

Data Informativity for the Closed-Loop Identification of MISO ARX Systems

Kévin Colin* Laurent Bako** Xavier Bombois**,**

* *Division of Decision and Control Systems, KTH Royal Institute of Technology, Sweden*

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

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

Abstract: In the Prediction Error identification framework, it is crucial that the collected data are informative with respect to the chosen model structure to get a consistent estimate. In this work, we focus on the data informativity property for the identification of multi-inputs single-output ARX systems in closed-loop and we derive a necessary and sufficient condition to verify if a given multisine external excitation combined with the feedback introduced by the controller yields informative data with respect to the chosen model structure.

Keywords: System Identification; Data Informativity; Prediction Error Method; Consistency

1. INTRODUCTION

In order to get an accurate identified model within the Prediction Error framework, it is crucial to ensure the consistency of the estimate. For this purpose, the data collected on the to-be-identified system must be informative with respect to (w.r.t.) the considered model structure. In this paper, we consider the fundamental problem of data informativity for the direct closed-loop identification of MISO ARX systems.

Data informativity is ensured 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 where necessary and sufficient conditions have been derived for both the open-loop and direct closed-loop case [Ljung 1999, Gevers et al. 2008; 2009, Bazanella et al. 2012]. When considering Multiple-Inputs Multiple-Outputs (MIMO) systems, the work [Colin et al. 2020b] tackles the analysis of data informativity for the open-loop identification of MIMO systems. For the direct closed-loop identification of MIMO systems, there has been attention towards determining the complexity of the MIMO controller to guarantee that a costless identification (without external excitation) in closed-loop will be informative [Bazanella et al. 2010, Ng et al. 1977]. When considering the costless identification of Multiple-Inputs single-Outputs (MISO) ARMAX systems, the paper [Yan and Zhu 2018] gives a condition for the data informativity under some restrictions on the controller and the model structure.

When the controller is not enough complex to yield informative data in the costless framework, an external excitation must be added. In [Bazanella et al. 2010], it is said that an external excitation signal with a strictly positive definite power spectrum matrix at all the frequen-

cies always yields informative data for the direct closed-loop identification. This condition is of course only sufficient and is moreover very restrictive. As an example, a multisine excitation will never respect this condition. In [Colin et al. 2020a], we developed a condition that is less restrictive and that allows to verify whether data generated with multisine or filtered white noise excitation are informative with respect to MIMO model structures (Finite Impulse Response, ARX, ARMAX, Box-Jenkins,...) with diagonal noise matrices. Even if this condition is less conservative than the one in [Bazanella et al. 2010], we have only been able to prove the sufficiency of the proposed condition in [Colin et al. 2020a].

In this paper, we show that this condition is in fact also necessary in the case of MIMO ARX model structures with diagonal noise matrices. Since the approach [Colin et al. 2020a] pertains to a channel-by-channel verification, we will here restrict attention to the MISO ARX case for the sake of brevity. For the same reason, we will here only consider multisine excitations.

Notations. For any matrix A , A^T denotes its transpose. The notation $\mathbf{0}_{n \times m}$ refers to the matrix of size $n \times m$ full of zeros. The set of natural numbers and the set of real-valued scalars will be denoted by \mathbb{N} and \mathbb{R} respectively. For two integers n and p such that $n \leq p$, the set $[[n, p]]$ is the set of consecutive integers between n and p , i.e., the set $\{n, n+1, \dots, p-1, p\}$. We will denote the Cartesian product by \times . For discrete-time systems, z denotes the forward-shift operator. For discrete-time quasi-stationary signals $x(t)$, we define by \bar{E} the following operator [Ljung 1999]

$$\bar{E}[x(t)] = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{t=1}^N E[x(t)] \quad (1)$$

where E is the expectation operator. The degree of a polynomial $P(x)$ is denoted $\deg(P)$. When $x = z^{-1}$, we say

that ρ is the delay of the polynomial $P(z^{-1})$ when the first nonzero coefficient is linked to $z^{-\rho}$, i.e., $P(z^{-1}) = p_\rho z^{-\rho} + p_{\rho+1} z^{-(\rho+1)} + \dots + p_n z^{-n}$ with $p_\rho \neq 0$.

2. PREDICTION ERROR FRAMEWORK

2.1 True system and closed-loop configuration

Consider a MISO system \mathcal{S} with an input vector $\mathbf{u} \in \mathbb{R}^p$ and an output $y \in \mathbb{R}$, described by

$$\mathcal{S}: y(t) = \mathbf{G}_0(z)\mathbf{u}(t) + H_0(z)e(t) \quad (2)$$

where $\mathbf{G}_0(z)$ is a row vector of p stable transfer functions, $H_0(z)$ is a stable, inversely stable and monic¹ transfer function and e is a zero-mean white noise.

As shown in Figure 1, the system \mathcal{S} is put under feedback control with a stabilizing linear controller $\mathbf{K}(z)$ which is a column vector of p transfer functions. The reference signal is set to 0. For identification purpose, a quasi-stationary external excitation $\mathbf{r} \in \mathbb{R}^p$ can be added to the control effort such that the input \mathbf{u} is given by

$$\mathbf{u}(t) = -\mathbf{K}(z)y(t) + \mathbf{r}(t) \quad (3)$$

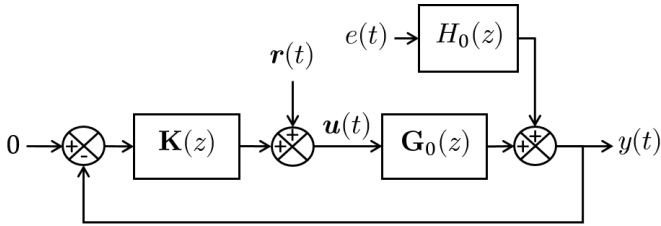


Figure 1. Closed-loop configuration.

We will assume that the excitation vector \mathbf{r} is uncorrelated with the white noise e and that there is no algebraic loop (i.e., there is at least a delay in the scalar transfer function $\mathbf{G}_0(z)\mathbf{K}(z)$). In this work, we focus on developing conditions to get a consistent estimate of $(\mathbf{G}_0(z), H_0(z))$ when considering the direct closed-loop identification approach, i.e., by using the data $\mathbf{x}(t) = (y^T(t), \mathbf{u}^T(t))^T$.

2.2 External excitation considered in this study

We will consider multisine external excitations \mathbf{r} where each entry r_k of \mathbf{r} ($k = 1, \dots, p$) is a multisine made up of sinusoids at s different frequencies ω_l ($l = 1, \dots, s$), i.e.,

$$r_k(t) = \sum_{l=1}^s \Lambda_{kl} \cos(\omega_l t + \Psi_{kl}) \quad k = 1, \dots, p \quad (4)$$

where Λ_{kl} and Ψ_{kl} are respectively the amplitude and the phase shift of the sinusoid at the frequency ω_l . Note that Λ_{kl} and Ψ_{kl} can be zero for some k ($k = 1, \dots, p$) but, for each $l = 1, \dots, s$, there exists (at least) a value of k for which $\Lambda_{kl} \neq 0$.

¹ This means that H_0 and H_0^{-1} are stable and $H_0(z = \infty) = 1$.

2.3 Model structure considered in this study

The system \mathcal{S} is identified within a full-order model structure $\mathcal{M} = \{(\mathbf{G}(z, \boldsymbol{\theta}), H(z, \boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \mathcal{D}_\theta\}$ where $\boldsymbol{\theta}$ is a parameter vector. Since we know that $\mathbf{G}_0(z)$ is stable and $H_0(z)$ is stable, inversely stable and monic, it is natural to consider the set \mathcal{D}_θ of parameters $\boldsymbol{\theta}$ that restricts the parameter vector $\boldsymbol{\theta}$ to those values for which $\mathbf{G}(z, \boldsymbol{\theta})$ is stable and $H(z, \boldsymbol{\theta})$ is monic, stable and inversely stable. The model structure is said to be full-order if $\exists \boldsymbol{\theta}_0 \in \mathcal{D}_\theta$ such that $(\mathbf{G}(z, \boldsymbol{\theta}_0), H(z, \boldsymbol{\theta}_0)) = (\mathbf{G}_0(z), H_0(z))$.

In this paper, we will consider MISO ARX model structures. In order to define this type of model structures, let us decompose the parameter vector $\boldsymbol{\theta}$ as follows $\boldsymbol{\theta} = (\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T)^T$ where $\boldsymbol{\beta}$ is a vector of dimension n_β made up by the parameters uniquely found in $G(z, \boldsymbol{\theta})$ and $\boldsymbol{\alpha}$ is a vector of dimension n_α containing the rest of the parameters. Then, in a MISO ARX model structure, we have that

$$\mathbf{G}(z, \boldsymbol{\theta}) = \mathbf{B}(z, \boldsymbol{\beta})/A(z, \boldsymbol{\alpha}) \quad (5)$$

$$H(z, \boldsymbol{\theta}) = 1/A(z, \boldsymbol{\alpha}) \quad (6)$$

where $\mathbf{B}(z, \boldsymbol{\beta})$ is a row vector of p FIR filters $B_k(z, \boldsymbol{\beta})$ and where $A(z, \boldsymbol{\alpha})$ is a monic FIR filter. Moreover, we will assume that each FIR filter $B_k(z, \boldsymbol{\beta})$ of $\mathbf{B}(z, \boldsymbol{\beta})$ is independently parametrized by a parameter vector $\boldsymbol{\beta}_k$, i.e., $B_k(z, \boldsymbol{\beta}) = B_k(z, \boldsymbol{\beta}_k)$. Finally, we will respectively denote by ρ_k and $\deg(B_k)$ the delay and the degree of $B_k(z, \boldsymbol{\beta}_k)$ ($k = 1, \dots, p$) and by $\deg(A) = n_\alpha$ the degree of $A(z, \boldsymbol{\alpha})$. Hence, the dimension n_β of $\boldsymbol{\beta}$ is equal to $n_\beta = p + \sum_{k=1}^p (\deg(B_k) - \rho_k)$.

Recall that the set \mathcal{D}_θ contains the parameter vectors $\boldsymbol{\theta}$ such that $\mathbf{G}(z, \boldsymbol{\theta})$ is stable and $H(z, \boldsymbol{\theta})$ is monic, stable and inversely stable. For MISO ARX model structures as defined above, the set \mathcal{D}_θ is made up of all $\boldsymbol{\theta} = (\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T)^T$ such that $1/A(z, \boldsymbol{\alpha})$ is stable. Therefore, the set \mathcal{D}_θ is equal to the Cartesian product of \mathbb{R}^{n_β} (since the row vector $\mathbf{B}(z, \boldsymbol{\beta})$ of FIR filters is always stable for any $\boldsymbol{\beta} \in \mathbb{R}^{n_\beta}$) and the subset of \mathbb{R}^{n_α} over which $1/A(z, \boldsymbol{\alpha})$ is stable. We will denote by \mathcal{D}_α this particular subset. Hence, we have that $\mathcal{D}_\theta = \mathcal{D}_\alpha \times \mathbb{R}^{n_\beta}$.

In the next paragraph, we define the estimator for the estimation of $\boldsymbol{\theta}_0$ and its consistency.

2.4 Estimator, consistency and data informativity

Assume that we have collected a set \mathcal{Z}^N of N input-output data defined by:

$$\mathcal{Z}^N = \left\{ \mathbf{x}(t) = (y(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, N \right\}$$

For each $(\mathbf{G}(z, \boldsymbol{\theta}), H(z, \boldsymbol{\theta})) \in \mathcal{M}$, we can define the one-step ahead predictor $\hat{y}(t, \boldsymbol{\theta})$ for the output $y(t)$ using \mathcal{Z}^N :

$$\hat{y}(t, \boldsymbol{\theta}) = W_y(z, \boldsymbol{\theta})y(t) + \mathbf{W}_u(z, \boldsymbol{\theta})\mathbf{u}(t) = \mathbf{W}(z, \boldsymbol{\theta})\mathbf{x}(t)$$

$$\mathbf{W}_u(z, \boldsymbol{\theta}) = H^{-1}(z, \boldsymbol{\theta})\mathbf{G}(z, \boldsymbol{\theta}) = \mathbf{B}(z, \boldsymbol{\beta}) \quad (7)$$

$$W_y(z, \boldsymbol{\theta}) = 1 - H^{-1}(z, \boldsymbol{\theta}) = 1 - A(z, \boldsymbol{\alpha}) \quad (8)$$

where $\mathbf{W}(z, \boldsymbol{\theta}) = (W_y(z, \boldsymbol{\theta}), \mathbf{W}_u(z, \boldsymbol{\theta}))$.

Then, we define the following estimator $\hat{\boldsymbol{\theta}}_N$ for $\boldsymbol{\theta}_0$:

$$\hat{\boldsymbol{\theta}}_N = \arg \min_{\boldsymbol{\theta} \in \mathcal{D}_\theta} \frac{1}{N} \sum_{t=1}^N \epsilon^2(t, \boldsymbol{\theta}) \quad (9)$$

where $\epsilon(t, \boldsymbol{\theta}) = y(t) - \hat{y}(t, \boldsymbol{\theta})$. In order to guarantee an accurate estimate $\hat{\boldsymbol{\theta}}_N$ for $\boldsymbol{\theta}_0$, $\hat{\boldsymbol{\theta}}_N$ must be consistent, i.e., it must converge to $\boldsymbol{\theta}_0$ with probability equal to 1 when $N \rightarrow +\infty$. To ensure the consistency, the data $\mathbf{x}(t) = (y(t), \mathbf{u}^T(t))^T$ must be informative with respect to \mathcal{M} [Ljung 1999, Bazanella et al. 2012].

Definition 1. (Data Informativity [Bazanella et al. 2010]). Consider the framework defined above with the data $\mathbf{x}(t) = (y(t), \mathbf{u}^T(t))^T$ collected on the true system \mathcal{S} in (2) operated in closed-loop with a stabilizing controller $\mathbf{K}(z)$ and a quasi-stationary external excitation \mathbf{r} (see (3)). Consider also a model structure \mathcal{M} yielding the predictor $\hat{y}(t, \boldsymbol{\theta}) = \mathbf{W}(z, \boldsymbol{\theta})\mathbf{x}(t)$. Define the set

$$\Delta_{\mathbf{W}} = \{ \Delta \mathbf{W}(z) = \mathbf{W}(z, \boldsymbol{\theta}') - \mathbf{W}(z, \boldsymbol{\theta}'') \mid (\boldsymbol{\theta}', \boldsymbol{\theta}'') \in \mathcal{D}_{\boldsymbol{\theta}}^2 \} \quad (10)$$

The data $\mathbf{x}(t)$ are said to be informative w.r.t. the model structure \mathcal{M} if, for all $\Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}}$, we have

$$\bar{E} [\|\Delta \mathbf{W}(z)\mathbf{x}(t)\|^2] = 0 \implies \Delta \mathbf{W}(z) \equiv \mathbf{0}_{1 \times (p+1)} \quad (11)$$

where the notation $\Delta \mathbf{W}(z) \equiv \mathbf{0}_{1 \times (p+1)}$ means that $\Delta \mathbf{W}(e^{j\omega}) = \mathbf{0}_{1 \times (p+1)}$ at almost all $\omega \in]-\pi, \pi]$ and where the operator \bar{E} is defined in (1). \square

Since MISO ARX model structures are globally identifiable at any $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\theta}}$ and that there is at least a delay in $\mathbf{G}_0(z)\mathbf{K}(z)$, then the data informativity implies the consistency of $\hat{\boldsymbol{\theta}}_N$ [Ljung 1999, Bazanella et al. 2012]. Hence, in this paper, our main objective is to develop a necessary and sufficient condition to *verify* whether closed-loop data generated with a given excitation \mathbf{r} of the type (4) are informative (and thus yield a consistent estimate (9)). For this purpose, we will show that the condition that we have developed in [Colin et al. 2020a] and of which we have proven the sufficiency for data informativity in the same paper is also a necessary condition for the data informativity in the case considered in this paper, i.e., the case of MISO ARX structures. In addition, we will also derive new conditions that must satisfy the controller $\mathbf{K}(z)$ and the external excitation \mathbf{r} in order to ensure the data informativity.

3. PREVIOUS RESULT ON THE DATA INFORMATIVITY

3.1 Preliminary notations

In order to recall the data informativity condition developed in [Colin et al. 2020a], let us introduce some notations. For the controller $\mathbf{K}(z)$, we will put all its entries on the least common factor of the denominators denoted $Y(z)$ such that it can be rewritten as follows

$$\mathbf{K}(z) = \mathbf{X}(z)/Y(z) \quad (12)$$

where $\mathbf{X}(z)$ is a column vector of p FIR filters $X_k(z)$ ($k = 1, \dots, p$) and we will denote by $\deg(X_k)$ and ρ_{X_k} the degree and delay of $X_k(z)$ respectively. For the monic polynomial $Y(z)$, we will denote by $\deg(Y)$ its degree. For the MISO ARX model structures, we introduce the following notations

$$B_k(z, \boldsymbol{\beta}_k) = \boldsymbol{\beta}_k^T \mathbf{Z}_{B_k}(z) \quad (13)$$

$$A(z, \boldsymbol{\alpha}) = 1 + \boldsymbol{\alpha}^T \mathbf{Z}_A(z) \quad (14)$$

where $\mathbf{Z}_{B_k}(z) = (z^{-\rho_k}, z^{-(\rho_k+1)}, \dots, z^{-\deg(B_k)})^T$ and $\mathbf{Z}_A(z) = (z^{-1}, z^{-2}, \dots, z^{-n_\alpha})^T$.

Consider now the following vector of polynomials

$$\boldsymbol{\Gamma}(z) = \begin{pmatrix} Y(z)\mathbf{Z}_A(z) \\ X_1(z)\mathbf{Z}_{B_1}(z) \\ \vdots \\ X_p(z)\mathbf{Z}_{B_p}(z) \end{pmatrix}$$

Let us denote by γ_{min} and γ_{max} the minimum of the delays and the maximum of the degrees of the entries of $\boldsymbol{\Gamma}(z)$. From the expression of $\boldsymbol{\Gamma}(z)$, the scalars γ_{min} and γ_{max} are given by

$$\gamma_{min} = \min(1, \rho_1 + \rho_{X_1}, \rho_2 + \rho_{X_2}, \dots, \rho_p + \rho_{X_p}) \quad (15)$$

$$\gamma_{max} = \max(\deg(A) + \deg(Y), \deg(B_1) + \deg(X_1), \deg(B_2) + \deg(X_2), \dots, \deg(B_p) + \deg(X_p)) \quad (16)$$

Therefore, we can rewrite $\boldsymbol{\Gamma}(z)$ as follows

$$\boldsymbol{\Gamma}(z) = \begin{pmatrix} \mathcal{A} \\ \mathcal{B} \end{pmatrix} \begin{pmatrix} z^{-\gamma_{min}} \\ z^{-(\gamma_{min}+1)} \\ \vdots \\ z^{-\gamma_{max}} \end{pmatrix} \quad (17)$$

where the entries of the matrices \mathcal{A} and \mathcal{B} depend on the entries of $Y(z)$ and $\mathbf{X}(z)$ respectively. The matrix \mathcal{A} is of dimension $n_\alpha \times (\gamma_{max} - \gamma_{min} + 1)$ and the matrix \mathcal{B} is of dimension $n_\beta \times (\gamma_{max} - \gamma_{min} + 1)$. An example of the construction of $\boldsymbol{\Gamma}(z)$, \mathcal{A} and \mathcal{B} will be given in Section 6.

When \mathbf{r} is a multisine of the type (4), we define \mathcal{C} as the column block matrix $\mathcal{C} = (\mathcal{C}_1^T, \dots, \mathcal{C}_p^T)^T$ where

$$\mathcal{C}_k^T = \begin{pmatrix} \bar{\Lambda}_{k1} e^{-j\rho_k\omega_1} & \dots & \bar{\Lambda}_{k1} e^{-j\deg(B_k)\omega_1} \\ \bar{\Lambda}_{k1}^* e^{j\rho_k\omega_1} & \dots & \bar{\Lambda}_{k1}^* e^{j\deg(B_k)\omega_1} \\ \vdots & \dots & \vdots \\ \bar{\Lambda}_{ks} e^{-j\rho_k\omega_s} & \dots & \bar{\Lambda}_{ks} e^{-j\deg(B_k)\omega_s} \\ \bar{\Lambda}_{ks}^* e^{j\rho_k\omega_s} & \dots & \bar{\Lambda}_{ks}^* e^{j\deg(B_k)\omega_s} \end{pmatrix} \quad (k = 1, \dots, p)$$

with $\bar{\Lambda}_{kl} = \Lambda_{kl} e^{j\Psi_{kl}}$ ($k = 1, \dots, p$, $l = 1, \dots, s$). The matrix \mathcal{C} has a dimension of $n_\beta \times 2s$.

3.2 Result for data informativity

Theorem 1. ([Colin et al. 2020a]). Consider Definition 1 and a MISO ARX model structure \mathcal{M} defined in Section 2.3. Assume that the true system \mathcal{S} is operated in closed-loop with a stabilizing linear controller $\mathbf{K}(z)$ with a nonzero external excitation \mathbf{r} of the type (4) and that we collect data $\mathbf{x}(t)$ on \mathcal{S} . Let us also define the sets

$$\Delta_{\boldsymbol{\beta}} = \{ \delta \boldsymbol{\beta} = \boldsymbol{\beta}' - \boldsymbol{\beta}'' \mid (\boldsymbol{\beta}', \boldsymbol{\beta}'') \in \mathbb{R}^{n_\beta} \times \mathbb{R}^{n_\beta} \} \quad (18)$$

$$\Delta_{\boldsymbol{\alpha}} = \{ \delta \boldsymbol{\alpha} = \boldsymbol{\alpha}' - \boldsymbol{\alpha}'' \mid (\boldsymbol{\alpha}', \boldsymbol{\alpha}'') \in \mathcal{D}_\alpha \times \mathcal{D}_\alpha \} \quad (19)$$

where \mathcal{D}_α is the set of all parameter vectors $\boldsymbol{\alpha}$ such that $1/A(z, \boldsymbol{\alpha})$ is stable. The data $\mathbf{x}(t)$ are informative w.r.t. \mathcal{M} if, and only if, at least one of the properties (20) and (21) holds for all $(\delta \boldsymbol{\alpha}^T \ \delta \boldsymbol{\beta}^T)^T \in \Delta_{\boldsymbol{\alpha}} \times \Delta_{\boldsymbol{\beta}}$

$$(a) \quad (\delta \boldsymbol{\alpha}^T, \delta \boldsymbol{\beta}^T) \begin{pmatrix} \mathcal{A} & \mathbf{0} \\ \mathcal{B} & \mathcal{C} \end{pmatrix} = \mathbf{0} \implies (\delta \boldsymbol{\alpha}^T \ \delta \boldsymbol{\beta}^T) = \mathbf{0}_{1 \times n_\theta} \quad (20)$$

$$(b) \quad (\delta \boldsymbol{\alpha}^T, \delta \boldsymbol{\beta}^T) \begin{pmatrix} \mathcal{A} \\ \mathcal{B} \end{pmatrix} = \mathbf{0} \implies (\delta \boldsymbol{\alpha}^T \ \delta \boldsymbol{\beta}^T) = \mathbf{0}_{1 \times n_\theta} \quad (21)$$

where \mathcal{A} , \mathcal{B} and \mathcal{C} are the matrices defined in Section 3.1. When $\mathbf{r} = \mathbf{0}$ (costless identification), the corresponding

data collected on \mathcal{S} are informative if and only if (21) holds for all $(\delta\alpha^T \ \delta\beta^T)^T \in \Delta_\alpha \times \Delta_\beta$. \square

Let us introduce the following matrices

$$\mathcal{P}^{(a)} = \begin{pmatrix} \mathcal{A} & \mathbf{0} \\ \mathcal{B} & \mathcal{C} \end{pmatrix} \quad \text{and} \quad \mathcal{P}^{(b)} = \begin{pmatrix} \mathcal{A} \\ \mathcal{B} \end{pmatrix} \quad (22)$$

which are the block-matrices involved in Theorem 1. It is clear that a sufficient condition to guarantee (20) (resp. (21)) for all $(\delta\alpha^T \ \delta\beta^T)^T \in \Delta_\alpha \times \Delta_\beta$ is that $\mathcal{P}^{(a)}$ (resp. $\mathcal{P}^{(b)}$) is full row rank.

This rank condition on $\mathcal{P}^{(a)}$ (resp. $\mathcal{P}^{(b)}$) would be a necessary and sufficient condition for (20) (resp. (21)) to hold for all $(\delta\alpha^T \ \delta\beta^T)^T \in \Delta_\alpha \times \Delta_\beta$ if the set $\Delta_\alpha \times \Delta_\beta$ contains an Euclidean ball of \mathbb{R}^{n_θ} centered at $\mathbf{0}_{n_\theta \times 1}$. Indeed, the property (20) (resp. (21)) holds for all $(\delta\alpha^T \ \delta\beta^T)^T \in \Delta_\alpha \times \Delta_\beta$ if and only if the intersection between the left nullspace of $\mathcal{P}^{(a)}$ (resp. $\mathcal{P}^{(b)}$) and the set $\Delta_\alpha \times \Delta_\beta$ is equal to the singleton $\{\mathbf{0}_{n_\theta \times 1}\}$. Moreover, the left nullspace of $\mathcal{P}^{(a)}$ (resp. $\mathcal{P}^{(b)}$) reduces to $\{\mathbf{0}_{n_\theta \times 1}\}$ when $\mathcal{P}^{(a)}$ (resp. $\mathcal{P}^{(b)}$) is full row rank and it is an hyperplane which contains $\mathbf{0}_{n_\theta \times 1}$ when $\mathcal{P}^{(a)}$ (resp. $\mathcal{P}^{(b)}$) is not full row rank. Consequently, when $\Delta_\alpha \times \Delta_\beta$ contains an Euclidean ball centered at $\mathbf{0}_{n_\theta \times 1}$, a necessary and sufficient condition for the intersection between the left nullspace of $\mathcal{P}^{(a)}$ (resp. $\mathcal{P}^{(b)}$) and $\Delta_\alpha \times \Delta_\beta$ to reduce to the singleton $\{\mathbf{0}_{n_\theta \times 1}\}$ is that $\mathcal{P}^{(a)}$ (resp. $\mathcal{P}^{(b)}$) is full row rank.

In this paper, we will show that $\Delta_\alpha \times \Delta_\beta$ contains such an Euclidean ball in the case of MISO ARX model structures. Consequently, for these model structures, a necessary and sufficient condition for data informativity is that at least one of the two matrices $\mathcal{P}^{(a)}$ and $\mathcal{P}^{(b)}$ is full row rank.

Because of the construction of Δ_β given in (18), we have that $\Delta_\beta = \mathbb{R}^{n_\beta}$. Moreover, the sets Δ_β and Δ_α are defined independently. Consequently, the set $\Delta_\alpha \times \Delta_\beta$ contains an Euclidean ball of \mathbb{R}^{n_θ} centered at $\mathbf{0}_{n_\theta \times 1}$ with a nonzero radius if and only if the set Δ_α contains an Euclidean ball of \mathbb{R}^{n_α} centered at $\mathbf{0}_{n_\alpha \times 1}$ with a nonzero radius. Because of the construction of the set Δ_α defined in (19) and of the fact that $\mathbf{0}_{n_\alpha \times 1} \in \mathcal{D}_\alpha$, we have that $\mathcal{D}_\alpha \subset \Delta_\alpha$. Consequently, if the set \mathcal{D}_α contains an Euclidean ball of \mathbb{R}^{n_α} centered at $\mathbf{0}_{n_\alpha \times 1}$ with a nonzero radius, so does Δ_α .

To summarize, if \mathcal{D}_α contains an Euclidean ball of \mathbb{R}^{n_α} centered at $\mathbf{0}_{n_\alpha \times 1}$ with a nonzero radius, so does Δ_α which in turn implies that $\Delta_\alpha \times \Delta_\beta$ contains an Euclidean ball of \mathbb{R}^{n_θ} centered at $\mathbf{0}_{n_\theta \times 1}$ with a nonzero radius. Let us therefore perform a geometrical study of the set \mathcal{D}_α .

Remark 1. For MISO ARX model structures, we can choose \mathcal{D}_θ equal to \mathbb{R}^{n_θ} since the predictor filter matrix $\mathbf{W}(z, \theta)$ is always stable (matrix of FIR filters). For that \mathcal{D}_θ , the rank condition on $\mathcal{P}^{(a)}$ and $\mathcal{P}^{(b)}$ discussed above is a necessary and sufficient condition for the data informativity. However, in this paper, we choose \mathcal{D}_θ corresponding to the priors that we have on the system, that is $\mathbf{G}_0(z)$ is stable and $H_0(z)$ is stable, inversely stable and monic. For this choice of \mathcal{D}_θ , the geometrical study of \mathcal{D}_α is thus necessary.

4. GEOMETRICAL STUDY OF THE SET \mathcal{D}_α

The set \mathcal{D}_α contains the parameter vectors $\alpha \in \mathbb{R}^{n_\alpha}$ such that $1/A(z, \alpha)$ is stable. Let us write this transfer function as follows

$$\frac{1}{A(z, \alpha)} = \frac{z^{n_\alpha}}{z^{n_\alpha} + \alpha_1 z^{n_\alpha-1} + \alpha_2 z^{n_\alpha-2} + \dots + \alpha_{n_\alpha}}$$

where α_l ($l = 1, \dots, n_\alpha$) are the coefficients contained in the parameter vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{n_\alpha})^T$. In the discrete-time domain, $1/A(z, \alpha)$ is stable if the n_α roots $z_1, z_2, \dots, z_{n_\alpha}$ of its denominator $z^{n_\alpha} + \alpha_1 z^{n_\alpha-1} + \alpha_2 z^{n_\alpha-2} + \dots + \alpha_{n_\alpha}$ have an absolute value strictly less than 1. The following lemma ensures that each root of a polynomial has a modulus smaller than 1 if the polynomial coefficients are taken to be sufficiently small in magnitude

Lemma 1. Consider a polynomial $P(x)$ of degree m of the form

$$P(x) = x^m + p_1 x^{m-1} + p_2 x^{m-2} + \dots + p_m$$

such that $|p_h| < 1/m \ \forall h \in \llbracket 1, m \rrbracket$. Then, the m roots x_1, x_2, \dots, x_m of $P(x)$ are such that $|x_h| < 1 \ \forall h \in \llbracket 1, m \rrbracket$. \square

Proof. See Appendix A. \blacksquare

As shown below, we can prove the main result of this paper based on Lemma 1.

Lemma 2. For any $n_\alpha \in \mathbb{N} \setminus \{0\}$ (i.e., for any nonzero degree of $A(z, \alpha)$), the set \mathcal{D}_α of parameters α such that $1/A(z, \alpha)$ is stable contains an Euclidean ball centered at $\mathbf{0}_{n_\alpha \times 1}$ with a nonzero radius. \square

Proof. By applying Lemma 1 to our case with $x = z$, $P = z^{n_\alpha} + \alpha_1 z^{n_\alpha-1} + \alpha_2 z^{n_\alpha-2} + \dots + \alpha_{n_\alpha}$, $m = n_\alpha$ and $p_h = \alpha_h$, we can conclude that all the transfer functions $1/A(z, \alpha)$ with a parameter vector α such that $|\alpha_h| < 1/n_\alpha \ \forall h \in \llbracket 1, n_\alpha \rrbracket$ are stable since the roots of the denominator have an absolute value strictly less than 1. Hence, all parameter vectors α satisfying this property belong to \mathcal{D}_α . Moreover, the set of parameter vectors α satisfying $|\alpha_h| < 1/n_\alpha \ \forall h \in \llbracket 1, n_\alpha \rrbracket$ is the interior of an hypercube centered at $\mathbf{0}_{n_\alpha \times 1}$ and where each line segment has a length of $2/n_\alpha$. The interior of this hypercube (which is a subset of \mathcal{D}_α) contains entirely, e.g., the Euclidean ball centered at $\mathbf{0}_{n_\alpha \times 1}$ with a radius equal to $0.9/n_\alpha$. Therefore, there exists an Euclidean ball centered at $\mathbf{0}_{n_\alpha \times 1}$ with a nonzero radius (e.g., a radius equal to $0.9/n_\alpha$) subset of \mathcal{D}_α , which concludes the proof. \blacksquare

Consequently, we obtain this important result

Theorem 2. Consider Theorem 1 and a MISO ARX model structure \mathcal{M} defined in Section 2.3. Assume that the true system \mathcal{S} is operated in closed-loop with a controller $\mathbf{K}(z)$ with a multisine excitation \mathbf{r} of the type (4) and that we collect data $\mathbf{x}(t)$ on \mathcal{S} . Then, the data $\mathbf{x}(t)$ are informative w.r.t. \mathcal{M} if and only if at least one of the two matrices $\mathcal{P}^{(a)}$ and $\mathcal{P}^{(b)}$ (see (22)) is full row rank. When $\mathbf{r} = \mathbf{0}$, we have data informativity if and only if the matrix $\mathcal{P}^{(b)}$ is full row rank. \square

In the next section, we give some hints/results on the combined complexity that the controller $\mathbf{K}(z)$ and the multisine external excitation \mathbf{r} should have in order to guarantee the data informativity.

5. ADDITIONAL CONDITIONS FOR THE DATA INFORMATIVITY

As aforementioned, the necessary and sufficient condition for the data informativity presented in this paper relies on the fact that a particular matrix is full row rank (Theorem 2). For a matrix to be full row rank, it is necessary that its number of columns is greater than or equal to its number of rows. With this last statement, we will see in the next theorem that there is a necessary combined complexity that the controller $\mathbf{K}(z)$ and the multisine external excitation \mathbf{r} should have in order to guarantee the data informativity.

Theorem 3. *Consider Theorem 1 and a MISO ARX model structure \mathcal{M} as defined in Section 2.3. Assume that the true system \mathcal{S} is operated in closed-loop with a stabilizing linear controller $\mathbf{K}(z)$ with a nonzero multisine external excitation \mathbf{r} of the type (4). Then, the data $\mathbf{x}(t)$ collected on \mathcal{S} are informative w.r.t. \mathcal{M} only if at least one of the following properties is satisfied*

- the following condition is guaranteed

$$\gamma_{max} - \gamma_{min} + 1 \geq \deg(A) + p + \sum_{k=1}^p (\deg(B_k) - \rho_k) \quad (23)$$

where γ_{min} and γ_{max} are defined in (15) and (16).

- the total number s of sinusoids in \mathbf{r} guarantees that

$$s \geq \frac{\deg(A) + p + \sum_{k=1}^p (\deg(B_k) - \rho_k) - \text{rank}(\mathcal{P}^{(b)})}{2} \quad (24)$$

When $\mathbf{r} = \mathbf{0}$, the condition (23) must be guaranteed in order to ensure the data informativity. \square

Proof. When $\mathbf{r} = \mathbf{0}$, the data $\mathbf{x}(t)$ are informative w.r.t. \mathcal{M} if and only if $\mathcal{P}^{(b)}$ defined in (22) is full row rank. For this matrix to be full row rank, its number of columns (equal to $\gamma_{max} - \gamma_{min} + 1$) must be greater than or equal to its number of rows (equal to $n_\alpha + n_\beta = \deg(A) + p + \sum_{k=1}^p (\deg(B_k) - \rho_k)$), which implies (23).

When $\mathbf{r} \neq \mathbf{0}$, then the data $\mathbf{x}(t)$ are informative w.r.t. \mathcal{M} if and only if at least one of the two matrices $\mathcal{P}^{(a)}$ and $\mathcal{P}^{(b)}$ (see (22)) is full row rank. For the matrix $\mathcal{P}^{(b)}$ to be full row rank, (23) must hold as proved earlier. If $\mathcal{P}^{(b)}$ is not full row rank, then the multisine excitation \mathbf{r} must be designed such that $\mathcal{P}^{(a)}$ is full row rank. We observe that $\mathcal{P}^{(b)}$ and $\mathcal{P}^{(a)}$ have the same number of rows (equal to $n_\alpha + n_\beta = \deg(A) + p + \sum_{k=1}^p (\deg(B_k) - \rho_k)$), but the number of columns of $\mathcal{P}^{(a)}$ is larger than the one of $\mathcal{P}^{(b)}$ (due to the matrix \mathcal{C} linked to the external excitation). Consequently, when $\mathcal{P}^{(b)}$ is not full row rank, the addition of the $2s$ extra columns from \mathcal{C} must compensate the rank deficiency of $\mathcal{P}^{(b)}$ in order to make $\mathcal{P}^{(a)}$ full row rank. This gives us the necessary condition in (24). \blacksquare

With Theorem 3, we obtain a similar result for the data informativity as in the SISO case [Gevers et al. 2008]. Indeed, when the controller is not enough complex to provide the data informativity alone ($\mathcal{P}^{(b)}$ not full row rank), the multisine external excitation \mathbf{r} must be enough rich (i.e., must contain enough sinusoids) in order to compensate the lack of complexity of the controller

$\mathbf{K}(z)$. Nonetheless, contrary to the SISO case, the richness conditions (23) and (24) are only necessary and not sufficient. Indeed, due to some unlucky choices of controller and multisine external excitation, the inequalities (23) and (24) may be satisfied but the matrices $\mathcal{P}^{(a)}$ and $\mathcal{P}^{(b)}$ may not be full row rank. Consequently, it is always important to verify the rank after having checked that one of both inequalities (24)-(23) is satisfied. In the next section, we provide an example of an unlucky choice of controller $\mathbf{K}(z)$ to illustrate this important point.

6. NUMERICAL EXAMPLE

6.1 True system to be identified

Consider the following MISO ARX system \mathcal{S} with $p = 2$ inputs by

$$y(t) = \underbrace{\begin{pmatrix} 0.48z^{-2} & z^{-1} \\ 1 + 0.25z^{-1} & 1 + 0.25z^{-1} \end{pmatrix}}_{\mathbf{G}_0(z)} \mathbf{u}(t) + \underbrace{\frac{1}{1 + 0.25z^{-1}}}_{H_0(z)} e(t)$$

The system \mathcal{S} is operated under feedback control with a stabilizing controller $\mathbf{K}(z)$ given by

$$\mathbf{K}(z) = \underbrace{\begin{pmatrix} 1 - 0.5z^{-1} \\ 0.0625 - 0.725z^{-1} + 0.3625z^{-2} \end{pmatrix}}_{\mathbf{X}(z)} \quad (25)$$

In this case, the common factor of the denominators is $Y(z) = 1$. We will identify \mathcal{S} within a MISO ARX model structure \mathcal{M} such that $\rho_1 = 2$, $\rho_2 = 1$, $\deg(B_1) = 2$, $\deg(B_2) = 1$ and $\deg(A) = 1$. The model structure is then full-order and the true parameter vector $\boldsymbol{\theta}_0$ is equal to $\boldsymbol{\theta}_0 = (0.25, 0.48, 1)^T$.

6.2 Costless identification ($\mathbf{r} = \mathbf{0}$)

Let us study if we can get informative data with a costless experiment, i.e., without external excitation ($\mathbf{r} = \mathbf{0}$). For this purpose, we have to construct and verify the rank of $\mathcal{P}^{(b)}$. To do so, let us construct the polynomial column vector $\boldsymbol{\Gamma}(z)$ given in (17). For the chosen model structure \mathcal{M} , we have that $\mathbf{Z}_A(z) = z^{-1}$, $\mathbf{Z}_{B_1}(z) = z^{-2}$ and $\mathbf{Z}_{B_2}(z) = z^{-1}$. Therefore,

$$\boldsymbol{\Gamma}(z) = \begin{pmatrix} \frac{z^{-1}}{z^{-2} - 0.5z^{-3}} \\ 0.0625z^{-1} - 0.725z^{-2} + 0.3625z^{-3} \end{pmatrix}$$

Hence, $\gamma_{min} = 1$, $\gamma_{max} = 3$ and so $\mathcal{P}^{(b)}$ is given by

$$\mathcal{P}^{(b)} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -0.5 \\ 0.0625 & -0.725 & 0.3625 \end{pmatrix}$$

The matrix $\mathcal{P}^{(b)}$ has 3 columns for 3 rows. Therefore, the controller is enough complex to satisfy the condition (23). However, we do not have data informativity since $\text{rank}(\mathcal{P}^{(b)}) = 2$ and so $\mathcal{P}^{(b)}$ is not full row rank. Here, we illustrate that the complexity condition (23) on the controller is only necessary and not sufficient for the data informativity, while for the SISO case it was necessary and sufficient [Gevers et al. 2008]. To confirm that we indeed do not have data informativity in this case, let us observe that the controller $\mathbf{K}(z)$ makes that $\hat{y}(t, \boldsymbol{\theta}') = \hat{y}(t, \boldsymbol{\theta}'') \forall t$ for $\boldsymbol{\theta}' = (0.3, -0.1, 0.2)^T$ and $\boldsymbol{\theta}'' = \boldsymbol{\theta}_0$.

6.3 Identification with an external excitation ($\mathbf{r} \neq \mathbf{0}$)

Since the costless identification cannot yield informative data, we need to add a multisine external excitation \mathbf{r} . In order to make $\mathcal{P}^{(a)}$ full row rank, we have to choose the number s of sinusoids for the excitation such that it satisfies (24). In this case, $s \geq 1$. Let us therefore choose a signal $\mathbf{r}(t) = (r_1(t), r_2(t))^T$ with $s = 1$ sinusoid at $\omega_1 = 0.1\text{rad/s}$ and with the following phasors $\bar{\Lambda}_{kl}$ ($k = 1, 2, l = 1$)

$$\bar{\Lambda}_{11} = 0 \quad \bar{\Lambda}_{21} = -1.7e^{j0.5}$$

i.e., $r_1(t) = 0$ and $r_2(t) = -1.7 \cos(0.1t + 0.5) \forall t$. Here, the matrix \mathcal{C} is given by

$$\mathcal{C} = \begin{pmatrix} \bar{\Lambda}_{11}e^{-2j\omega_1} & \bar{\Lambda}_{11}^*e^{2j\omega_1} \\ \bar{\Lambda}_{21}e^{-j\omega_1} & \bar{\Lambda}_{21}^*e^{j\omega_1} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -1.7e^{j0.4} & -1.7e^{-j0.4} \end{pmatrix}$$

Let us calculate the rank of the corresponding matrix $\mathcal{P}^{(a)}$. With this excitation, the matrix $\mathcal{P}^{(a)}$ has a rank of 3: it is full row rank. Therefore, from Theorem 2, the data $\mathbf{x}(t)$ generated with this excitation \mathbf{r} are thus informative w.r.t. \mathcal{M} . It is interesting to observe that we can yield informative data by only exciting one signal in \mathbf{r} , as was observed in [Mišković et al. 2008].

6.4 Monte-Carlo simulations

In order to verify that we have the data informativity with the multisine external excitation, we have applied it to the true system in 1000 identification experiments (with different realizations of the white noise e , assumed to be zero-mean Gaussian with a variance of 1) and we have identified $\hat{\boldsymbol{\theta}}_N$ (see (9)) for each experiment with $N = 10000$ input-output data. For each input vector, we have computed the mean of these 1000 estimates. The mean of these 1000 estimates is $(0.2496, 0.4802, 0.9999)^T$ which is very close to $\boldsymbol{\theta}_0 = (0.25, 0.48, 1)^T$, suggesting the consistency and so the data informativity w.r.t. \mathcal{M} .

7. CONCLUSION

In this paper, we have studied the data informativity for the direct closed-loop identification of MISO ARX systems with multisine external excitation. We have seen that the condition in [Colin et al. 2020a] is a necessary and sufficient one in this case. Based on this result, we have developed additional conditions to help the user in her/his choice of the external excitation to ensure the data informativity.

REFERENCES

- Bazanella, A.S., Bombois, X., and Gevers, M. (2012). Necessary and sufficient conditions for uniqueness of the minimum in Prediction Error Identification. *Automatica*, 48(8), 1621 – 1630.
- Bazanella, A.S., Gevers, M., and Mišković, L. (2010). Closed-loop identification of MIMO systems: a new look at identifiability and experiment design. *European Journal of Control*, 16(3), 228–239.
- Colin, K., Bombois, X., Bako, L., and Morelli, F. (2020a). Closed-loop identification of MIMO systems in the Prediction Error framework: Data informativity analysis. *Automatica*, 121.
- Colin, K., Bombois, X., Bako, L., and Morelli, F. (2020b). Data informativity for the open-loop identification of MIMO systems in the prediction error framework. *Automatica*, 117.
- Gevers, M., Bazanella, A., and Mišković, L. (2008). Informative data: How to get just sufficiently rich? In *2008 47th IEEE Conference on Decision and Control.*, 3425–3430. IEEE.
- Gevers, M., Bazanella, A.S., Bombois, X., and Mišković, L. (2009). Identification and the Information Matrix: How to Get Just Sufficiently Rich? *IEEE Transactions on Automatic Control*, 54(12), 2828–2840.
- Ljung, L. (1999). *System identification: Theory for the user*. Prentice Hall information and system sciences series. Prentice Hall PTR, Upper Saddle River (NJ), second edition.
- Mišković, L., Karimi, A., Bonvin, D., and Gevers, M. (2008). Closed-loop identification of multivariable systems: With or without excitation of all references? *Automatica*, 44(8), 2048–2056.
- Ng, T.S., Goodwin, G.C., and Anderson, B.D. (1977). Identifiability of MIMO linear dynamic systems operating in closed loop. *Automatica*, 13(5), 477–485.
- Yan, W. and Zhu, Y. (2018). Informative Conditions for Identification of MISO ARMAX Model in Closed-loop Systems. In *18th IFAC Symposium on System Identification SYSID 2018*, volume 51, 455 – 460.

Appendix A. PROOF OF LEMMA 1

To prove Lemma 1, we will need the following result

Lemma 3. Let $m \in \mathbb{N} \setminus \{0\}$. Define the function f_m by

$$f_m : \mathbb{R}^+ \rightarrow \mathbb{R} \\ x \rightarrow f_m(x) = x^m - \frac{1}{m} \sum_{i=1}^m x^{m-i} \quad (\text{A.1})$$

Then, $f_m(x) < 0$ if and only if $0 \leq x < 1$. \square

Proof. Let $m \in \mathbb{N} \setminus \{0\}$. First, we have that $f_m(1) = 0$ and $f_m(0) = -1/m < 0$. Secondly, when $0 < x < 1$, we have that $x^m < x^{m-i} \forall i \in \llbracket 1, m \rrbracket$. By summing these m inequalities, we obtain $mx^m < \sum_{i=1}^m x^{m-i}$ and so $f_m(x) < 0$. Finally, when $x > 1$, we have that $x^m > x^{m-i} \forall i \in \llbracket 1, m \rrbracket$. By summing these m inequalities, we obtain $mx^m > \sum_{i=1}^m x^{m-i}$ and so $f_m(x) > 0$. We conclude that $f_m(x) < 0$ if and only if $0 \leq x < 1$. \blacksquare

Let us now prove Lemma 1. Consider the following polynomial $Q(x) = P(x) - x^m$ where $P(x)$ is the polynomial defined in the lemma statement. By recalling that the roots of $P(x)$ are denoted by x_1, x_2, \dots, x_m , then we have that, for all $h \in \llbracket 1, m \rrbracket$,

$$|Q(x = x_h)| = \left| \underbrace{P(x = x_h)}_0 - x_h^m \right| = \left| \sum_{i=1}^m p_i x_h^{m-i} \right| \leq \sum_{i=1}^m |p_i| |x_h|^{m-i}$$

Since $|p_i| < 1/m$ for all $i \in \llbracket 1, m \rrbracket$, then we have that, for all $h \in \llbracket 1, m \rrbracket$, $|x_h|^m < 1/m \sum_{i=1}^m |x_h|^{m-i}$ which is equivalent to $f_m(|x_h|) < 0$ where f_m is the function defined in Lemma 3. By applying Lemma 3, we conclude that $|x_h| < 1$. Since the latter holds for all $h \in \llbracket 1, m \rrbracket$, then the proof follows. \blacksquare