq}(z)$ and with $l_{kq}^{(h)}$ the coefficients of the polynomials. For all

$m_k \in [\![\rho_k, \rho_k+\eta_k]\!]$, we can thus express $w(z)u_k(t-m_k)$ as

$$
\begin{aligned}
& w(z)u_k(t-m_k) \\
& = \sum_{q=1}^{r} L_{kq}(z)v_q(t-m_k) \\
& = \sum_{h=d_{k1}}^{n_{k1}} l_{k1}^{(h)} v_1(t-h-m_k) + \cdots + \sum_{h=d_{kr}}^{n_{kr}} l_{kr}^{(h)} v_r(t-h-m_k)
\end{aligned}
$$

Let us consider the above expansions for all entries in ϕ_u and let us define a vector ϕ_{fwn} containing all elements $v_q(t-h-m_k)$ of the white noise sequences v_q $(q=1,\cdots,r)$ that are present in these expansions. For a given q, ϕ_{fwn} will thus contain all elements of $v_q(t)$ from $v_q(t-\Gamma_q)$ till $v_q(t-\Upsilon_q)$ with $\Gamma_q = \min_{k\in[\![1,n_u]\!]} d_{kq}+\rho_k$ and $\Upsilon_q = \max_{k\in[\![1,n_u]\!]} n_{kq}+\eta_k+\rho_k$. The vector ϕ_{fwn} of dimension $\xi = r+\sum_{q=1}^{r}\left(\Upsilon_q-\Gamma_q\right)$ is thus given by:

$$
\phi_{fwn}(t) = \begin{pmatrix} \mathbf{v}_1(t) \\ \vdots \\ \mathbf{v}_r(t) \end{pmatrix} \text{ with } \mathbf{v}_q(t) = \begin{pmatrix} v_q(t-\Gamma_q) \\ v_q(t-\Gamma_q-1) \\ \vdots \\ v_q(t-\Upsilon_q) \end{pmatrix} \tag{26}
$$

Using ϕ_{fwn} and the expression for $w(z)u_k(t-m_k)$, we have that:

$$
w(z)\phi_{u_k}(t) = \underbrace{\begin{pmatrix} \mathscr{L}_{k1} & \cdots & \mathscr{L}_{kr} \end{pmatrix}}_{\mathscr{L}_k} \phi_{fwn}(t)
$$

where $\mathscr{L}_{kq}$ $(q=1,..,r)$ is a matrix of dimension $(\eta_k+1)\times(\Upsilon_q-\Gamma_q+1)$ defined as:

$$
\mathscr{L}_{kq} = \begin{pmatrix} \mathbf{0}_{(\eta_k+1)\times(d_{kq}+\rho_k-\Gamma_q)} & \tilde{\mathscr{L}}_{kq} & \mathbf{0}_{(\eta_k+1)\times(\Upsilon_q-n_{kq}-\eta_k-\rho_k)} \end{pmatrix}
$$

$$
\tilde{\mathscr{L}}_{kq} = \begin{pmatrix} l_{kq}^{(d_{kq})} & \cdots & \cdots & \cdots & l_{kq}^{(n_{kq})} & & \mathbf{0} \\ & \ddots & \cdots & \cdots & \cdots & \ddots & \\ \mathbf{0} & & l_{kq}^{(d_{kq})} & \cdots & \cdots & \cdots & l_{kq}^{(n_{kq})} \end{pmatrix}
$$

We can therefore also write that:

$$
w(z)\phi_u(t) = \mathscr{L}\phi_{fwn}(t) \tag{27}
$$

where $\mathscr{L} = \begin{pmatrix} \mathscr{L}_1^T & \cdots & \mathscr{L}_{n_u}^T \end{pmatrix}^T$ is a matrix of dimension $\mu\times\xi$. To illustrate the construction of this matrix $\mathscr{L}$, let us consider the following example.

Example 12 *Consider for the input regressor* $n_u = 3$, $(\rho_1,\rho_2,\rho_3) = (1,1,3)$, $(\eta_1,\eta_2,\eta_3) = (0,2,1)$, $r = 2$, $w(z) = 1$ *and*

$$
N(z) = L(z) = \begin{pmatrix} 0.2-z^{-2} & z^{-1}-0.8z^{-2}+3z^{-3} \\ 0.5z^{-2} & -z^{-1}+2z^{-2} \\ 0 & 0.4z^{-1}+5z^{-2} \end{pmatrix}
$$

The input regressor ϕ_u *is thus made up of* $u_1(t-1)$, $u_2(t-1)$, $u_2(t-2)$, $u_2(t-3)$, $u_3(t-3)$ *and* $u_3(t-4)$ *(i.e.* $\mu = 6$*). To determine* $\phi_{fwn}(t)$, *we observe that* $\Gamma_1 = 1$, $\Gamma_2 = 2$, $\Upsilon_1 = 5$ *and* $\Upsilon_2 = 6$. *Consequently,* $\phi_{fwn}(t) = \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{v}_2(t) \end{pmatrix}$, *where*

$$
\mathbf{v}_1(t) = \begin{pmatrix} v_1(t-1) \\ v_1(t-2) \\ v_1(t-3) \\ v_1(t-4) \\ v_1(t-5) \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2(t) = \begin{pmatrix} v_2(t-2) \\ v_2(t-3) \\ v_2(t-4) \\ v_2(t-5) \\ v_2(t-6) \end{pmatrix}
$$

We can now construct successively the matrices $\tilde{\mathscr{L}}_{kq}$ $(k=1,\cdots,n_u,\ q=1,\cdots,r)$ *with* $n_u = 3$ *and* $r = 2$ *and the matrix* $\mathscr{L}$:

$$
\tilde{\mathscr{L}}_{11} = \begin{pmatrix} 0.2 & 0 & -1 \end{pmatrix} \qquad \tilde{\mathscr{L}}_{12} = \begin{pmatrix} 1 & -0.8 & 3 \end{pmatrix}
$$

$$
\tilde{\mathscr{L}}_{21} = \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{pmatrix} \qquad \tilde{\mathscr{L}}_{22} = \begin{pmatrix} -1 & 2 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 2 \end{pmatrix}
$$

$$
\tilde{\mathscr{L}}_{31} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \tilde{\mathscr{L}}_{32} = \begin{pmatrix} 0.4 & 5 & 0 \\ 0 & 0.4 & 5 \end{pmatrix}
$$

$$
\mathscr{L} = \left(\begin{array}{c|c} \mathscr{L}_{11} & \mathscr{L}_{12} \\ \hline \mathscr{L}_{21} & \mathscr{L}_{22} \\ \hline \mathscr{L}_{31} & \mathscr{L}_{32} \end{array}\right) = \left(\begin{array}{ccc|ccc} \mathbf{0}_{1\times0} & \tilde{\mathscr{L}}_{11} & \mathbf{0}_{1\times2} & \mathbf{0}_{1\times0} & \tilde{\mathscr{L}}_{12} & \mathbf{0}_{1\times2} \\ \hline \mathbf{0}_{3\times2} & \tilde{\mathscr{L}}_{21} & \mathbf{0}_{3\times0} & \mathbf{0}_{1\times0} & \tilde{\mathscr{L}}_{22} & \mathbf{0}_{3\times1} \\ \hline \mathbf{0}_{2\times2} & \tilde{\mathscr{L}}_{31} & \mathbf{0}_{2\times1} & \mathbf{0}_{1\times2} & \tilde{\mathscr{L}}_{32} & \mathbf{0}_{2\times0} \end{array}\right)
$$

$$
= \left(\begin{array}{ccccc|ccccc} 0.2 & 0 & -1 & \underline{0} & \underline{0} & 1 & -0.8 & 3 & \underline{0} & \underline{0} \\ \hline \underline{0} & \underline{0} & 0.5 & 0 & 0 & -1 & 2 & 0 & 0 & \underline{0} \\ \underline{0} & \underline{0} & 0 & 0.5 & 0 & 0 & -1 & 2 & 0 & \underline{0} \\ \underline{0} & \underline{0} & 0 & 0 & 0.5 & 0 & 0 & -1 & 2 & \underline{0} \\ \hline \underline{0} & \underline{0} & 0 & 0 & \underline{0} & \underline{0} & \underline{0} & 0.4 & 5 & 0 \\ \underline{0} & \underline{0} & 0 & 0 & \underline{0} & \underline{0} & \underline{0} & 0 & 0.4 & 5 \end{array}\right)
$$

where the underlined zeros correspond to the zeros of the $\mathbf{0}_{(\eta_k+1)\times(\Upsilon_q-n_{kq}-\eta_k-\rho_k)}$ *and* $\mathbf{0}_{(\eta_k+1)\times(d_{kq}+\rho_k-\Gamma_q)}$ *for all* $\mathscr{L}_{kq}$. □

Remark 13 *If all the coefficients $l_{kq}^{(h)}$ are nonzero in the expression* (25) *for the filters $L_{kq}(z)$ ($k = 1, \cdots, n_u$ and $q = 1, \cdots, r$), all the columns in $\mathscr{L}$ will be nonzero (as it was the case for $\mathscr{Q}$, see Remark 11). However, if some of these coefficients appear to be zero, this can no longer be the case. As an example, the second column of $\mathscr{L}$ in Example 12 is identically zero. This happens because $L_{11}(z)$ misses the term in z^{-1} and, consequently, the white noise element $v_1(t-2)$ will never appear in the expansions describing $w(z)u_k(t-m_k)$ given below* (25). □

6.4 Necessary and sufficient condition for ϕ_u to be PE

We can now combine Lemma 10 with the matrices $\mathscr{Q}$ and $\mathscr{L}$ introduced in the previous subsections to deduce our main theorem giving a necessary and sufficient condition for ϕ_u to be PE:

Theorem 14 *Consider the regressor ϕ_u of dimension μ (see Theorem 7) and the expressions* (24) *and* (27) *for this regressor corresponding, respectively, to the multisine case (see* (2)*) and the filtered white noise case (see* (3)*). Then, $\phi_u(t)$ is PE if and only if*

- *$\mathscr{Q}$ is full row rank (i.e. rank$(\mathscr{Q}) = \mu$) in the multisine case*
- *$\mathscr{L}$ is full row rank (i.e. rank$(\mathscr{L}) = \mu$) in the filtered white noise case.* □

PROOF. Let us first consider the multisine case. We first observe that $\phi_{sin}(t)$ is (always) PE since $\bar{E}[\phi_{sin}(t)\phi_{sin}^*(t)] = \frac{1}{4}\mathbf{I}_{2p} \succ 0$. Consequently, the result follows from a direct application of Lemma 10. Let us now consider the filtered white noise case. We observe that $\phi_{fwn}(t)$ is (always) PE since it contains elements of independent white noise sequences. Consequently, an application of Lemma 10 yields the equivalence between a full row rank matrix $\mathscr{L}$ and the fact that $w(z)\phi_u(t)$ is PE. By noticing that $w(z)$ is a stable and inversely stable filter, this completes the proof. Indeed, for such $w(z)$, it is straightforward to show using condition (c) of Lemma 5 that $\phi_u(t)$ is PE if and only if $w(z)\phi_u(t)$ is PE. ■

Given the expressions (2) and (3) for u, it is easy to construct, respectively, $\mathscr{Q}$ and $\mathscr{L}$ and check whether the corresponding input signal will yield a persistently exciting ϕ_u and thus informative data for the considered model structure $\mathscr{M}$. Indeed, these matrices are a function of the model structures parameters ρ_k and η_k and of the input vector parametrization. This input vector parametrization is characterized by the amplitudes Λ_{kl} the phase shifts Ψ_{kl} and the frequencies ω_l (for the multisine case) and by the matrix N (filtered white noise case).

We can, e.g., apply the above theorem on the situation described in Example 12. In this example, the matrix $\mathscr{L}$ is full row rank, i.e., rank$(\mathscr{L}) = \mu = 6$. Consequently, the input vector u of dimension 3 generated via $N(z)$ with the white noises v_1 and v_2 will therefore yield informative data for any model structure corresponding to the regressor ϕ_u used in Example 12. It is interesting to note that this is the case even though $\det(\Phi_u(\omega)) = 0$ at all ω.

Theorem 14 gives us thus a tool to verify whether a given input vector will yield informative data for a given model structure. However, it does not give much hints on how to determine the input vector to yield full row rank matrices $\mathscr{Q}$ and $\mathscr{L}$. Such hints will be given in the next sections. We will first start by analyzing the multisine case and then the filtered white noise case.

Remark 15 *If the system is MIMO, the matrix $\mathscr{Q}$ or $\mathscr{L}$ must be constructed for each channel and it should be verified that these n_y matrices are full row rank.*

Remark 16 *In Section 5.2, it was discussed that choosing an input vector u with linearly independent elements (u is PE) is a good practice and is furthermore a necessary condition to yield a persistently exciting ϕ_u for the majority of model structures. We have given in that section a necessary and sufficient condition for u to be PE in the filtered white noise case. We can now do the same for the multisine case. If u is described as* (2)*, u is PE if and only if the following phasor matrix $\mathscr{A}$ is full row rank:*

$$\mathscr{A} = \begin{pmatrix} \overline{\Lambda}_{11} & \overline{\Lambda}_{11}^* & \cdots & \overline{\Lambda}_{1p} & \overline{\Lambda}_{1p}^* \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \overline{\Lambda}_{n_u1} & \overline{\Lambda}_{n_u1}^* & \cdots & \overline{\Lambda}_{n_up} & \overline{\Lambda}_{n_up}^* \end{pmatrix} \tag{28}$$

Let us indeed observe that $u(t) = \mathscr{A}\phi_{sin}(t)$ with $\phi_{sin}(t)$ as defined in (23)*. Consequently, the above result directly follows from Lemma 10 and the fact that $\phi_{sin}(t)$ is PE (see the proof of Theorem 14).*

7 Multisines yielding data informativity

To help us constructing a multisine input vector yielding a persistently exciting ϕ_u (and thus data informativity), we have already discussed a number of aspects in Section 5.2. In particular, we cannot have a persistently exciting ϕ_u if the number p_k of non-zero sinusoids in each u_k does not respect (20). This condition can also be explained by the fact that the matrix $\mathscr{Q}$ introduced in section 6.2 can only be full row rank if all matrices $\mathscr{Q}_k$ are full row rank. Since $\mathscr{Q}_k$ has $\eta_k + 1$ rows and contains $2p_k$ non-zero columns, it is necessary that (20) holds for $\mathscr{Q}_k$ to be full row rank. Using a similar reasoning on the full matrix $\mathscr{Q}$, we can now derive the following additional condition on the total number p of sinusoids in the input vector u.

Lemma 17 *Consider a regression vector $\phi_u(t)$ of dimension μ (see Theorem 7). Suppose that the corresponding input signal u of dimension n_u is a multisine containing p different frequencies (see* (2)*). Then, for ϕ_u to be PE, it* *is necessary to have that*

$$p \geq \frac{\mu}{2} \tag{29}$$

PROOF. The number of columns in $\mathcal{Q}$ is $2p$ and they are all nonzero (see Remark 11) while its number of rows is given by μ. Consequently, it is clear that (29) is a necessary condition for $\mathcal{Q}$ to be full row rank. ■

Recall that $\mu = \sum_{k=1}^{n_u}(\eta_k + 1)$ is larger than $\eta_k + 1$ and that $p \geq p_k\ \forall k$. It is therefore clear that this additional condition is in no way implied by (20). Actually, when $p = p_k$ for each k, it is (29) that implies (20).

In the SISO case, for multisine excitation, a necessary and sufficient condition for data informativity is that the scalar multisine input contains a number of frequencies/sinusoids that is larger than the half of the dimension of the input regressor corresponding to the considered model structure. We have an interesting analogy in the MISO case. For ϕ_u to be PE, the number p of frequencies in the multisine input vector must be larger than the half of the dimension of the input regressor ϕ_u and the number p_k of frequencies in each scalar input u_k of this input vector must also be larger than the half of the dimension of the corresponding part ϕ_{u_k} of ϕ_u. The main difference is however that the combination of these conditions are not sufficient for data informativity in the MISO case (even if we add condition (ii) of Lemma 8). Indeed, some phasor and frequency choices can yield a rank deficient matrix $\mathcal{Q}$ or, equivalently, a regressor ϕ_u that is not PE. An example is a phasor choice that would yield $u_1(t) = -2\ u_2(t-2)$ while $u_1(t)$ and $u_2(t-2)$ both lie in $\phi_u(t)$ (see the next example).

However, the above phenomenon can generally be avoided if we choose the phasor value in some random manner. We can thus propose the following procedure. Determine the regressor ϕ_u corresponding to the considered model structure and choose a value of p satisfying (29), choose a set of p different frequencies and choose values of p_k $(k = 1, \cdots, n_u)$ satisfying $p_k \leq p$ and (20). This defines a class of multisine input vector for which a number of phasors have to be determined. Let us determine these phasors in a random and independent manner[3] and let us check whether the corresponding $\mathcal{Q}$ is indeed full row rank. If it is not the case, a new realization is performed, etc.

Example 18 *Consider an input regressor* $\phi_u(t)$ *with* $n_u = 2$ *and with* $(\rho_1, \rho_2) = (0,0)$ *and* $(\eta_1, \eta_2) = (0,2)$. *Therefore, the regression vector is given by* $\phi_u(t) = \Big(u_1(t)\ \ u_2(t)\ \ u_2(t-1)\ \ u_2(t-2)\Big)^T$ *which has dimension* $\mu = 4$. *Consider that we generate the input with* $p = 2$ *frequencies* $\omega_1 = 0.1$ *and* $\omega_2 = 0.3$ *and that* $p_1 = p_2 = p = 2$. *This defines the following class of input vectors:*

$$u_1(t) = \Lambda_{11}\ \cos(0.1t + \Psi_{11}) + \Lambda_{12}\ \cos(0.3t + \Psi_{12})$$
$$u_2(t) = \Lambda_{21}\ \cos(0.1t + \Psi_{21}) + \Lambda_{22}\ \cos(0.3t + \Psi_{22})$$

This input vector class respects both (29) *and* (20).

Consider first the following choice for the four phasors $\bar{\Lambda}_{kl} = \Lambda_{kl}e^{j\Psi_{kl}}$ $(k = 1,2,\ l = 1,2)$*:* $\bar{\Lambda}_{11} = e^{j0.3}$, $\bar{\Lambda}_{12} = -1$, $\bar{\Lambda}_{21} = -0.5\ e^{j0.5}$ *and* $\bar{\Lambda}_{22} = 0.5\ e^{j0.6}$. *This phasor choice respects condition (ii) of Lemma 8 since* $u_1(t)$ *and* $u_2(t)$ *are linearly independent. Using Remark 16, we indeed observe that the phasor matrix* $\mathcal{A}$ *in* (28)*:*

$$\mathcal{A} = \begin{pmatrix} e^{j0.3} & e^{-j0.3} & -1 & -1 \\ -0.5e^{j0.5} & -0.5e^{-j0.5} & 0.5e^{j0.6} & 0.5e^{-j0.6} \end{pmatrix}$$

is full row rank.

Even though all three necessary conditions are respected, the input regressor $\phi_u(t)$ *is nevertheless not PE since we can observe that* $u_1(t) = -2u_2(t-2)\ \forall t$ *and that both* $u_1(t)$ *and* $u_2(t-2)$ *lies in* ϕ_u. *This can be also verified by constructing the corresponding matrix* $\mathcal{Q}$*:*

$$\mathcal{Q} = \begin{pmatrix} e^{j0.3} & e^{-j0.3} & -1 & -1 \\ -0.5e^{j0.5} & -0.5e^{-j0.5} & 0.5e^{j0.6} & 0.5e^{-j0.6} \\ -0.5e^{j0.4} & -0.5e^{-j0.4} & 0.5e^{j0.3} & 0.5e^{-j0.3} \\ -0.5e^{j0.3} & -0.5e^{-j0.3} & 0.5 & 0.5 \end{pmatrix}$$

We observe that the fourth row is obtained by multiplying the first one by -0.5, *i.e.,* $rank(\mathcal{Q}) < \mu = 4$. *Consequently,* ϕ_u *is not PE with this phasor choice. However, let us just modify the phasor* $\bar{\Lambda}_{11}$ *to* $\bar{\Lambda}_{11} = e^{j0.2}$ *and the matrix* $\mathcal{Q}$ *then becomes full row rank. We can also test the random input vector construction introduced above: we have generated 1000 realizations for the four amplitudes* Λ_{kl} *and the four phase shifts* Ψ_{kl} $(k = 1,2,\ l = 1,2)$ *using independent continuous uniform distributions, the amplitudes varying between* 0 *and* 5 *and the phase shifts between* 0 *and* 2π. *For these 1000 realizations, the rank of the corresponding matrix* $\mathcal{Q}$ *was indeed equal to* $\mu = 4$. *Consequently, the 1000 realizations of the input vector* u *would yield informative data for any model structure corresponding to the regressor* ϕ_u *used in this example. As an example, in the first realization, the four phasors were* $\bar{\Lambda}_{11} = 4.14\ e^{j6.21}$, $\bar{\Lambda}_{12} = 3.15\ e^{j2.16}$, $\bar{\Lambda}_{21} = 1.49\ e^{j1.51}$ *and* $\bar{\Lambda}_{22} = 3.09\ e^{j2.50}$.□

For certain classes of multisine input vector u, the verification that $\mathcal{Q}$ is full row rank is however not necessary: the data informativity can indeed be guaranteed for all phasor choices. An example of such a class is the one

[3] Another option would be to use optimal experiment design to determine the phasors yielding the smallest covariance matrix for a given input power.

where the scalar inputs u_k do not share any common frequencies.

Lemma 19 *Consider the data $s(t) = (u^T(t), y^T(t))^T$ obtained by applying a multisine input vector u (see (2)) on a MISO system with n_u inputs. Consider a full-order model structure $\mathcal{M}$ for this MISO system and the notations introduced in Section 5.1. Then, the data $s(t)$ are informative with respect to $\mathcal{M}$ if, for each $k = 1, \cdots, n_u$,*

(i) the scalars input u_k contains a number p_k of non-zero sinusoids that satisfies:

- $p_k \geq \frac{deg(\mathring{B}_k)+1}{2}$ *when $\mathcal{M}$ is ARX or FIR.*
- $p_k \geq \frac{deg(\mathring{B}_k)+deg(\mathring{F}_k)+1}{2}$ *when $\mathcal{M}$ is BJ or OE.*

(ii) the p_k frequencies of the sinusoids in u_k are different from the p_m frequencies of the sinusoids in all u_m $(m \neq k)$.

Note that, if (i) and (ii) are respected, $s(t)$ are informative whatever the values of the p_k frequencies, the p_k amplitudes (non-zero) and the p_k phase-shifts defining u_k $(k = 1, \cdots, n_u)$. □

PROOF. See Appendix D

It is to be noted that, in the BJ/OE case, the regressor ϕ_u (see (19)) corresponding to the multisine input vectors defined in Lemma 19 will not be necessarily PE since the necessary conditions for this property (i.e. (19) and (29)) will not be always fulfilled. This is however not a contradiction since a persistently exciting ϕ_u is only a sufficient condition for data informativity in the MISO BJ/OE case.

Remark 20 *The approaches presented in this section can be also applied in the MIMO case by noticing that (29) is then replaced by $p \geq \mu_{max}/2$ where μ_{max} is the largest value of μ for the regressors ϕ_u corresponding to each channel and that (20) is replaced by $p_k \geq (\eta_{k,max} + 1)/2$ where $\eta_{k,max}$ is the largest value of η_k for the regressors ϕ_{u_k} corresponding to each channel. For Lemma 19, the condition (i) must be respected for each channel of the considered MIMO system.* □

8 Filtered white noise excitation yielding informative data

As opposed to the multisine case, condition (i) of Lemma 8 is always respected in the filtered white noise case. However, the condition on the relation between the number of rows and of non-zero columns in a full row rank matrix will, like in the multisine case, lead to a necessary condition for ϕ_u to be PE. This condition will impose a constraint on the complexity of $N(z)$. Let us recall that, for the filtered white noise case, there can be some zero-columns among the ξ columns of $\mathcal{L}$ (see Remark 13).

Lemma 21 *Consider a regression vector $\phi_u(t)$ of dimension μ (see Theorem 7) and that corresponds to an input vector generated as in (3). Suppose that the matrix $\mathcal{L}$ of dimension $\mu \times \xi$ corresponding to this input vector (see Section 6.3) has χ non-zero columns ($\chi \leq \xi$). Then, for ϕ_u to be PE, it is necessary that $\chi \geq \mu$.*

PROOF. The number of non-zero columns in $\mathcal{L}$ is χ while its number of rows is given by μ. Consequently, it is necessary that $\chi \geq \mu$ for $\mathcal{L}$ to be full-row rank. ■

As shown in Section 6.3, the number χ of non-zero columns in $\mathcal{L}$ (which is, in the vast majority of the cases, equal to ξ, i.e., the dimension of $\phi_{fwn}(t)$) can be increased by increasing the complexity of the filters in $N(z)$ or by increasing the dimension r of the vector $v(t)$. Consequently, the above lemma gives ingredients on how to generate the input vector u in the filtered white noise case. In, e.g., Example 12, this necessary condition is indeed respected ($\chi = 9 > \mu = 6$) which is logical since we already observed that $\mathcal{L}$ is full column rank. It is however important to stress that $\chi \geq \mu$ is only a necessary condition for ϕ_u to be PE (even if we add condition (ii) of Lemma 8). In other words, even though the complexity of $N(z)$ coupled with the dimension of v is sufficient to ensure $\chi \geq \mu$, there are particular values of the filter matrix $N(z)$ that yield a rank deficient $\mathcal{L}$. To avoid such a phenomenon, similarly as in the multisine case, we can choose a filter structure that satisfies the condition in Lemma 21, generate the corresponding filter coefficients in a random manner (and check subsequently whether the corresponding $\mathcal{L}$ is full-row rank).

Example 22 *Consider an input regressor $\phi_u(t)$ with $n_u = 2$ and with $(\rho_1, \rho_2) = (0, 0)$ and $(\eta_1, \eta_2) = (2, 2)$. The input regressor $\phi_u(t)$ is thus given by $\phi_u(t) = (u_1(t), u_1(t-1), u_1(t-2), u_2(t), u_2(t-1), u_2(t-2))^T$ and is of dimension $\mu = 6$. Consider that we wish to generate the input vector u using:*

$$u(t) = \begin{pmatrix} l_{11} + l_{12}\ z^{-1} \\ l_{21}\ z^{-2} + l_{22}\ z^{-3} \end{pmatrix} v_1(t)$$

with $l_{11} = 1$, $l_{12} = -0.5$, $l_{21} = -2$ and $l_{22} = 1$. If we follow the procedure of Section 6.3, the vector $\phi_{fwn}(t)$ corresponding to this matrix $N(z) = L(z)$ is a vector of dimension $\xi = 6$ containing $v_1(t)$, $v_1(t-1)$, $v_1(t-2)$, $v_1(t-3)$, $v_1(t-4)$ and $v_1(t-5)$. Since the expansion (25) of the two polynomial filters in $N(z)$ do not miss any coefficients, we will have $\chi = \xi = 6$ and, consequently, the necessary condition of Lemma 21 is respected. We also observe that condition (ii) of Lemma 8 is respected since u_1 and u_2 are linearly independent (the rows of $N(z)$ being linearly independent). However, we also observe that $u_2(t) = -2u_1(t-2)$ $\forall t$ and that both $u_2(t)$ and $u_1(t-2)$ lie in ϕ_u. Consequently, the input regressor ϕ_u is not PE and the corresponding $\mathcal{L}$ is not full row rank. However, if we just modify l_{11} to $l_{11} = 2$, the *corresponding matrix $\mathscr{L}$ becomes full row rank. We also apply the random input vector construction by generating 1000 realizations for the four coefficients l_{kq} ($k = 1,2$, $q = 1,2$) using independent zero-mean uniform distributions where each l_{kq} can vary between -1 and 1. For these 1000 realizations, the rank of the corresponding matrix $\mathscr{L}$ was indeed equal to $\mu = 6$. Consequently, the 1000 realizations of the input vector u would yield informative data for any model structure corresponding to the regressor ϕ_u used in this example. As an example, in the first realization, the FIR filters were $l_{11} = -0.398$, $l_{12} = 0.471$, $l_{21} = -0.320$ and $l_{22} = -0.911$.* □

Like in the multisine case, in certain cases, the verification that $\mathscr{L}$ is full row rank is however not necessary. An example of such a class is the one where the number r of independent white noises generating u is larger than or equal to the dimension of u, i.e., $r \geq n_u$. In this case, the verification of the rank of the matrix $\mathscr{L}$ can be replaced by the simpler verification of the rank of $N(e^{j\omega})$.

Lemma 23 *Consider the data $s(t) = (u^T(t), y^T(t))$ obtained by applying an input vector u on a MISO system of arbitrary order and consider a full-order model structure $\mathscr{M}$ for this MISO system. Suppose that the input vector of dimension n_u is generated as a filtered white noise $u = Nv$ with v a vector of r independent white noises with $r \geq n_u$. Then, the data $s(t)$ are informative with respect to $\mathscr{M}$ if rank($N(e^{j\omega})) = n_u$ for almost all ω.* □

PROOF. The power spectrum $\Phi_u(\omega)$ of the input vector u is given by $\Phi_u(\omega) = N(e^{j\omega})\Sigma_v N^*(e^{j\omega})$ where $\Sigma_v = Ev(t)v^T(t)$ is a positive definite diagonal matrix. Under the condition on $N(e^{j\omega})$ in the statement of the lemma, we observe that $\Phi_u(\omega)$ is thus positive definite at almost all frequencies. As shown, e.g., below Theorem 2, an input vector u having this property will yield informative data for all model structures. This completes the proof. ■

A special case of the above lemma is the case where $r = n_u$ and where $N(z)$ is a diagonal matrix of transfer functions (i.e. $N_{ij}(z) = 0$ when $i \neq j$). In this case, $N(e^{j\omega})$ will always be full rank. This is the relatively classical choice of an input vector u where all scalar multisine inputs u_k ($k = 1\cdots, n_u$) are mutually uncorrelated. Note that the scalar inputs u_k were also mutually uncorrelated in the case treated in Lemma 19. Note also that, when $r < n_u$, the condition rank($N(e^{j\omega})) = n_u$ can never be respected.

Remark 24 *The approaches presented in this section can be also applied in the MIMO case by noticing that the condition in Lemma 21 can be replaced by $\chi \geq \mu_{max}$ where μ_{max} is the largest value of μ for the regressors ϕ_u corresponding to each channel. The input vector choice discussed in Lemma 23 also yield informative data for arbitrary MIMO model structures.*

9 Numerical example

9.1 True system to be identified

Consider the following MISO OE system $\mathscr{S}$ with $n_y = 1$ and $n_u = 2$ given by

$$y(t) = \begin{pmatrix} \frac{0.34}{1-0.53z^{-1}} & \frac{z^{-1}}{1-0.8z^{-1}+0.15z^{-2}} \end{pmatrix} \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} + e(t)$$

where the variance σ_0^2 of e is here chosen equal to $\sigma_0^2 = 0.0001$. We will identify $\mathscr{S}$ within a full-order model structure $\mathscr{M}$, where all the transfers in $G(z,\theta)$ are independently parametrized as in (16) with $(\rho_1, \rho_2) = (0,1)$ and

$$\begin{aligned} \mathring{B}_1 &= \theta_1 & \mathring{B}_2 &= \theta_3 \\ \mathring{F}_1 &= 1 + \theta_2 z^{-1} & \mathring{F}_2 &= 1 + \theta_4 z^{-1} + \theta_5 z^{-2} \end{aligned}$$

Consequently, we have $\theta^T = \tilde{\theta}^T = \begin{pmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_5 \end{pmatrix}$. Let us now determine the regressor ϕ_u corresponding to this model structure $\mathscr{M}$. For this purpose, observe that:

- $\eta_1 = \deg(\mathring{B}_1) + \deg(\mathring{F}_1) + 2\deg(\mathring{F}_2) = 5$
- $\eta_2 = \deg(\mathring{B}_2) + \deg(\mathring{F}_2) + 2\deg(\mathring{F}_1) = 4$

Consequently, the regressor ϕ_u is given by $\phi_u = (\phi_{u_1}^T, \phi_{u_2}^T)^T$ with

$$\phi_{u_1}^T(t) = \begin{pmatrix} u_1(t) & u_1(t-1) & u_1(t-2) & u_1(t-3) & u_1(t-4) & u_1(t-5) \end{pmatrix}$$
$$\phi_{u_2}^T(t) = \begin{pmatrix} u_2(t-1) & u_2(t-2) & u_2(t-3) & u_2(t-4) & u_2(t-5) \end{pmatrix}$$

The dimension μ of ϕ_u is here equal to $\mu = 11$.

9.2 Identification with multisine

In this subsection, we consider that $u(t)$ is given by (2).

As shown in Theorem 7, the persistency of ϕ_u is only a sufficient condition for data informativity. However, we will show that this sufficient condition is not too restrictive. This will be done as follows. We will present an example where ϕ_u is not PE and where the data are not informative. Moreover, we will also show that if the dimension of ϕ_u is reduced by removing $u_2(t-5)$ in $\phi_{u_2}^T$, this reduced ϕ_u would be PE. In this example, the regressor ϕ_u defined in Theorem 7 is the regressor that can only detect the non informativity of these data.

Case 1: We choose $u_1(t)$ as a multisine containing $p = p_1 = 6$ frequencies $\omega_1 = 0.002\pi$, $\omega_2 = 0.02\pi$, $\omega_3 = 0.1\pi$, $\omega_4 = 0.2\pi$, $\omega_5 = 0.4\pi$ and $\omega_6 = 0.8\pi$ and with phasors $\overline{\Lambda}_{1l} = 1$ ($l = 1, \cdots, 6$). The signal u_2 is chosen in such a way that

$$(G_2(z,\theta') - G_2(z,\theta_0))u_2(t) = -(G_1(z,\theta') - G_1(z,\theta_0))u_1(t)$$

for $\theta' = (10, 0.9, -3, -0.6, 0.8)^T$ different from $\theta_0 = (0.34, -0.53, 1, -0.8, 0.15)^T$.

This yields:

$$\begin{pmatrix} \overline{\Lambda}_{11} & \overline{\Lambda}_{21} \\ \overline{\Lambda}_{12} & \overline{\Lambda}_{22} \\ \overline{\Lambda}_{13} & \overline{\Lambda}_{23} \\ \overline{\Lambda}_{14} & \overline{\Lambda}_{24} \\ \overline{\Lambda}_{15} & \overline{\Lambda}_{25} \\ \overline{\Lambda}_{16} & \overline{\Lambda}_{26} \end{pmatrix} = \begin{pmatrix} 1 & 0.8475e^{j0.0132} \\ 1 & 0.8526e^{j0.1315} \\ 1 & 0.9633e^{j0.5926} \\ 1 & 1.1325e^{j0.9076} \\ 1 & 0.3875e^{j2.4547} \\ 1 & 8.3699e^{-j2.1509} \end{pmatrix}$$

This phasor and frequency choice respects both (20) and (29). It is however clear that these data are not informative (see Theorem 2) and this non-informativity can be detected via Theorem 7. Indeed, the regressor ϕ_u defined in Section 9.1 is not PE since the matrix $\mathscr{Q}$ corresponding to this ϕ_u and the above phasor and frequency choice has a rank equal to 10 while $\mu = 11$. This regressor is in fact the only one that can detect the non-informativity. Indeed, if we would consider the regressor ϕ_u of dimension 10 obtained by removing $u_2(t-5)$ from ϕ_u, this regressor would be PE. Note that we can generate infinite number of pathological cases as the previous one by considering all $\theta' \in \mathscr{D}_\theta$.

Case 2: Consider now a multisine input vector u with the same $p = 6$ frequencies as in Case 1, but with the following choice of phasors:

$$\begin{pmatrix} \overline{\Lambda}_{11} & \overline{\Lambda}_{21} \\ \overline{\Lambda}_{12} & \overline{\Lambda}_{22} \\ \overline{\Lambda}_{13} & \overline{\Lambda}_{23} \\ \overline{\Lambda}_{14} & \overline{\Lambda}_{24} \\ \overline{\Lambda}_{15} & \overline{\Lambda}_{25} \\ \overline{\Lambda}_{16} & \overline{\Lambda}_{26} \end{pmatrix} = \begin{pmatrix} 2.54e^{j0.32} & 3.67 \\ -0.56^{j2.41} & -2.25e^{-j0.51} \\ 7.34 & 5.84 \\ 2.3 & 6.73e^{-j1.7} \\ -8e^{-j1.45} & 9.78 \\ -0.45e^{j0.87} & 6.54e^{j3.01} \end{pmatrix}$$

For this choice of phasors, the matrix $\mathscr{Q}$ in (24) has full row rank, i.e., rank$(\mathscr{Q}) = 11$. This multisine input vector will thus yield informative data.

Case 3: Consider now a multisine input vector with $p = 3$ sinusoid frequencies $\omega_1 = 0.002\pi$, $\omega_2 = 0.02\pi$, $\omega_3 = 0.1\pi$ and with the following phasor choice:

$$\begin{pmatrix} \overline{\Lambda}_{11} & \overline{\Lambda}_{21} \\ \overline{\Lambda}_{12} & \overline{\Lambda}_{22} \\ \overline{\Lambda}_{13} & \overline{\Lambda}_{23} \end{pmatrix} = \begin{pmatrix} 2.54e^{j0.32} & 0 \\ 0 & -2.25e^{-j0.51} \\ 0 & 5.84 \end{pmatrix}$$

The scalar inputs u_1 and u_2 do not share any common frequencies and we have that $p_1 = 1 = \frac{\deg(\mathring{B}_1)+\deg(\mathring{F}_1)+1}{2}$ and $p_2 = 2 \geq \frac{\deg(\mathring{B}_2)+\deg(\mathring{F}_2)+1}{2}$. Consequently, we are in the case of the class of multisine input vector proposed in Lemma 19. This class yields informative data independently of the choice of phasors and frequencies.

9.3 Identification with filtered white noise

In this subsection, we consider that $u(t)$ is given by (3).

Case 4: Consider an input vector generated with $r = 1$ independent white noise $v_1(t)$, i.e.,

$$u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = \begin{pmatrix} N_{11}(z) \\ N_{21}(z) \end{pmatrix} v_1(t)$$

where $N_{21}(z) = z^{-2}$ and $N_{11}(z) = \cdots$

$$\frac{-0.4z^{-1} + 0.95z^{-2} - 0.61z^{-3} + 0.17z^{-4}}{1 - 1.1z^{-1} - 0.22z^{-2} + 0.41z^{-3} - 0.008z^{-4} - 0.034z^{-5}}$$

In this case, the condition of Lemma 21 is respected. We therefore verify the rank of the corresponding matrix $\mathscr{L}$ and we obtain rank$(\mathscr{L}) = 11 = \mu$. Consequently, this input vector will yield informative data.

Case 5: Let us now consider an input vector u with $r = n_u = 2$ and where $N(z) = L(z)/w(z)$ is characterized by $w(z) = 1 - 0.5z^{-1} + 0.6z^{-2}$ and

$$L(z) = \begin{pmatrix} -z^{-1} + 0.2z^{-2} & -0.5 + 0.15z^{-1} \\ 2 - 0.4z^{-1} & 0.6z^{-2} + 0.18z^{-3} \end{pmatrix}$$

We are here in the situation of Lemma 23 and we can easily verify that the frequency response $N(e^{j\omega})$ of $N(z)$ is full-row rank for all $\omega \in]-\pi, \pi]$.

9.4 Monte-Carlo simulations

In order to confirm that the input choice in Cases 2, 3, 4 and 5 yield informative data, we have applied each of these input vectors to the true system in 1000 identification experiments (with different realizations of the white noise e) and we have identified the estimate $\hat{\theta}_N$ (see (8)) for each experiment. For each input vector, we have computed the mean of these 1000 estimates and we have in each case observed that this mean is almost equal to the true parameter vector θ_0. For example, the observed mean obtained in Case 3 is $(0.3402, -0.5298, 1.0000, -0.8000, 0.15000)^T$ that must be compared to $\theta_0 = (0.34, -0.53, 1, -0.8, 0.15)^T$. The discrepancy between the observed mean and θ_0 is in fact the largest in Case 3 (this can be explained by the fact that the three frequencies of the multisine are chosen in a small interval of $[0, \pi]$ inducing a larger variance for $\hat{\theta}_N$).

10 Conclusion

In this paper, we have derived a condition that allows one to check whether a given input vector will yield informative data to identify a model of a MIMO/MISO system in open loop in a given model structure. We have done that for the classical model structures used in prediction-error identification and for both multisine and filtered white noise input vectors. We have seen that this condition can be easily checked a-posteriori based on the input vector parametrization. We have also given hints on how to construct an input vector that will yield informative data. In the future, we wish to extend our results to the closed-loop case.

References

[1] Märta Barenthin, Xavier Bombois, Hakan Hjalmarsson, and Gérard Scorletti. Identification for control of multivariable systems: Controller validation and experiment design via LMIs. *Automatica*, 44(12):3070–3078, December 2008.

[2] Alexandre Sanfelice Bazanella, Xavier Bombois, and Michel Gevers. Necessary and sufficient conditions for uniqueness of the minimum in prediction error identification. *Automatica*, 48(8):1621 – 1630, August 2012.

[3] Alexandre Sanfelice Bazanella, Michel Gevers, and Ljubiša Miškovic. Closed-loop identification of MIMO systems: a new look at identifiability and experiment design. *European Journal of Control*, 16(3):228–239, 2010.

[4] Xavier Bombois, Gérard Scorletti, Michel Gevers, Paul MJ Van den Hof, and Roland Hildebrand. Least costly identification experiment for control. *Automatica*, 42(10):1651–1662, October 2006.

[5] Kévin Colin, Xavier Bombois, Laurent Bako, and Federico Morelli. Data informativity for the identification of MISO FIR system with filtered white noise excitation. Accepted for presentation at CDC 2019. https://hal.archives-ouvertes.fr/hal-02161598v1.

[6] Kévin Colin, Xavier Bombois, Laurent Bako, and Federico Morelli. Informativity: how to get just sufficiently rich in the MISO FIR case? In *2019 18th European Control Conference (ECC)*, pages 351–356, June 2019. https://hal.archives-ouvertes.fr/hal-02070880v1.

[7] Paul M.J. Van den Hof, Arne Dankers, Peter S.C. Heuberger, and Xavier Bombois. Identification of dynamic models in complex networks with prediction error methods—basic methods for consistent module estimates. *Automatica*, 49(10):2994 – 3006, October 2013.

[8] Michel Gevers, Alexandre Bazanella, and Ljubiša Miškovic. Informative data: How to get just sufficiently rich? In *2008 47th IEEE Conference on Decision and Control.*, December 2008.

[9] Michel Gevers, Alexandre Sanfelice Bazanella, Xavier Bombois, and Ljubiša Miškovic. Identification and the information matrix: How to get just sufficiently rich? *IEEE Transactions on Automatic Control*, 54(12):2828–2840, December 2009.

[10] Henrik Jansson and Håkan Hjalmarsson. Optimal experiment design in closed loop. *IFAC Proceedings Volumes*, 38(1):488–493, 2005.

[11] Abhishankar Kumar, M. Nabil, and Sridharakumar Narasimhan. Economical Input Design for Identification of Multivariate Systems. *IFAC-PapersOnLine*, 48(28):1313 – 1318, 2015. 17th IFAC Symposium on System Identification SYSID 2015.

[12] Lennart Ljung. *System identification: Theory for the user.* Prentice Hall information and system sciences series. Prentice Hall PTR, Upper Saddle River (NJ), second edition edition, 1999.

[13] Tung-Sang Ng, Graham C Goodwin, and Brian DO Anderson. Identifiability of MIMO linear dynamic systems operating in closed loop. *Automatica*, 13(5):477–485, September 1977.

[14] Torsten Soderstrom and Petre Stoica. *System Identification.* Prentice Hall, 1989.

A Proof of Theorem 2

To prove Theorem 2, we will need the following lemma.

Lemma 25 *Consider a quasi-stationary input vector u of the type* (2) *or* (3). *Consider also a stable and inversely stable matrix $V(z)$ of transfer functions ($V(z) \neq 0$) and another stable matrix $W(z)$ of transfer functions such that $V(z)W(z)$ and $V(z)W(z)u(t)$ are valid operations yielding, respectively, a matrix of transfer functions and a vector of signals. Then, we have the following equivalences*

$$\bar{E}\left[||V(z)W(z)u(t)||^2\right] = 0 \Longleftrightarrow \bar{E}\left[||W(z)u(t)||^2\right] = 0$$

$$V(z)W(z) \equiv 0 \Longleftrightarrow W(z) \equiv 0$$

PROOF. Let us begin by the first equivalence. Since $V(z)$ is inversely stable, $\bar{E}\left[||V(z)W(z)u(t)||^2\right] = 0$ implies that $\bar{E}\left[||V^{-1}(z)V(z)W(z)u(t)||^2\right] = \bar{E}\left[||W(z)u(t)||\right]^2 = 0$. Now, since $V(z)$ is stable, $E\left[||W(z)u(t)||^2\right] = 0$ implies that $\bar{E}\left[||V(z)W(z)u(t)||^2\right] = 0$. This completes the proof of the first equivalence. We can follow a similar reasoning for the second equivalence, i.e., $V(z)W(z) \equiv 0 \Longrightarrow V^{-1}(z)V(z)W(z) = W(z) \equiv 0 \Longrightarrow V(z)W(z) \equiv 0$. ■

Let us now prove Theorem 2. By using (1) and (4), $\Delta W(z)s(t)$ can be rewritten as

$$\begin{aligned}&\Delta W(z)s(t)\\&= \left(\Delta W_u(z) + \Delta W_y(z)G_0(z)\right)u(t) + \Delta W_y(z)H_0(z)e(t)\end{aligned}$$

where $\Delta W_u(z)$ and $\Delta W_y(z)$ are defined similarly as $\Delta W(z)$.

Since $u(t)$ and $e(t)$ are assumed independent, $\bar{E}\left[||\Delta W(z)s(t)||^2\right] = 0$ is equivalent to

$$\begin{aligned}\bar{E}\left[||(\Delta W_u(z) + \Delta W_y(z)G_0(z))u(t)||^2\right] &= 0\\ \bar{E}\left[||\Delta W_y(z)H_0(z)e(t)||^2\right] &= 0\end{aligned}$$

Observe that, due to the whiteness of the vector e and the invertibility of $H_0(z)$, the spectrum $\Phi_{\bar{v}}(\omega)$ of $\bar{v}(t) = H_0(z)e(t)$ has the property that $\Phi_{\bar{v}}(\omega) \succ 0$ at all frequencies ω. Consequently, we can also say that $\bar{E}\left[||\Delta W(z)s(t)||^2\right] = 0$ is equivalent to the combination of the following two conditions:

$$\bar{E}\left[||\Delta W_u(z)u(t)||^2\right] = 0 \tag{A.1}$$

$$\Delta W_y \equiv 0 \tag{A.2}$$

Due to the expression of $W_y(z)$ in (6), (A.2) yields $H(z,\theta') = H(z,\theta'')$. Consequently, $\Delta W_u(z)$ (see (5)) can be rewritten as $\Delta W_u(z) = Y\left(z,\eta'\right)^{-1}\Delta X(z)$ with ΔX as defined in the statement of the theorem for the different model structures and with $Y\left(z,\eta'\right) = \mathbf{I}_{n_y}$ for ARX/FIR model structures and $Y\left(z,\eta'\right) = H\left(z,\eta'\right)$ for BJ/OE model structures.

Since we restrict attention to $\theta' = (\tilde{\theta}'^T,\eta'^T)^T \in \mathscr{D}_\theta$, $Y\left(z,\eta'\right)$ is a (non-zero) stable and inversely-stable matrix of transfer functions in both situations. Consequently, by Lemma 25, $\bar{E}\left[||\Delta W_u(z)u(t)||^2\right] = 0$ is equivalent to $\bar{E}\left[||\Delta X(z)u(t)||^2\right] = 0$.

From the reasoning above, for all $\Delta W(z) \in \Delta_\mathbf{W}$, the left-hand side of (8) is equivalent to $\bar{E}\left[||\Delta X(z)u(t)||^2\right] = 0$. From the reasoning above, we can also conclude that, for all $\Delta W(z) \in \Delta_\mathbf{W}$, the left-hand side of (8) always implies $\Delta W_y(z) \equiv 0$. Consequently, the right-hand side (8) can be restricted to $\Delta W_u(z) \equiv 0$ or, equivalently, via Lemma 25, to $\Delta X(z) \equiv 0$. In other words, the condition (13) in Theorem 2 (for all $\Delta X(z) \in \Delta_\mathbf{X}$) is equivalent to the condition (8) (for all $\Delta W(z) \in \Delta_\mathbf{W}$) in the definition of the data informativity. This completes thus the proof. ■

B Proof of Theorem 3

We will prove that the condition (13) for $\Delta X(z) \in \Delta_\mathbf{X}$ is equivalent to the fact that the condition (14) holds for all $\Delta X_{i:} \in \Delta_{\mathbf{X,i}}$ and for all $i = 1,\cdots,n_y$.

Let us first observe that $\Delta X(z) \equiv 0$ is equivalent to $\Delta X_{i:}(z) \equiv 0$ $(i = 1,\cdots,n_y)$. Secondly, let us prove that $\bar{E}\left[||\Delta X(z)u(t)||^2\right] = 0$ is equivalent to $\bar{E}\left[||\Delta X_{i:}(z)u(t)||^2\right] = 0$ $(i = 1,\cdots,n_y)$. For this purpose, let us rewrite the term $\bar{E}\left[||\Delta X(z)u(t)||^2\right]$ as follows:

$$\bar{E}\left[||\Delta X(z)u(t)||^2\right] = \sum_{i=1}^{n_y}\bar{E}\left[||\Delta X_{i:}(z)u(t)||^2\right]$$

with $\Delta X_{i:}(z)$ as defined in the theorem. Since the term $\bar{E}\left[||\Delta X_{i:}(z)u(t)||^2\right]$ is positive $(i = 1,\cdots,n_y)$, we have indeed that $\bar{E}\left[||\Delta X(z)u(t)||^2\right] = 0$ is equivalent to $\bar{E}\left[||\Delta X_{i:}(z)u(t)||^2\right] = 0$ $(i = 1,\cdots,n_y)$.

We have thus proven that the condition (13) for $\Delta X(z) \in \Delta_\mathbf{X}$ is equivalent to

$$\bar{E}\left[||\Delta X_{i:}(z)u(t)||^2\right] = 0\ (i = 1,\cdots,n_y) \Longrightarrow \Delta X_{i:}(z) \equiv 0\ (i = 1,\cdots,n_y)$$

for $\Delta X_{i:} \in \Delta_{\mathbf{X,i}}$ $(i = 1,\cdots,n_y)$. Noting that there are no common parameters in $\Delta X_{i:}(z)$ and $\Delta X_{j:}(z)$ $(j \neq i)$, the latter is equivalent to the fact that the condition (14) holds for all $\Delta X_{i:} \in \Delta_{\mathbf{X,i}}$ and for all $i = 1,\cdots,n_y$. This completes thus the proof. ■

C Proof of Theorem 7

First note that Theorem 2 also applies to the case $n_y = 1$ and let us distinguish the ARX/FIR case and the BJ/OE case in the sequel.

Let us first consider ARX/FIR model structures and let us first observe that, using (17) and the fact that $\tilde{\theta} = (\tilde{\theta}_1^T,\tilde{\theta}_2^T,\cdots,\tilde{\theta}_{n_u}^T)^T$, we have that $B_k(z,\tilde{\theta}_k)u_k(t) = \phi_{u_k}(t)^T\tilde{\theta}_k$ and $B(z,\tilde{\theta})u(t) = \phi_u^T(t)\tilde{\theta}$ with $\phi_{u_k}(t)$ and $\phi_u(t)$ as defined in (19). Let us now rewrite the left-hand side of condition (13), i.e., $\bar{E}\left[||\Delta B(z)u(t)||^2\right] = 0$ where $\Delta B(z)$ is a row vector of transfer functions of dimension k whose entries $\Delta B_k(z)$ are polynomials defined as $\Delta B_k(z) = z^{-\rho_k}\left(\mathring{B}_k(z,\tilde{\theta}'_k) - \mathring{B}_k(z,\tilde{\theta}''_k)\right)$. We have thus that

$$\begin{aligned}\Delta B_k(z)u_k(t) &= \phi_{u_k}(t)^T(\tilde{\theta}'_k - \tilde{\theta}''_k)\\ \Delta B(z)u(t) &= \phi_u(t)^T(\tilde{\theta}' - \tilde{\theta}'')\end{aligned} \tag{C.1}$$

Consequently, $\bar{E}\left[||\Delta B(z)u(t)||^2\right] = 0$ is equivalent to

$$\left(\tilde{\theta}' - \tilde{\theta}''\right)^T\bar{E}\left[\phi_u(t)\phi_u(t)^T\right]\left(\tilde{\theta}' - \tilde{\theta}''\right) = 0 \tag{C.2}$$

Let us now observe that the right-hand side of (13) is here equivalent to $\tilde{\theta}' = \tilde{\theta}''$ (see (17)). Recall also that $\mathscr{D}_{\tilde{\theta}} = \mathbb{R}^\mu$ (see Section 5.1 and recall that, here, the dimension $\tilde{n}$ of $\tilde{\theta}$ is equal to the dimension μ of ϕ_u). Using the above reasoning and Theorem 2, we will thus have informativity if and only if, for any $\tilde{\theta}'$ and $\tilde{\theta}''$ in $\mathbb{R}^\mu$, (C.2) implies $\tilde{\theta}' = \tilde{\theta}''$. This latter condition is equivalent to $\bar{E}\left[\phi_u(t)\phi_u(t)^T\right] \succ 0$ or that $\phi_u(t)$ is PE (Lemma 5). This completes thus the proof for the ARX/FIR model structures.

Let us now consider BJ/OE model structures and let us here also rewrite the left-hand side of condition (13), i.e., $\bar{E}\left[||\Delta G(z)u(t)||^2\right] = 0$. For this purpose, let us observe that the transfer functions $\Delta G_k(z)$ in the row vector $\Delta G(z) = (\Delta G_k(z))_{k\in[\![1,n_u]\!]}$ are given by

$$\Delta G_k(z) = z^{-\rho_k}\left(\frac{\mathring{B}_k(z,\tilde{\theta}'_k)}{\mathring{F}_k(z,\tilde{\theta}'_k)} - \frac{\mathring{B}_k(z,\tilde{\theta}''_k)}{\mathring{F}_k(z,\tilde{\theta}''_k)}\right)$$

Using obvious short-hand notations, the latter equation is rewritten as:

$$\Delta G_k(z) = z^{-\rho_k}\left(\frac{\mathring{B}'_k}{\mathring{F}'_k} - \frac{\mathring{B}''_k}{\mathring{F}''_k}\right)$$

Let us now put all entries of $\Delta G(z)$ on the same denominator which will be denoted by $N(z)$, i.e., $N(z) = \prod_{k=1}^{n_u}\mathring{F}'_k\mathring{F}''_k$. Consequently, we have that $\Delta G(z) = \frac{M(z)}{N(z)}$ where $M(z) = (M_k)_{k\in[\![1,n_u]\!]}$ is a vector of polynomials $M_k(z)$ having the following expression:

$$M_k(z) = z^{-\rho_k}(\mathring{B}'_k\mathring{F}''_k - \mathring{B}''_k\mathring{F}'_k)\prod_{m=1,m\neq k}^{n_u}\mathring{F}'_m\mathring{F}''_m \tag{C.3}$$

Each $M_k(z)$ is thus a polynomial of delay ρ_k and of degree $\rho_k + \eta_k$ (with η_k as defined in the statement of the theorem) and can thus be rewritten as follows:

$$M_k(z) = \sum_{m=\rho_k}^{\eta_k+\rho_k}\delta_k^{(m)} z^{-m} \tag{C.4}$$

where the coefficients $\delta_k^{(m)}$ are a known function of $\tilde{\theta}'$ and $\tilde{\theta}''$. Combining the above elements successively yields:

$$\Delta G(z)u(t) = \frac{1}{N(z)}\left(\sum_{k=1}^{n_u}\sum_{m=\rho_k}^{\eta_k+\rho_k}\delta_k^{(m)}u_k(t-m)\right)$$

$$\Delta G(z)u(t) = \frac{1}{N(z)}\underbrace{\left(\phi_u(t)^T\delta\right)}_{=x(t)}$$

with δ the vector made up of the concatenation of all $\delta_k^{(m)}$ and ϕ_u as given in the statement of the theorem.

Since $\tilde{\theta}'$ and $\tilde{\theta}''$ are in $\mathcal{D}_{\tilde{\theta}} = \{\tilde{\theta} \mid G(z,\tilde{\theta}) \text{ is stable}\}$, $N(z)$ is a stable and inversely stable filter. Consequently, via Lemma 25 in Appendix A, the left-hand side of (13) (i.e. $\bar{E}\left[||\Delta G(z)u(t)||^2\right] = 0$) is equivalent to

$$\bar{E}\left[x^2(t)\right] = \delta^T\bar{E}\left[\phi_u(t)\phi_u(t)^T\right]\delta = 0 \tag{C.5}$$

When ϕ_u is PE, i.e., when $\bar{E}[\phi_u(t)\phi_u^T(t)] \succ 0$ (Lemma 5), we have that, for any δ generated as above with $\tilde{\theta}'$ and $\tilde{\theta}'' \in \mathcal{D}_{\tilde{\theta}}$, (C.5) implies $\delta = 0$ which in turn implies that $M(z) \equiv 0$ and thus that $\Delta G(z) \equiv 0$ (i.e. the right-hand side of (13)). The sufficient condition for the BJ/OE case in this theorem is then a consequence of Theorem 2. Note that this condition is not necessary in the BJ/OE case since the vectors δ that are generated by all $\tilde{\theta}'$ and $\tilde{\theta}'' \in \mathcal{D}_{\tilde{\theta}}$ will not cover the whole vectorial space R^{μ}. Observe indeed that the dimension μ of δ is (much) larger than the dimension of $\tilde{\theta}'$ and $\tilde{\theta}''$.

D Proof of Lemma 19

Let us consider Theorem 2 for the MISO case ($n_y = 1$) and let us define, for each $k = 1, \cdots, n_u$, the set

$$\Delta\mathbf{X_k} = \left\{\Delta X_k = X_k(z,\tilde{\theta}'_k) - X_k(z,\tilde{\theta}''_k) \text{ with } \tilde{\theta}'_k \text{ and } \tilde{\theta}''_k \in \mathcal{D}_{\tilde{\theta}_k}\right\}$$

with $\mathcal{D}_{\tilde{\theta}_k} = \{\tilde{\theta}_k \mid X_k(z,\tilde{\theta}_k) \text{ is stable}\}$ and $X_k(z,\tilde{\theta}_k) = B_k(z,\tilde{\theta}_k)$ (ARX/FIR case) or $X_k(z,\tilde{\theta}_k) = G_k(z,\tilde{\theta}_k)$ (BJ/OE case).

Now, because of condition (ii), let us observe that the scalar inputs u_k are all mutually uncorrelated. Consequently, the condition (13) can be successively rewritten as:

$$\sum_{k=1}^{n_u}\bar{E}\left[\Delta X_k(z)u_k(t)\right]^2 = 0 \Longrightarrow \Delta X_k \equiv 0 \;\forall k \in [\![1, n_u]\!] \tag{D.1}$$

$$\bar{E}\left[\Delta X_k(z)u_k(t)\right]^2 = 0 \;\; k \in [\![1, n_u]\!] \Longrightarrow \Delta X_k \equiv 0 \;\forall k \in [\![1, n_u]\!] \tag{D.2}$$

Since they are no common parameters in ΔX_k and ΔX_m ($m \neq k$), we have that (D.2) is also equivalent to the fact that the following condition holds for each $k = 1, \cdots, n_u$:

$$\bar{E}\left[\Delta X_k(z)u_k(t)\right]^2 = 0 \Longrightarrow \Delta X_k \equiv 0 \tag{D.3}$$

From the above reasoning and Theorem 2, u yields informative data with respect to $\mathcal{M}$ if and only if, for each $k = 1, \cdots, n_u$, (D.3) holds for all $\Delta X_k \in \Delta\mathbf{X_k}$.

By Theorem 2 applied to the SISO case ($n_u = 1, n_y = 1$) for each individual k, that (D.3) holds for all $\Delta X_k \in \Delta\mathbf{X_k}$ is equivalent to the fact that u_k yields informative data for a SISO model structure described[4] by $B_k(z,\tilde{\theta}_k)$ (in the ARX/FIR case) and by $G_k(z,\tilde{\theta}_k)$ (in the BJ/OE case). Lemma 19 then follows from the classical necessary and sufficient conditions for data informativity in the SISO case [8]. ■

[4] The matrices A and H do not play any role for data informativity (Theorem 2).