Data Informativity for the Identication of MISO FIR Systems with Filtered White Noise Excitation

Kévin Colin, Xavier Bombois, Laurent Bako, Federico Morelli

▶ To cite this version:

HAL Id: hal-02161598
https://hal.archives-ouvertes.fr/hal-02161598

Submitted on 20 Jun 2019

Data Informativity for the Identification of MISO FIR Systems with Filtered White Noise Excitation

Kevin Colin[1], Xavier Bombois[1], Laurent Bako[1] and Federico Morelli[1*]

June 20, 2019

This technical report presents a pre-print version of a paper that has been submitted for the proceedings of CDC 2019.

Abstract

For Prediction Error Identification, there are two main ingredients to get a consistent estimate: one of them is the data informativity with respect to (w.r.t.) the considered model structure. One common criterion used for the informativity is the positive definiteness of the input density spectral power (DSP) matrix at all frequencies. This criterion is not appropriate for multisine excitation but can be used for filtered white noise excitation for many identification problems. However, this criterion is not necessary and its application for some identification problems might not be possible. In this paper, we propose a necessary and sufficient condition for the data informativity in the case of multiple-inputs single-output (MISO) finite impulse response (FIR) model structure in open-loop.

1 Introduction

When identifying a system with the Prediction Error Method, there are two properties to respect in order to guarantee a consistent estimation: global identifiability of the model structure and data informativity w.r.t. this model structure [1] [2]. In this paper, we focus on the data informativity criterion in the open-loop configuration.

This data informativity criterion has been extensively studied in the case of SISO systems [1] [2] [3]. However, even if one could have expected that this study had also been carried out for the multivariate case and in particular for the MISO systems, this is surprisingly not the case, up to our knowledge. In the literature, we only find the condition that the power spectrum matrix of the input vector must be strictly positive definite at each frequency [2] [3]. This condition is

---

[*1] Laboratoire Ampère, UMR CNRS 5005, Ecole Centrale de Lyon, Université de Lyon, Ecully, France, `kevin.colin@ec-lyon.fr`

sufficient, but not necessary for data informativity as e.g. shown in [4] where we derive conditions for data informativity in the MISO case when multisine excitation are used. In this paper, we continue the analysis of the data informativity of MISO systems by considering the case of an input vector where each element is a stochastic process. In other words, we consider the case where each input is generated as a filtered version of a number of white noise processes and we derive necessary and sufficient condition under which the data informativity is guaranteed.

From this condition, we will, e.g., observe that input vectors that have a power spectrum matrix with a zero determinant at each frequency can lead to informative data.

It is to be noted that the latter case will be quite unusual if the user can freely choose the input vector (one would then generally go for independent input processes leading to positive definite power spectrum matrix at each frequency). However, if the input vector is not chosen but given, it is important to know precisely in which situation we will have informative data and in which cases not. Such a situation is e.g. encountered when the two stage approach is used for dynamic network identification [5].

In the case of the identification of MISO FIR systems in open loop, the informativity condition is equivalent to the condition that the input regressor is persistently exciting (PE), equivalent to the linear independence in the signal space [6] of the signals in this regressor. By choosing an appropriate formulation of the input expression, we find a simpler necessary and sufficient condition which is the right-invertibility of a certain matrix containing the coefficients of the filter generating the inputs. We also develop additional conditions to guide the user in its filter choice to guarantee the data informativity. A numerical example is provided to illustrate the developed results.

Notations. For all matrix A, A^T denotes its transpose, A^* its conjugate transpose and $A \succ 0$ means that A is strictly positive definite. The notation $\mathbf{0}_{n \times m}$ refers to the matrix of size $n \times m$ full of zeros. For all integers m and p such that $m \leq p$, the set $[\![m, p]\!]$ is the set of consecutive integers between m and p. We denote j the complex number such that $j^2 = -1$. The symbol $\otimes$ denotes the Kronecker product and $\star$ denotes the convolution product. For a polynomial P in the variable X, we denote $p^{(n)}$ the coefficient of the monomial X^n. For quasistationary signals x [1], we define the operator $\bar{E}[x(t)] = \lim_{N \to +\infty} \frac{1}{N} \sum_{t=1}^{N} E[x(t)]$ where E is the expectation operator.

2   Prediction error of MISO FIR system and problem considered

Consider a linear system $\mathcal{S}$ with p inputs and one output [1] [2]:

$$\mathcal{S} \ : \ y(t) = G_0(z)\mathbf{u}(t) + e_0(t) \tag{2.1}$$

where z is the forward-shift operator, $G_0(z) = \big(G_{0_1}(z), \cdots, G_{0_p}(z)\big)$ is a stable transfer function matrix, $\mathbf{u}(t) = (u_1(t) \cdots u_p(t))^T$ is the input vector, $y(t)$ is the scalar output and $e_0(t)$ is a white noise with variance σ_0^2. For the sake of simplicity, all $G_{0_i}(z)$ are FIR systems with the same order and are given by

$$G_{0_i}(z) = \beta_{i1}z^{-1} + \cdots + \beta_{im}z^{-m} \tag{2.2}$$

where all β_{ik} are the true parameters to estimate.

The system $\mathcal{S}$ in (2.1) is identified within its full-order parametrized model structure denoted

$M(\theta) = \{G(z,\theta) = (G_1(z,\theta) \cdots G_p(z,\theta)) \mid \theta \in \mathbb{R}^{pm}\}$ where θ is the parameter vector and such that $\exists \theta_0 \in \mathbb{R}^{pm}$ such that $G(z,\theta_0) = G_0(z)$ (full-order). In this case, each $G_i(z,\theta)$ has the following form:

$$G_i(z,\theta) = b_{i1} z^{-1} + \cdots + b_{im} z^{-m} \tag{2.3}$$

where all b_{ik} with $i \in [\![1,p]\!]$ and $k \in [\![1,m]\!]$ are the parameters to identify. We will denote $\theta_i = (b_{i_1} \cdots b_{i_m})^T$. Hence, the parameter vector is given by $\theta^T = \begin{pmatrix} \theta_1^T & \cdots & \theta_p^T \end{pmatrix}$.

From $M(\theta)$, we define the one-step ahead predictor:

$$\hat{y}(t,\theta) \triangleq G(z,\theta)\mathbf{u}(t) = \sum_{i=1}^{p} G_i(z,\theta_i) u_i(t) \tag{2.4}$$

$$= \sum_{i=1}^{p} \phi_{u_i}^T(t)\theta_i \tag{2.5}$$

where $\phi_{u_i}^T(t) = (u_i(t-1) \cdots u_i(t-m))$ is a regressor of the scalar input $u_i(t)$. Denote

$$\phi_{\mathbf{u}}(t) = \begin{pmatrix} \phi_{u_1}(t) \\ \vdots \\ \phi_{u_p}(t) \end{pmatrix} = \mathbf{u}(t) \otimes \begin{pmatrix} z^{-1} \\ \vdots \\ z^{-m} \end{pmatrix} \tag{2.6}$$

the input regressor derived from $M(\theta)$. With this notation, (2.5) becomes $\hat{y}(t,\theta) = \phi_{\mathbf{u}}^T(t)\theta$. Define the prediction error:

$$\epsilon(t,\theta) = y(t) - \hat{y}(t,\theta) = y(t) - \phi_{\mathbf{u}}^T(t)\theta \tag{2.7}$$

Prediction Error identification consists in computing the optimal θ, denoted $\hat{\theta}_N$, minimizing a least-squares cost-function on the prediction error by using N input-output data generated from $\mathcal{S}$:

$$\hat{\theta}_N = \arg \min_{\theta} V_N(\theta) \tag{2.8}$$

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \epsilon(t,\theta)^2 \tag{2.9}$$

If $M(\theta)$ is full-order, globally identifiable [1] [3] and the data are informative w.r.t. $M(\theta)$ [1] [2] [3], then $\hat{\theta}_N$ will be a consistent estimate of θ_0 [1]. In other words, the asymptotic identification criterion, i.e.

$$\arg \min_{\theta} \bar{E}[\epsilon(t,\theta)^2]$$

has unique solution which is equal to θ_0. The MISO FIR model structure is globally identifiable [3]. Hence, we only need to study the data informativity w.r.t. $M(\theta)$ to guarantee the consistency of $\hat{\theta}_N$.

In *open loop*, the data informativity w.r.t. the MISO FIR model structure $M(\theta)$ is equivalent to the input informativity.

Definition 2.1

The input $\mathbf{u}(t)$ *is said to be informative w.r.t. the MISO FIR structure* $M(\theta)$ *if, we have:*

$$\bar{E}\left\{[G(z,\theta') - G(z,\theta'')]\mathbf{u}(t)\right\}^2 = 0 \Rightarrow \theta' = \theta'' \tag{2.10}$$

for any θ' and θ'' in $\mathbb{R}^{pm}$ ☐

In the sequel, we will consider that each input $u_i(t)$ is generated using r *independent zero-mean* white noise processes $e_q(t)$ with $q \in [\![1, r]\!]$, each of them filtered by a rational and stable filter denoted $F_{iq}(z)$:

$$u_i(t) = \sum_{q=1}^{r} F_{iq}(z)e_q(t) \tag{2.11}$$

Each $e_q(t)$ is assumed independent of $e_0(t)$. Regroup the white noises $e_q(t)$ in a vector $\mathbf{e}(t) = \begin{pmatrix} e_1(t) & \cdots & e_r(t) \end{pmatrix}^T$ and construct the input $\mathbf{u}(t)$ such that $\mathbf{u}(t) = F(z)\mathbf{e}(t)$ where $F(z) = (F_{iq})_{(i,q) \in [\![1,p]\!] \times [\![1,r]\!]}$ is a rational stable filter matrix. We will assume that there is no zero-column in $F(z)$, i.e. each $e_q(t)$ is used at least once for the generation of $\mathbf{u}(t)$. The power spectrum matrix of $\mathbf{u}$, denoted $\Phi_{\mathbf{u}}$, is given by $\Phi_{\mathbf{u}}(\omega) = F(e^{j\omega})\Lambda_0 F^*(e^{j\omega})$ where $\bar{E}[\mathbf{e}(t)\mathbf{e}^T(t)] = \Lambda_0 \succ 0$. By using Parseval theorem, the left hand equality of (2.10) is equivalent to

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \Delta G(e^{j\omega})\Phi_{\mathbf{u}}(\omega)\Delta G^*(e^{j\omega})d\omega = 0 \tag{2.12}$$

where $\Delta G(z) = G(z, \theta') - G(z, \theta'')$. Hence, it is straightforward to see that, if the input power matrix $\Phi_{\mathbf{u}}$ is strictly positive definite at all frequencies, then $\Delta G(e^{j\omega}) = 0 \ \forall\omega$, leading to the consistency. However, such property is not necessary and not verified for all $\mathbf{u}(t) = F(z)\mathbf{e}(t)$ leading to informativity. For instance, with $p = 2$, consider

$$F(z) = \begin{pmatrix} 1 + 2z^{-1} \\ z^{-1} + 3z^{-2} \end{pmatrix} \text{ and } \Lambda_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.13}$$

The input power matrix is given for all $\omega \in [-\pi, \pi]$ by

$$\Phi_{\mathbf{u}}(\omega) = \begin{pmatrix} 2e^{-j\omega} + 2e^{j\omega} + 5 & 7e^{j\omega} + 3e^{j2\omega} + 2 \\ 7e^{-j\omega} + 3e^{-j2\omega} + 2 & 3e^{-j\omega} + 3e^{j\omega} + 10 \end{pmatrix}$$

It has a zero-determinant for all frequencies: we cannot verify the data informativity with the positive definiteness of $\Phi_{\mathbf{u}}$ in this case. However, we will show that, with this filter and when $m = 2$, we have data informativity (see **Case 4**, in Section 4).

The motivation of this paper is to find a necessary and sufficient condition to verify the data informativity for any filter matrix $F(z)$ and for any number r of white noises in $\mathbf{e}(t)$. We have this following necessary and sufficient condition in the general case:

Theorem 2.1

[7] [4] *The following propositions are equivalent:*

(a) *The input vector $\mathbf{u}(t)$ is informative w.r.t. $M(\theta)$.*

(b) *$\bar{E}[\phi_{\mathbf{u}}(t)\phi_{\mathbf{u}}^T(t)] \succ 0$.*

(c) *The signals in $\phi_{\mathbf{u}}(t)$ are linearly independent in the signal space, i.e. $\phi_{\mathbf{u}}^T(t)\alpha = 0 \ \forall t \Leftrightarrow \alpha = 0$.* ☐

Proof: • Let us first prove that $(a) \Leftrightarrow (b)$.

As we have $G(z,\theta) = \phi_{\mathbf{u}}^T(t)\theta$, the equality $\bar{E}\left\{[G(z,\theta') - G(z,\theta'')]\mathbf{u}(t)\right\}^2 = 0$ is equivalent to $\Delta\theta^T \bar{E}[\phi_u(t)\phi_u^T(t)]\Delta\theta = 0$ with $\Delta\theta = \theta' - \theta''$. Hence, the property (2.10) holds if and only if $\bar{E}[\phi_{\mathbf{u}}(t)\phi_{\mathbf{u}}^T(t)] \succ 0$.

• Let us now prove that $(b) \Leftrightarrow (c)$.

For any $\alpha \in \mathbb{R}^{pm}$, $\bar{E}[\phi_{\mathbf{u}}(t)\phi_{\mathbf{u}}^T(t)] \succ 0$ is equivalent to $\bar{E}[\phi_u(t)^T\alpha]^2 = 0 \Leftrightarrow \alpha = 0$. However the quantity $\bar{E}[\phi_u(t)^T\alpha]^2$ is the power of any signal $\phi_u(t)^T\alpha$. Therefore, as the power of any signal is equal to 0 if and only if this signal is identically equal to 0, we have $\bar{E}[\phi_u(t)^T\alpha]^2 = 0 \Leftrightarrow \phi_u(t)^T\alpha = 0 \ \forall t$. Therefore $\phi_u(t)$ is PE if and only if $\phi_u(t)^T\alpha = 0 \ \forall t$ which completes the proof. ∎

A quasistationary vector-valued signal regressor $\phi(t)$ verifying $\bar{E}[\phi(t)\phi^T(t)] \succ 0$ is called persistently exciting (PE). Applied to $\phi_{\mathbf{u}}(t)$, it is a necessary and sufficient condition for the data informativity w.r.t. MISO FIR model structure in open-loop (and so is a necessary and sufficient condition for the consistency). We are going to apply this condition in the case of filtered white noise excitation for any number of white noises in $\mathbf{e}(t)$. Hence, the problem of this paper is the following one:

Problem 2.1

Find necessary and sufficient conditions on $F(z)$ such that the regressor $\phi_{\mathbf{u}}(t)$ in (2.6) is PE for any number r of different white noise processes in $\mathbf{e}(t)$. □

3   Input regressor persistency in the case of white noise excitation

3.1   Development of a necessary and sufficient condition

In this section, we are going to develop a necessary and sufficient condition for the persistency of the input regressor $\phi_{\mathbf{u}}(t)$ when each input $u_i(t)$ is given by (2.11).

In the sequel, we will decompose $F(z)$ as follows:

$$F(z) = \frac{L(z)}{w(z)} \tag{3.1}$$

where the scalar filter $w(z)$ is the obtained denominator after putting all filters in $F(z)$ on the same denominator and $L(z)$ a $p \times r$ matrix of FIR filters. We will denote L_{iq} the (i,q)-entry of $L(z)$ and each L_{iq} will be written as follows

$$L_{iq}(z) = \sum_{h=0}^{n_{iq}} l_{iq}^{(h)} z^{-h} \tag{3.2}$$

Remark 3.1

*One can also think of writing $F(z)$ with the Infinite Impulse Response (IIR) of the rational filters, however it seems more complex to develop a necessary and sufficient condition which easily answers Problem 2.1. We will show that both ways lead to the same result in one numerical example (see **Case 1** in section 4).* □

By denoting n_{max} the maximal order among all filters $L_{iq}(z)$, we rewrite $L(z)$ as follows

$$L(z) = \sum_{h=0}^{n_{max}} L^{(h)} z^{-h} \text{ where } L^{(h)} = \begin{pmatrix} l_{11}^{(h)} & \cdots & l_{1r}^{(h)} \\ \vdots & \vdots & \vdots \\ l_{p1}^{(h)} & \cdots & l_{pr}^{(h)} \end{pmatrix} \tag{3.3}$$

By using (2.6), $\phi_{\mathbf{u}}(t)$ obeys the following relation:

$$w(z)\phi_{\mathbf{u}}(t) = \left(L(z) \otimes \begin{pmatrix} z^{-1} \\ \vdots \\ z^{-m} \end{pmatrix} \right) \mathbf{e}(t) \tag{3.4}$$

Let us rewrite the expression in (3.4). First, we will consider a permuted version $\tilde{\phi}_{\mathbf{u}}(t)$ of $\phi_{\mathbf{u}}(t)$ such that it has this following form

$$w(z)\tilde{\phi}_{\mathbf{u}}(t) = \begin{pmatrix} \mathbf{u}(t-1) \\ \vdots \\ \mathbf{u}(t-m) \end{pmatrix} = \begin{pmatrix} L(z)\mathbf{e}(t-1) \\ \vdots \\ L(z)\mathbf{e}(t-m) \end{pmatrix} \tag{3.5}$$

Remark 3.2

It is clear that if Theorem 3.1 holds for $\tilde{\phi}_{\mathbf{u}}$, it also holds for $\phi_{\mathbf{u}}$ (and conversely). $\square$

Hence, combining (3.3) and (3.5) leads to

$$w(z)\tilde{\phi}_{\mathbf{u}}(t) = \sum_{h=0}^{n_{max}} \begin{pmatrix} L^{(h)}\mathbf{e}(t-h-1) \\ \vdots \\ L^{(h)}\mathbf{e}(t-h-m) \end{pmatrix}$$

equivalent to

$$w(z)\tilde{\phi}_{\mathbf{u}}(t) = \mathbb{F} \underbrace{\begin{pmatrix} \mathbf{e}(t-1) \\ \vdots \\ \mathbf{e}(t-m-n_{max}) \end{pmatrix}}_{\mathbf{v}(t)} \tag{3.6}$$

where

$$\mathbb{F} = \begin{pmatrix} \mathbb{L}_{base} & & & \mathbf{0}_{p\times(m-1)r} \\ \mathbf{0}_{p\times r} & \mathbb{L}_{base} & & \mathbf{0}_{p\times(m-2)r} \\ \mathbf{0}_{p\times r} & \mathbf{0}_{p\times r} & \mathbb{L}_{base} & \mathbf{0}_{p\times(m-3)r} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{p\times r} & \mathbf{0}_{p\times r} & \mathbf{0}_{p\times r} & \cdots & \mathbb{L}_{base} \end{pmatrix} \tag{3.7}$$

$$\mathbb{L}_{base} = \begin{pmatrix} L^{(0)} & L^{(1)} & \cdots & L^{(n_{max})} \end{pmatrix} \tag{3.8}$$

Note that the matrix $\mathbb{F}$ *is not diagonal-block*, that is why we decided to represent it as in (3.7). We provide a numerical example to illustrate this point. Consider that $w(z) = 1$, $p = 3$, $m = 2$, $r = 2$ and:

$$L(z) = \begin{pmatrix} 1 & -z^{-3} \\ 2+2z^{-3} & 5z^{-2} \\ 0 & 2z^{-1}+3z^{-2}-4z^{-3} \end{pmatrix}$$

Hence, $n_{max} = 3$,

$$L^{(0)} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 0 \end{pmatrix} \qquad L^{(1)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 2 \end{pmatrix}$$

$$L^{(2)} = \begin{pmatrix} 0 & 0 \\ 0 & 5 \\ 0 & 3 \end{pmatrix} \qquad L^{(3)} = \begin{pmatrix} 0 & -1 \\ 2 & 0 \\ 0 & -4 \end{pmatrix}$$

and

$$L_{base} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 2 & 0 & 0 & 0 & 0 & 5 & 2 & 0 \\ 0 & 0 & 0 & 2 & 0 & 3 & 0 & 4 \end{pmatrix}$$

Therefore, $\mathbb{F}$ is given by

$$\mathbb{F} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 5 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 3 & 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 5 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 3 & 0 & 4 \end{pmatrix}$$

It is clear that it is not block-diagonal.

The matrix $\mathbb{F}$ that links $\mathbf{v}(t)$ to $\tilde{\phi}_{\mathbf{u}}(t)$ is a matrix of scalars. In this case, we have the following result:

Theorem 3.1

The regressor $\tilde{\phi}_{\mathbf{u}}(t)$ defined above in (3.5) is PE if and only if the matrix $\mathbb{F}$ in (3.6) is full-row rank.   □

Proof: Since $w(z)$ is a 1×1 stable FIR filter that applies to all elements of $\tilde{\phi}_{\mathbf{u}}(t)$, it is clear that $\tilde{\phi}_{\mathbf{u}}(t)$ is PE if and only if $w(z)\tilde{\phi}_{\mathbf{u}}(t)$ is PE. Let us now prove that $\phi_w(t) = w(z)\tilde{\phi}_{\mathbf{u}}(t)$ is PE if and only if $\mathbb{F}$ is full row-rank.

From (3.6), $\bar{E}[\phi_w(t)\phi_w^T(t)] = \mathbb{F}\bar{E}[\mathbf{v}(t)\mathbf{v}^T(t)]\mathbb{F}^T$. As $\bar{E}[\mathbf{v}(t)\mathbf{v}^T(t)] \succ 0$ (due to white noise properties), $\bar{E}[\phi_w(t)\phi_w^T(t)] \succ 0$ if and only if $\mathbb{F}$ is full-row rank.   ■

Theorem 3.1 combined with Remark 3.2 answers Problem 2.1: it is a necessary and sufficient condition to get consistency for the identification of MISO FIR model structure in open-loop for any number of white noises in $\mathbf{e}(t)$. It depends explicitly on the filters coefficient and is a similar criterion as in the multisine case [4] (rank verification of a certain matrix of scalars).

For some choices, the matrix $\mathbb{F}$ will not be full-row rank. We can deduce some possible reasons to understand why it happens. It is explained in the next paragraph.

3.2   Development of additional conditions

One has to be careful in the choice of $L(z)$ to provide the right-invertibility of $\mathbb{F}$. Indeed, the rank verification is an *a posteriori* condition.

But, first, one has to remark that, depending on the choice of $F(z)\mathbf{e}(t)$, there can be some zero-columns in $\mathbb{F}$. A zero-column column in $\mathbb{F}$ means that the corresponding generated white noise element in $\mathbf{v}(t)$ is not present in $\tilde{\phi}_{\mathbf{u}}(t)$. To illustrate this remark, let us consider back the example given in Section 3. There are two zero-columns in $\mathbb{F}$. As $\mathbb{F}$ multiplies $\mathbf{v}(t)$ and from the construction of $\mathbf{v}(t)$ in (3.6), these zero-columns mean that the correspond generated white noises $e_2(t-1)$ and $e_1(t-3)$ are not present in $\phi_{\mathbf{u}}(t)$. The generated white noises present in $\phi_{\mathbf{u}}(t)$ obtained from $e_1(t)$ are $e_1(t-1)$, $e_1(t-2)$, $e_1(t-4)$ and $e_1(t-5)$. The ones obtained from $e_2(t)$ are $e_2(t-2)$, $e_2(t-3)$, $e_2(t-4)$ and $e_2(t-5)$.

Theorem 3.2

If $\mathbb{F}$ defined in (3.7) is full-row rank, then

- *(i) the rows of $L(z)$ (or equivalently $F(z)$) must be linearly independent.*

- *(ii) the number of generated white noises in $\mathbf{v}(t)$ present in $\phi_{\mathbf{u}}(t)$ is greater than or equal to the number of parameters to be identified.* □

Proof: (*i*) From the construction of $\mathbb{F}$, when it is full-row rank, $\mathbb{L}_{base}$ must be full-row rank. With the expression of $\mathbb{L}_{base}$ in (3.7), it corresponds to the linear independence of the rows of $L(z)$ (or $F(z)$ by dividing $L(z)$ by $w(z)$).

(*ii*) To be full-row rank, $\mathbb{F}$ must have at least a number of *non-zero columns* greater than or equal to its row number. There are pm rows in $\mathbb{F}$, corresponding to the number of parameters to be identified. There are $m + n_{max}$ columns, corresponding to the number of generated white noise elements of the form $e_q(t-k)$ in $\mathbf{v}(t)$. However, some of these columns can be zero-columns. Then, the conclusion follows. ∎

The conditions derived in Theorem 3.2 are necessary *but not sufficient*. This remark will be illustrated with a numerical example (**Case 3** in section 4). In the next paragraph, we give an interpretation and a synthesis of the obtained results.

3.3   Synthesis of the developed results

In this paper, we have developed a necessary and sufficient conditions to verify the data informativity w.r.t. a MISO FIR model structure (and so the consistency) in open-loop excited by filtered white noise with any number of white noise processes: it is given in Theorem 3.1 and corresponds to the rank verification of a matrix depending on the coefficients of $L(z)$, on r and m. But, first, the filter matrix $F(z)$ must be decomposed as in (3.1). Then, Theorem 3.1 can be used on the FIR filter matrix $L(z)$. For both numerical examples in Section 3 and in (2.13) and by considering that $m = 2$, the rank of $\mathbb{F}$ is respectively equal to 4 and 6 which was the number of parameters to be identified in each case. For both cases, we have consistency.

We have also derived conditions to guide better the user in its filter choices, given in Theorem 3.2. *If these conditions are not met, the consistency will never be guaranteed.* The first one is the linear independence of the rows of $F(z)$ which can be easily avoided when we do the filter choice. The second one is linked to the number of generated white noise elements through the filtering of $\mathbf{e}(t)$ by $L(z)$: it must be at least equal to the number of parameters to be identified. This criterion is directly linked to the number of white noises and the filter orders as follows:

- when the number of white noises in $\mathbf{e}(t)$ is greater than the number of inputs, i.e. $r \geq p$, then there is at least one generated white noise element from the filtering of each $e_q(t)$[1]. In other words, it means there are at least $r \geq p$ generated white noise elements of the form $e_q(t-k)$ present in $\mathbf{u}(t)$, for each $q = 1, \cdots, r$. Hence, there are at least $rm \geq pm$ generated white noise elements present in $\phi_{\mathbf{u}}(t)$ (due to the m times regression): condition (ii) of Theorem 3.2 is always guaranteed. However, this does not necessarily mean that the consistency is reached.

- when $r < p$, one should consider the condition (ii) of Theorem 3.2 to verify that the chosen filter is enough complex to generate the right number of white noise elements present in $\phi_{\mathbf{u}}(t)$. Indeed, if the filter is not enough complex, the consistency will never be reached regardless of the filter coefficients values.

However, *these conditions are not sufficient*: it seems than we cannot avoid the rank verification of Theorem 3.1. But, in the general case, if the number of generated white noises is sufficient and the coefficient filters are chosen randomly, the consistency will be reached.

Remark 3.3

> Note that, when $r = p$ and $F(z)$ is such that $det(F(e^{j\omega})) \neq 0 \ \forall\omega$, then it is guaranteed that $\Phi_{\mathbf{u}}(\omega) > 0 \ \forall\omega$ and the consistency is also reached.

In the next section, we give other numerical examples.

4   Numerical example

For the numerical example, as in [4], we consider the MISO FIR system $\mathcal{S}$ described in (2.1) with $p = 2$ inputs, $m = 2$ parameters per G_{0_i}:

$$G_{0_1}(z) = 0.22z^{-1} - 0.63z^{-2} \text{ and } G_{0_2}(z) = z^{-1} + 0.95z^{-2}$$

i.e. the same example as in [4]. We use filtered white noise excitation

$$\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = \begin{pmatrix} F_{11}(z) & \cdots & F_{1r}(z) \\ F_{21}(z) & \cdots & F_{2r}(z) \end{pmatrix} \begin{pmatrix} e_1(t) \\ \vdots \\ e_r(t) \end{pmatrix}$$

where $r \in \mathbb{N}^*$, all $F_{iq}(z)$ are rational stable filters. We are going to propose several input choices to test the conditions developed in this paper. For each choice, we identify $G(z, \theta)$ with N input-output data and we compute the rank of $\mathbb{F}$ in (3.7). From Theorem 3.1, there is data informativity w.r.t. the MISO FIR model structure (and so consistency) if and only if $\text{rank}(\mathbb{F}) = pm = 4$.

For FIR systems, the modeling error $\hat{\theta}_N - \theta_0$ of the computed parameters is normally distributed around 0 [1]. Thus the relative error $(\hat{\theta}_N - \theta_0)/\theta_0$ is normally distributed around 0%. We consider $N = 1000$ and $\sigma_0^2 = 0.1$ for small variance of the relative error to facilitate the consistency verification. The notation $err(b_{ik})$ refers to the relative error between the computed parameter b_{ik} and the true parameter β_{ik}. The white noise $\mathbf{e}(t)$ is chosen Gaussian such that $E[\mathbf{e}(t)\mathbf{e}(t)] = I_r$ where I_r is the identity matrix.

---

[1]It is true as we assumed that there is no zero-column in $F(z)$.

4.1 Cases with $r = 1$ white noise

Case 1: Let us consider
$$F(z) = \begin{pmatrix} 1 \\ \frac{1}{1-0.5^{-1}} \end{pmatrix} = \frac{L(z)}{w(z)}$$

where $L(z) = \begin{pmatrix} 1 - 0.5^{-1} \\ 1 \end{pmatrix}$ and $w(z) = 1 - 0.5z^{-1}$.

With these filters, the different generated white noises in $\phi_{\mathbf{u}}$ are $e_1(t-1)$, $e_1(t-2)$ and $e_1(t-3)$: it is less than the number of parameters to be identified which are $pm = 4$. Hence, from Theorem 3.2, there is no consistency. This result is verified in simulation, in Table 1, with the high values of relative errors.

One can think of using the IIR of the filters without using the decomposition $L(z)/w(z)$. Indeed, in this case, the number of generated white noises will be infinite in $\mathbf{u}(t)$ and so in $\phi_{\mathbf{u}}(t)$ and we would circumvent the problem shown above. In this example, $1/(1 - 0.5z^{-1})$ can rewritten with its IIR. The impulse response of a filter of the form $1/(1 - bz^{-1})$ with $|b| < 1$ is given by $h_k = b^k$ for all $k \in \mathbb{N}$ and $h_k = 0$ for all $k < 0$. Hence, $u_2(t)$ is given by

$$u_2(t) = (h \star e_1)(t) = \sum_{k=0}^{+\infty} b^k e_1(t-k) = \sum_{k=0}^{+\infty} 0.5^k e_1(t-k) \tag{4.1}$$

The input regressor is given by

$$\tilde{\phi}_{\mathbf{u}}(t) = \begin{pmatrix} u_1(t-1) \\ u_2(t-1) \\ u_1(t-2) \\ u_2(t-2) \end{pmatrix} = \begin{pmatrix} e_1(t-1) \\ \sum_{k=1}^{+\infty} 0.5^{k-1} e_1(t-k) \\ e_1(t-2) \\ \sum_{k=2}^{+\infty} 0.5^{k-2} e_1(t-k) \end{pmatrix}$$
$$= \mathbb{F}\mathbf{v}(t)$$

where

$$\mathbb{F}^T = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0.5 & 1 & 1 \\ 0 & 0.5^2 & 0 & 0.5 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0.5^{k-1} & 0 & 0.5^{k-2} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \text{ and } \mathbf{v} = \begin{pmatrix} e_1(t-1) \\ e_1(t-2) \\ e_1(t-3) \\ \vdots \\ e_1(t-k) \\ \vdots \end{pmatrix} \tag{4.2}$$

From Theorem 3.1, we will have consistency if and only if $\mathbb{F}$ is full-row rank. Both first columns are linearly independent. But the columns after are all linearly dependent. Hence, we will not be able to find $pm = 4$ linearly independent columns in order to get $\mathbb{F}$ full-row rank: the identification will be non-consistent. We meet the same conclusion.

Case 2: We want to excite $\mathcal{S}$ with

$$F(z) = \begin{pmatrix} 1 + 0.5z^{-1} \\ \frac{-2+0.5z^{-2}}{1-0.5z^{-1}} \end{pmatrix} = \frac{L(z)}{w(z)}$$

where $L(z) = \begin{pmatrix} 1 - 0.25z^{-2} \\ -2 + 0.5z^{-2} \end{pmatrix}$ and $w(z) = 1 - 0.5z^{-1}$.

The second row of $L(z)$ can be obtained by multiplying the first one by -2. From Theorem 3.2, we conclude to the non-consistency. This explains the high values of relative errors found in simulation, presented in Table 1.

Case 3: Consider

$$F(z) = \begin{pmatrix} \frac{1-0.5z^{-1}}{1-0.5z^{-1}} \\ \frac{-4z^{-1}+2z^{-2}}{1-0.5z^{-1}} \end{pmatrix} = \frac{1}{w(z)} \begin{pmatrix} 1-0.5z^{-1} \\ -4z^{-1}+2z^{-2} \end{pmatrix}$$

Here, $\text{rank}(\mathbb{F}) = 3 \neq 4$: the identification is again non-consistent, as shown by the simulation in Table 1. However, both necessary conditions of Theorem 3.2 are verified. Indeed, we generate $e_1(t-1)$, $e_1(t-2)$, $e_1(t-3)$ and $e_1(t-4)$ and the rows of $L(z)$ are linearly independent. But, if we look closer at $L(z)$, we see that $-4z^{-1}L_{11}(z) = L_{22}(z)$, resulting in $u_1(t-1) = -0.25u_2(t-2) \, \forall t$. As shown in Theorem 2.1, the persistency is equivalent to the linear independence of the signal in $\phi_{\mathbf{u}}(t)$, which is not the case here. Hence, it is not consistent. This case is interesting as it allows us to insist on the fact that both conditions in Theorem 3.2 are only necessary and not sufficient.

Case 4: Consider the example given in (2.13). Here, $F(z) = L(z)$ as $w(z) = 1$ and both necessary conditions in Theorem 3.2 are verified. We obtain $\text{rank}(\mathbb{F}) = 4$: we have consistency, as illustrated by the simulation in Table 1.

4.2   Cases with $r > 1$ independent white noises

First, consider the case where $r = p = 2$.

Case 5: We take $w(z) = 1$ and for $L(z)$ the following filters: $L_{11}(z) = 1 + 2z^{-1}$, $L_{12}(z) = 2z^{-1} + 4z^{-2}$, $L_{21}(z) = 2 + z^{-1}$ and $L_{22}(z) = 4z^{-1} + 2z^{-2}$. Here, $\text{rank}(\mathbb{F}) = 3$: the identification is not consistent, as shown in Table 1. However, it is easy to verify that both conditions of Theorem 3.2 are guaranteed with this filter choice. This example is interesting, because we show that even for the case $r = p$, there can be some pathological cases where we meet both conditions of Theorem 3.2 but the identification is not consistent. It is to be noted that, in this case, $\det(F(e^{j\omega})) = 0 \, \forall \omega$ and so the sufficient condition of Remark 3.3 is not satisfied. So this remark is not in contradiction to this observed result.

We consider $w(z) = 1 - 0.5z^{-1}$ for the last two cases.

Case 6: Consider for $L(z)$ the following filters: $L_{11}(z) = 9$, $L_{12}(z) = 1$, $L_{21}(z) = 2$ and $L_{22}(z) = -4$. Here, $\text{rank}(\mathbb{F}) = 4$: the identification is consistent, as shown in Table 1.

We can also use more white noises processes than the number of inputs, for instance $r = 3$.

Case 7: Finally we consider, $L_{11}(z) = 9$, $L_{12}(z) = 1$, $L_{13}(z) = 4$, $L_{21}(z) = 2$, $L_{22}(z) = -4$ and $L_{23}(z) = -4$. Again, $\text{rank}(\mathbb{F}) = 4$: we have consistency, as illustrated in Table 1.

Table 1: Parameter relative errors for the 6 cases.

	$err(b_{11})$	$err(b_{12})$	$err(b_{21})$	$err(b_{22})$
Case 1	284.73%	-67.89%	-40.5%	-67.28%
Case 2	178.12%	-101.48%	-54.15%	-100.12%
Case 3	-0.012%	-55.41%	-11.45%	-0.003%
Case 4	0.29%	0.043%	-0.039%	-0.028%
Case 5	280.91%	-49.05%	-30.9%	-65.05%
Case 6	0.53%	0.084%	-0.041%	0.098%
Case 7	-0.055%	-0.056%	0.048%	-0.041%

5   Conclusion

This paper considers the problem of data informativity for MISO FIR systems in open-loop when this system is excited by filtered white noise. This problem is mainly straightforward when the input power spectrum matrix positive definiteness criterion can be used. When it is not the case, we need to develop a less restrictive condition which is the aim of this paper. The informativity is equivalent to the persistency of the regressor derived from the MISO FIR model structure. As for the multisine case [4], it is equivalent to the linear independence of the elements in the input regressor derived from this model structure. A necessary and sufficient condition depending on the number of considered white noise processes and on the filter coefficients. One result is that we can identify such systems with inputs generated with less white noises than the number of inputs. However, the necessary conditions developed in Theorem 3.2 show that, in this case, there is a needed complexity for the filter orders. A numerical example has been proposed to verify several results developed in this paper. In the future, we want to extend this study in closed-loop and also for other model structures such as Box-Jenkins, ARMAX, ARX and Ouput-Error. We want also to study if it is interesting to consider less white noise processes for experiment design.

References

[1] L. Ljung, *System Identification: Theory for the User*, 2nd ed.   Pearson Education, 1998.

[2] T. Soderstrom and P. Stoica, *System Identification.*   Prentice Hall, 1989.

[3] M. Gevers, A. S. Bazanella, X. Bombois, and L. Miskovic, "Identification and the information matrix: How to get just sufficiently rich?" *IEEE Transactions on Automatic Control*, vol. 54, no. 12, pp. 2828–2840, Dec 2009.

[4] K. Colin, X. Bombois, L. Bako, and F. Morelli, "Informativity: how to get just sufficiently rich for the Identification of MISO FIR Systems with Multisine Excitation?" 2019, To appear in ECC 2019, HAL Id : hal-02070880.

[5] P. M. V. den Hof, A. Dankers, P. S. Heuberger, and X. Bombois, "Identification of dynamic models in complex networks with prediction error methods—basic methods for consistent module estimates," *Automatica*, vol. 49, no. 10, pp. 2994 – 3006, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0005109813003592

[6] G. Cariolaro, *Unified signal theory.*   Springer Science & Business Media, 2011.

[7] M. Gevers, A. Bazanella, and L. Miskovi, "Informative data: How to get just sufficiently rich?" in *2008 47th IEEE Conference on Decision and Control.*   IEEE, 2008, pp. 3425–3430.