Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Data Informativity for the Identification of particular Parallel Hammerstein Systems

K. Colin [*] **X. Bombois** [*,**] **L. Bako** [*] **F. Morelli** [*]

[*] *Laboratoire Ampère, UMR CNRS 5005, Ecole Centrale de Lyon, Université de Lyon, France*
[**] *Centre National de la Recherche Scientifique (CNRS), France*

Abstract: To obtain a consistent estimate when performing an identification with Prediction Error, it is important that the excitation yields informative data with respect to the chosen model structure. While the characterization of this property seems to be a mature research area in the linear case, the same cannot be said for nonlinear systems. In this work, we study the data informativity for a particular type of Hammerstein systems for two commonly-used excitations: white Gaussian noise and multisine. The real life example of the MEMS gyroscope is considered.

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

1. INTRODUCTION

In the Prediction Error (PE) identification framework, for the estimate to be consistent, the prediction error must be different for different values of the to-be-identified parameter vector. If the data used for the identification ensures this property, we say that the data are informative.

The data informativity has been extensively studied in the case of Linear Time-Invariant (LTI) systems. This has been done both for Single Input Single Output (SISO) systems (Ljung, 1999; Bazanella et al., 2012; Gevers et al., 2007, 2008, 2009) and for Multiple Inputs Multiple Outputs (MIMO) systems (Bazanella et al., 2010; Colin et al., 2019a,b,c,d).

While the PE identification framework is generally used for LTI systems, it can also be used as an efficient tool to identify certain classes of nonlinear systems. This is, e.g., the case for block-oriented systems with static nonlinearities (Hammerstein/Wiener systems) (Giri and Bai, 2010).

Adapting the PE framework to this type of systems entails a number of challenges. As an example, the identification problem boils down to a complex non-convex optimization problem. Consequently, a good initialization of this optimization problem is crucial and the best linear approximation framework can be used for this purpose (Schoukens et al., 2011, 2015; Schoukens and Tiels, 2017). Another issue (the one we will consider in the present paper) is that we have to ensure that the data used for the identification are informative (i.e. yields a different prediction error for all values of the to-be-identified parameter vector). Up to our knowledge, this problem has never been studied in the literature.

In this paper, we will tackle the data informativity for a particular type of block-oriented systems with two branches described by $y(t) = G_0(z)u(t) + P_0(z)u^n(t) + v(t)$ where u is the excitation, y the output, v the measurement noise and $n \in \mathbb{N}$. This system has a parallel Hammerstein structure with one monomial nonlinearity. For instance, the linear dynamics of the micro-electromechanical structure (MEMS) gyroscope with capacitive instrumentation can be modeled quite accurately with such representation with $n = 2$ (Saukoski, 2008; Kempe, 2011).

For this type of Hammerstein systems, we derive results in order to verify whether a given set of data is informative or not. We do that by rewriting the system as an equivalent system with one output and two inputs, i.e., $u_1 = u$ and $u_2 = u^n$ and we use data informativity results for Multiple Input Single Output (MISO) systems to tackle the data informativity problem for the considered parallel Hammerstein system.

Notations. For all matrices 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. For quasi-stationary signals x (Ljung, 1999), 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. Finally, for discrete-time systems, z denotes the forward-shift operator.

2. PREDICTION ERROR FRAMEWORK

Consider the parallel Hammerstein system $\mathcal{S}$ with one input u and one output y described by

$$\mathcal{S}: \quad y(t) = G_0(z)u(t) + P_0(z)u^n(t) + H_0(z)e(t) \qquad (1)$$

where $G_0(z)$ and $P_0(z)$ are stable transfer functions, $H_0(z)$ is a stable, inversely stable and monic [1] transfer function, e a white noise and n an integer that will be assumed to be known and such that $n \geq 2$. We will suppose that u and e are independent (i.e. the identification experiment

[1] i.e., $H_0(z = \infty) = 1$

is performed in open-loop). Therefore, u^n and e are also independent.

As already mentioned in the introduction, we want to identify a model of $\mathcal{S}$ by using the PE identification framework. For this purpose, one could consider a parametrized nonlinear model structure $\mathcal{M} = \{(f(u(t),\tilde{\theta}), H(z,\eta)) \mid \theta = (\tilde{\theta}^T, \eta^T)^T \in \mathcal{D}_\theta\}$ with $f(u(t),\tilde{\theta})$ a nonlinear function in $u(t)$ defined by $f(u(t),\tilde{\theta}) = G(z,\tilde{\theta}_G)u(t) + P(z,\tilde{\theta}_P)u^n(t)$ where $G(z,\tilde{\theta}_G)$, $P(z,\tilde{\theta}_P)$ and $H(z,\eta)$ are parametrized rational transfer functions, θ the parameter vector and $\tilde{\theta} = (\tilde{\theta}_G^T, \tilde{\theta}_G^T)^T$. The set $\mathcal{D}_\theta \subset \mathbb{R}^m$ restricts $\theta = (\tilde{\theta}_G^T, \tilde{\theta}_P^T, \eta^T)^T$ to those values for which $G(z,\tilde{\theta}_G)$ and $P(z,\tilde{\theta}_P)$ are stable and $H(z,\eta)$ is stable and inversely stable [2].

Instead of choosing this model structure, one can notice that the system $\mathcal{S}$ in (1) is equivalent to the following MISO system with the input vector $\mathbf{u} = (u, u^n)^T$:

$$\mathcal{S}: \quad y(t) = (G_0(z), P_0(z))\mathbf{u}(t) + H_0(z)e(t) \tag{2}$$

Therefore, it can be identified within the following MISO (linear) Box-Jenkins (BJ) model structure $\mathcal{M}'$ described by

$$\mathcal{M}' = \{(G(z,\tilde{\theta}_G), P(z,\tilde{\theta}_P), H(z,\eta)) | \ \theta = (\tilde{\theta}_G^T, \tilde{\theta}_P^T, \eta^T)^T \in \mathcal{D}_\theta\} \tag{3}$$

In the sequel, we will denote by μ_G (resp. μ_P) the dimension of $\tilde{\theta}_G$ (resp. $\tilde{\theta}_P$). Moreover, we will consider the so-called full-order assumption for $\mathcal{M}'$, i.e., $\exists\theta_0 = (\tilde{\theta}_{0,G}^T, \tilde{\theta}_{0,P}^T, \eta_0^T)^T \in \mathcal{D}_\theta$ such that $(G(z,\tilde{\theta}_{0,G}), P(z,\tilde{\theta}_{0,P}), H(z,\eta_0)) = (G_0(z), P_0(z), H_0(z))$. Finally, we will assume that the model structure $\mathcal{M}'$ is globally identifiable at θ_0, by considering that there is no pole-zero cancellation at θ_0 (Bazanella et al., 2012).

Assume that we have a set of N input-output data $Z^N = \{x(t) = (\mathbf{u}^T(t), y(t))^T \mid t = 1, \cdots, N\}$ collected on $\mathcal{S}$. From each $(G(z,\tilde{\theta}_G), P(z,\tilde{\theta}_P), H(z,\eta)) \in \mathcal{M}'$, we construct the one-step ahead predictor $\hat{y}(t,\theta)$ given by

$$\hat{y}(t,\theta) = W_{\mathbf{u}}(z,\theta)\mathbf{u}(t) + W_y(z,\theta)y(t) = W(z,\theta)x(t) \tag{4}$$

where

$$W_{\mathbf{u}}(z,\theta) = H^{-1}(z,\eta)(G(z,\tilde{\theta}_G), P(z,\tilde{\theta}_P)) \tag{5}$$

$$W_y(z,\theta) = 1 - H^{-1}(z,\eta) \tag{6}$$

$$W(z,\theta) = (W_{\mathbf{u}}(z,\theta), W_y(z,\theta)). \tag{7}$$

Based on the dataset Z^N, we compute the optimal parameter vector denoted $\hat{\theta}_N$ minimizing a least-square criterion on the prediction error $\epsilon(t,\theta) = y(t) - \hat{y}(t,\theta)$:

$$\hat{\theta}_N = \arg\min_{\theta\in\mathcal{D}_\theta} V_N(\theta, Z^N) \tag{8}$$

$$V_N\left(\theta, Z^N\right) = \frac{1}{N}\sum_{t=1}^{N}\epsilon^2(t,\theta) \tag{9}$$

We want $\hat{\theta}_N$ to be a consistent estimate of the true parameter vector θ_0, i.e., $\hat{\theta}_N \to \theta_0$ with a probability equal to 1 when $N \to \infty$. For that, it is important that $\mathcal{M}'$ be globally identifiable at θ_0 and that the data be informative with respect to (w.r.t.) $\mathcal{M}'$. The definition of the latter adapted to our problem is given below:

[2] $H(z,\eta)$ is moreover assumed monic, i.e., $H(z=\infty,\eta) = 1$.

***Definition** 1.* Consider the framework defined above with $\mathbf{u}^T = (u, u^n)$ and with the data $x(t) = (\mathbf{u}^T(t), y(t))^T$ collected by applying a quasi-stationary input u to the true system $\mathcal{S}$ in (1)-(2). Consider the model structure $\mathcal{M}'$ defined in (3) yielding the predictor $\hat{y}(t,\theta) = W(z,\theta)x(t)$. Define the set $\mathbf{\Delta_W} = \{\Delta W(z) = W(z,\theta') - W(z,\theta'') \mid \theta' \text{ and } \theta'' \text{ in } \mathcal{D}_\theta\}$. The data $x(t)$ are said to be informative w.r.t. the model structure $\mathcal{M}'$ when, for all $\Delta W(z) \in \mathbf{\Delta_W}$, we have

$$\bar{E}\left[||\Delta W(z)x(t)||^2\right] = 0 \implies \Delta W(z) \equiv \mathbf{0}_{1\times 3} \tag{10}$$

where $\Delta W(z) \equiv \mathbf{0}_{1\times 3}$ means that $\Delta W(e^{j\omega}) = \mathbf{0}_{1\times 3}$ at all or almost all ω. □

3. DATA INFORMATIVITY FOR MISO SYSTEMS IN OPEN-LOOP WITH TWO INPUTS

As already mentioned in the previous section, the nonlinear SISO system $\mathcal{S}$ with the input u in (1) can be rewritten as the MISO linear system with the input vector $\mathbf{u} = (u, u^n)^T$ in (2). Since we want a consistent estimate of $(G_0(z), P_0(z), H_0(z))$, the objective of this paper is to develop results in order to verify if a given experiment with an excitation u will yield informative data $x(t) = (\mathbf{u}^T(t), y(t))^T$ w.r.t. the MISO BJ model structure $\mathcal{M}'$. To derive these results, we will recall some data informativity results for MISO systems. These results are valid for arbitrary input vectors, i.e., $\mathbf{u}$ must not be necessarily equal to $(u, u^n)^T$. For this purpose, let us introduce the following notations: $X(z,\tilde{\theta}) = (G(z,\tilde{\theta}_G), P(z,\tilde{\theta}_P))$ and $\Delta X(z) = X(z,\tilde{\theta}') - X(z,\tilde{\theta}'')$ where $\tilde{\theta} = (\tilde{\theta}_G^T, \tilde{\theta}_P^T)^T$. Based on $\mathcal{D}_\theta$, we define the set $\mathcal{D}_{\tilde{\theta}} = \{\tilde{\theta} \mid \theta = (\tilde{\theta}^T, \eta^T)^T \in \mathcal{D}_\theta\}$. We also define the set $\mathbf{\Delta_X} = \{\Delta X(z) = X(z,\tilde{\theta}') - X(z,\tilde{\theta}'') \mid \tilde{\theta}' \text{ and } \tilde{\theta}'' \in \mathcal{D}_{\tilde{\theta}}\}$.

***Theorem** 1.* Consider data $x(t) = (\mathbf{u}^T(t), y(t))^T$ collected on a MISO system (2) in open loop. Consider also a model structure $\mathcal{M}'$ for this MISO system (see (3)) and the set $\mathbf{\Delta_X}$ defined above. Then, the data $x(t)$ are informative w.r.t. $\mathcal{M}'$ if and only if, for all $\Delta X(z) \in \mathbf{\Delta_X}$, we have

$$\bar{E}\left[||\Delta X(z)\mathbf{u}(t)||^2\right] = 0 \implies \Delta X(z) \equiv \mathbf{0}_{1\times 2} \tag{11}$$

□

Proof. See (Colin et al., 2019c). ■

In (Colin et al., 2019c), we have developed efficient conditions to verify whether a given input vector yields informative data (i.e., satisfies the condition in Theorem 1). As we will see in the sequel, these conditions may be more difficult to verify in the case of an input vector $\mathbf{u} = (u, u^n)^T$. Therefore, we present the following lemmas which give simpler (but more conservative) conditions for data informativity.

***Lemma** 1.* Consider Theorem 1. Assume that we collect the data $x(t) = (\mathbf{u}^T(t), y(t))^T$ by applying a quasi-stationary input vector $\mathbf{u}$ to the system $\mathcal{S}$ in (2). Assume that the power spectrum matrix $\Phi_{\mathbf{u}}(\omega)$ of the input $\mathbf{u}$ is strictly positive definite for almost all frequencies ω. Then, the data $x(t)$ are informative w.r.t. a full-order BJ model structure $\mathcal{M}'$ for $\mathcal{S}$ (see (3)). This result holds whatever the orders of the transfer functions G_0, P_0, and H_0 in (2) may be. □

Proof. From Parseval theorem, the left-hand side of (11) is equivalent to $\frac{1}{2\pi}\int_{-\pi}^{\pi}\Delta X(e^{j\omega})\Phi_{\mathbf{u}}(\omega)\Delta X^*(e^{j\omega})d\omega = 0$. Since $\Phi_{\mathbf{u}}(\omega)$ is positive definite for almost all frequencies ω, we have that $\bar{E}\left[||\Delta X(z)\mathbf{u}(t)||^2\right] = 0$ always implies $\Delta X(z) \equiv \mathbf{0}_{1\times 2}$ irrespectively of the complexity of $\Delta X \in \mathbf{\Delta_X}$, i.e., a full-order BJ model structure $\mathcal{M}'$ describing any system $\mathcal{S}$. This concludes the proof. ■

Lemma 1 cannot be used to verify the data informativity for multisine input vectors (since $\Phi_{\mathbf{u}}$ is never strictly positive definite in this case). Lemma 2 presents a data informativity condition for this type of excitation vectors.

Lemma 2. Consider Theorem 1 and suppose that the data $x(t) = (\mathbf{u}^T(t), y(t))^T$ are generated with a multisine input vector $\mathbf{u} = (u_1, u_2)^T$. Denote by Φ_{u_1} (resp. Φ_{u_2}) the power spectral density (PSD) of u_1 (resp. u_2). Suppose that Φ_{u_1} (resp. Φ_{u_2}) is non-zero at s_1 (resp. s_2) frequencies in $]-\pi,\pi]$. Finally, suppose that Φ_{u_1} and Φ_{u_2} are both non-zero at $s_{1,2}$ frequencies in this interval ($s_{1,2}$ can be equal to zero). Then, the data $x(t)$ are informative w.r.t. $\mathcal{M}'$ (see (3)) if $s_1 \geq \mu_G$ and $s_2 - s_{1,2} \geq \mu_P$ where μ_G and μ_P are defined below (3). □

Proof. Let us first observe that the multisine u_2 can be decomposed as follows $u_2 = u_2^{(/\!\!/ u_1)} + u_2^{(\perp u_1)}$ where

- $u_2^{(/\!\!/ u_1)}$ is the multisine whose PSD shares the same frequencies with the PSD of u_1, i.e., the multisine part of u_2 that is totally correlated to u_1. The PSD of $u_2^{(/\!\!/ u_1)}$ is non-zero in $s_{1,2}$ frequencies.
- $u_2^{(\perp u_1)}$ is the multisine whose PSD does not share any frequency with the PSD of u_1, i.e., the multisine part of u_2 that is not correlated to u_1 (and to $u_2^{(/\!\!/ u_1)}$ too by construction). The PSD of $u_2^{(\perp u_1)}$ is non-zero in $s_2 - s_{1,2}$ frequencies.

Consequently, the left hand side of (11) is equivalent to the following equation system, for all $\Delta X(z) = (\Delta G(z), \Delta P(z)) \in \mathbf{\Delta_X}$,

$$\begin{cases} \bar{E}\left[||\Delta G(z)u_1(t) + \Delta P(z)u_2^{(/\!\!/ u_1)}(t)||^2\right] = 0 \\ \bar{E}\left[||\Delta P(z)u_2^{(\perp u_1(t))}(t)||^2\right] = 0 \end{cases} \quad (12)$$

We have to prove that (12) implies $\Delta X = (\Delta G, \Delta P) \equiv \mathbf{0}_{1\times 2}$. The PSD of $u_2^{(\perp u_1(t))}$ is non-zero in at least μ_P different frequencies in $]-\pi,\pi]$. Therefore, the second equation of (12) implies $\Delta P(z) \equiv 0$ (Ljung, 1999; Gevers et al., 2008). By injecting the latter in the first equation of (12), we obtain $\bar{E}\left[||\Delta G(z)u_1(t)||^2\right] = 0$. Since the PSD of u_1 is non-zero in at least μ_G frequencies in the set $]-\pi,\pi]$ and so $\bar{E}\left[||\Delta G(z)u_1(t)||^2\right] = 0$ implies that $\Delta G(z) \equiv 0$, which concludes the proof from Theorem 1. ■

As already mentioned, Lemmas 1 and 2 pertain to an arbitrary input vector $\mathbf{u}$. They can therefore also be used in the case where the input vector $\mathbf{u} = (u_1, u_2)^T$ is of the form $u_1 = u$ and $u_2 = u^n$. This fact will be used in the sequel to derive data informativity results for the Hammerstein system $\mathcal{S}$ in (1).

4. WHITE GAUSSIAN NOISE EXCITATION

As shown in Appendix A, if the input signal u of (1) is chosen as a zero-mean Gaussian white noise, the PSD of $\mathbf{u} = (u, u^n)^T$ is strictly positive definite at (almost) all frequencies (whatever the value of $n \geq 2$). Consequently, we have the following result.

Theorem 2. Consider Theorem 1. Assume that we collect the data $x(t) = (\mathbf{u}^T(t), y(t))^T$ with $\mathbf{u} = (u, u^n)^T$ by applying a *zero-mean white Gaussian noise* u to the system $\mathcal{S}$ in (1). Then, the data $x(t)$ are informative with respect to a full-order BJ model structure $\mathcal{M}'$ describing any system $\mathcal{S}$. □

Proof. See Appendix A. ■

This type of stochastic excitation is interesting since it can allow to identify any system of the type (1).

5. MULTISINE EXCITATION

In this section, u is a sum of m cosinusoids given by

$$u(t) = \sum_{l=1}^{m} A_l \cos(\omega_l t + \phi_l) \quad (13)$$

where $A_l > 0$ and ϕ_l are respectively the amplitude and the phase-shift of the cosinusoid at the *non-zero* frequency ω_l belonging to the normalized frequency [3] interval $]0,\pi[$. Since u is a multisine, u^n is also a multisine, excited at more cosinusoids than u. In (Colin et al., 2019c), we gave a condition to verify if a given multisine input vector $\mathbf{u}$ will yield informative data w.r.t. $\mathcal{M}'$ by verifying the rank of a matrix depending on the model structure complexity and on the amplitudes, phase-shifts and frequencies of the cosinusoids in $\mathbf{u}$. To apply it to our problem, we have to compute the amplitudes, phase-shifts and frequencies of u^n which can become computationally expensive when n and m increase. Fortunately, using Lemma 2, we can derive a sufficient condition for data informativity that only requires the knowledge of the number of frequencies present in u^n (and not its full expression).

Theorem 3. Consider that the system $\mathcal{S}$ in (1) is excited with the multisine (13) where $\omega_l \in]0,\pi[$ $(l = 1,\cdots,m)$. Define s_2 and $s_{1,2}$ as in Lemma 2 for $\Phi_{u_1} = \Phi_u$ and $\Phi_{u_2} = \Phi_{u^n}$. Then, the data are informative w.r.t. $\mathcal{M}'$ (see (3)) if $2m \geq \mu_G$ and if $s_2 - s_{1,2} \geq \mu_P$ where μ_G and μ_P are defined below (3). □

Proof. Straightforward consequence of Lemma 2. ■

As already mentioned, to use the (sufficient) data informativity condition of Theorem 3, we only need to know how many frequencies are present in u^n and to compare them with the ones in u (see (13)).

For this purpose, one could compute the Fast Fourier Transform (FFT) of the sequence $\{u^n(t) \mid t = 1,\cdots,M\}$ for a sufficiently large value of M. Another procedure to this end will be derived in the next subsection. In this procedure, we will formally suppose that $\phi_l = 0$ $(l = 1,\cdots,m)$ in (13). See Remark 1 for further details.

[3] The normalized frequency ω_l is obtained from the true frequency $\tilde{\omega}_l$ by $\omega_l = \tilde{\omega}_l / f_s$ where f_s is the sampling frequency.

5.1 Method for the computation of s_2

Let us thus consider (13) with $A_l > 0$, $\phi_l = 0$ and $\omega_l \in]0, \pi[$ $(l = 1, \cdots, m)$ and let us observe that $u^n(t) = u^{n-1}(t)u(t)$, $u^{n-1}(t) = u^{n-2}(t)u(t)$, $\cdots$, $u^2(t) = u(t)u(t)$. Therefore, to get the cosinusoids of u^n, we have first to determine the ones of u^2 from u, then the ones of u^3 from u^2, $\cdots$, the ones of u^{n-1} from u^{n-2} and finally the ones of u^n from u^{n-1}.

Let us now study how to obtain the frequencies of u^p from the ones of u^{p-1} $(p = 2, \cdots, n)$. For this purpose, we know that u^{p-1} is a multisine and so $u^p(t) = u^{p-1}(t)u(t)$ can be written as the sum of the products of each cosinusoid in u by each cosinusoid in u^{p-1}. By using the fact that $\cos(\alpha)\cos(\beta) = 1/2(\cos(\alpha+\beta)+\cos(\alpha-\beta))$ $\forall(\alpha, \beta) \in \mathbb{R}^2$, all the products in this sum can be written as the sum of two cosinusoids.

By doing this, we see that the set of frequencies in u^p can be determined by adding and substracting ω_l $(l = 1, \cdots, m)$ to the frequencies of each cosinusoid present in u^{p-1}. Of course, in the obtained set, we have to remove the duplicates (e.g., the frequency $-\omega'$ is equivalent to the frequency ω' and the frequency $\omega' + 2k\pi$ with $k \in \mathbb{Z}$ is equivalent to ω'). We then obtain a set of frequencies in the interval $[0, \pi]$. Let us denote by m' the number of frequencies in this set. Then, $s_2 = 2m'$ if the set contains neither the frequency 0 nor π while $s_2 = 2m' - 1$ if it contains either 0 or π and $s_2 = 2m' - 2$ if it contains both 0 and π.

The above procedure supposes that no terms in the sum of products can cancel out. This is the reason why we suppose that $A_l > 0$ and $\phi_l = 0$. Consequently, all the terms in the summation will be characterized by a positive amplitude and a zero phase-shift and no cancellations can then occur.

Let us illustrate the above procedure on an example.

***Example** 1.* Consider $n = 3$, $m = 2$ and

$$u(t) = \cos(\omega_1 t) + \cos(\omega_2 t)$$

with $\omega_1 = 0.2$ and $\omega_2 = 0.3$. To compute s_2, we need the frequencies of the cosinusoids in u^3. For this purpose, we need first to determine the ones in u^2 from u.

► For each frequency in u (i.e., 0.2 and 0.3), we add and subtract $\omega_1 = 0.2$ and $\omega_2 = 0.3$. With ω_1 we obtain the terms 0.4, 0, 0.5, 0.1 and with ω_2 we have 0.5, -0.1, 0.6, 0. By removing the duplicates, we obtain the following frequencies in u^2: 0, 0.1, 0.4, 0.5 and 0.6.

► For each frequency in u^2 (i.e., 0, 0.1, 0.4, 0.5 and 0.6), we add and subtract $\omega_1 = 0.2$ and $\omega_2 = 0.3$. With ω_1, we obtain the terms 0.2, -0.2, 0.3, -0.1, 0.6, 0.2, 0.7, 0.3, 0.8 and 0.4 and with ω_2, we have 0.3, -0.3, 0.4, -0.2, 0.7, 0.1, 0.8, 0.2, 0.9 and 0.3. By removing the duplicates, we obtain the 8 following frequencies in u^3: 0.1, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8 and 0.9, all belonging to $[0, \pi]$. We see that $m' = 8$ and that $s_2 = 16$. Since the frequencies ± 0.2 and ± 0.3 are also in Φ_u, we have that $s_{1,2} = 4$ and thus that $s_2 - s_{1,2} = 12$. □

***Remark** 1.* As already mentioned, the above procedure is not guaranteed to yield the right s_2 if the phase-shifts are not all equal to zero. To show this let us consider the following multisine excitation $u(t) = 1/2\cos(0.4t) + \cos(0.6t) + \cos(t + \pi/2)$ and let us apply the above procedure for $n = 2$: we obtain the set of frequencies 0, 0.1, 0.2, 0.8, 0.9, 1, 1.1, and 1.2. However, the frequency 1 does not appear in u^2 due to the cancellations of the terms at this frequency. Note, however, that such a situation is rare in practice and we can therefore carefully use the procedure in this section when the phase-shifts are non-zero. Nonetheless, a verification using the FFT can be always useful. □

5.2 Result without computation of s_2

In some cases, we even do not have to compute s_2 to check the data informativity.

***Lemma** 3.* Consider that the excitation of (1) (with $n \geq 2$) is given by $u(t) = \sum_{l=1}^{m} A_l \cos(\omega_l t)$ with $A_l > 0$ and $\omega_l \in]0, \pi[$ $(l = 1, \cdots, m)$ and such that the m frequencies $n\omega_l$ are

- all different between them (modulo 2π).
- all different from π and 0 (modulo 2π).
- all different from each ω_j $(j = 1, \cdots, m)$ (modulo 2π).

Then, the data $x(t) = (\mathbf{u}^T(t), y(t))$ are informative with respect to $\mathcal{M}'$ if $m \geq \max(\mu_G/2, \mu_P/2)$ where μ_G and μ_P are defined below (3). □

Proof. Following the procedure in Section 5.1, we see that $n\omega_1$, $\cdots$, $n\omega_m$ are frequencies of the multisine u^n. Since these m positive frequencies are all different (modulo 2π) and different from π (modulo 2π), this implies that the PSD of u^n will be non-zero at at least $2m$ frequencies in $]-\pi, \pi]$, i.e., $s_2 \geq 2m$. Due to the fact that the m frequencies $n\omega_l$ are (modulo 2π) different from the frequencies ω_j $(j = 1, \cdots, m)$ in u, we have also that $s_2 - s_{1,2} \geq 2m$. The result then follows from Theorem 3. ■

***Remark** 2.* The idea of Lemma 3 is to give a lower bound for $s_2 - s_{1,2}$ which is $2m$. However, this bound is conservative. In Example 1 in Section 5.1 where $n = 3$ and $m = 2$, we have seen that $s_2 - s_{1,2} = 12$ while the lower bound is equal to $2m = 4$. □

5.3 Synthesis of the results

Let us now summarize the different results of this section by giving a general approach to verify the informativity w.r.t. $\mathcal{M}'$ for a given multisine u. First, we verify that $m \geq \mu_G/2$. If it is not the case, the data $x(t)$ are certainly not informative with respect to $\mathcal{M}'$.

If $m \geq \mu_G/2$, we can check whether the (conservative) condition of Lemma 3 is satisfied. If it is not the case, we compute s_2 and $s_{1,2}$ (using the FFT approach or the procedure of Section 5.1) and we verify the (less conservative) condition of Theorem 3. If this condition is still not validated, we need to compute the full expression of u^n (with all amplitudes and phase-shifts) and use the results of (Colin et al., 2019c) on this expression and the one of u (see (13)).

6. REAL LIFE EXAMPLE: MEMS GYROSCOPE

6.1 MEMS gyroscope description

The MEMS gyroscope is an inertial sensor used to measure angular rates by using Coriolis effect (see (Kempe, 2011; Saukoski, 2008) for details of its working principle). By focusing on its main dynamics and from physics laws, the MEMS can be modeled by (1) with $n = 2$ and where [4]

- $G_0(z)$ illustrates a parasite electrical bond between the excitation and measurement capacitive circuits:
$$G_0(z) = 10^{-2}\frac{9.47z^{-1} + 6.69z^{-2} - 16.21z^{-3}}{1 - 0.685z^{-1} + 0.175z^{-2} - 0.0415z^{-3}}$$
- $P_0(z)$ describes the mechanical motion of the MEMS which is a resonance with a high quality factor:
$$P_0(z) = 10^{-3}\frac{2.39z^{-1} - 5.47z^{-2}}{1 - 0.743z^{-1} + z^{-2}}$$

We will consider that e is Gaussian with a variance of 10^{-3} and $H_0(z) = 1$ for the sake of simplicity and to verify easily the consistency of the estimator. All data used in this example are simulated from the above true system in order to verify the consistency. The sampling frequency f_s is equal to 62500Hz. Finally, we choose the full-order model structure $\mathcal{M}'$ with the same complexity as $(G_0(z), P_0(z), H_0(z))$. Consequently, it is globally identifiable at θ_0 and the number of parameters to be identified in $G(z, \theta)$ (resp. in $P(z, \theta)$) is equal to $\mu_G = 6$ (resp. $\mu_P = 4$).

6.2 White Gaussian noise excitation

From Theorem 2, the data are informative with respect to any rational model structure $\mathcal{M}'$ when u is a white Gaussian noise. We do 100 Monte-Carlo simulations to illustrate the consistency with 100 realizations of the noise e, with a white Gaussian excitation u of variance 1 and a data number $N = 5000$ for each identification. By computing the mean of the 100 computed parameter vectors, we obtain the following model:

$$\begin{cases} G(z, \hat{\theta}_G) = 10^{-2}\dfrac{9.48z^{-1} + 6.68z^{-2} - 16.21z^{-3}}{1 - 0.685z^{-1} + 0.175z^{-2} - 0.0416z^{-3}} \\ P(z, \hat{\theta}_P) = 10^{-3}\dfrac{2.32z^{-1} - 5.42z^{-2}}{1 - 0.743z^{-1} + z^{-2}} \end{cases}$$

The closeness between the identified model and the true one suggests consistency of the estimator.

6.3 Multisine excitation

We consider the following multisine excitation for u
$$u(t) = \cos(\omega_1 t) + \cos(\omega_2 t) + 0.5\cos(\omega_3 t)$$
where ω_1, ω_2 and ω_3 are the normalized frequencies [5] are given by $\omega_1 = 0.08$, $\omega_2 = 1.19$ and $\omega_3 = 1.92$, corresponding to the true frequencies $\tilde{\omega}_1 = 503$ rad/s, $\tilde{\omega}_2 = 74362$ rad/s and $\tilde{\omega}_3 = 120001$ rad/s respectively. First, we have indeed $m = 3 \geq \mu_G/2$. In this case, the sinusoid frequencies have been chosen such that we can verify the condition of Lemma 3. We have indeed that $m = 3 \geq \max(\mu_G/2, \mu_P/2)$ and so the data will be informative w.r.t. $\mathcal{M}'$ with this excitation. Let us verify it by doing 100 Monte-Carlo simulations to illustrate the consistency with 100 realizations of the noise e, with a data number $N = 5000$ for each identification. By computing the mean of the 100 computed parameter vectors, we obtain the following model:

$$\begin{cases} G(z, \hat{\theta}_G) = 10^{-2}\dfrac{9.53z^{-1} + 6.65z^{-2} - 16.21z^{-3}}{1 - 0.691z^{-1} + 0.176z^{-2} - 0.0420z^{-3}} \\ P(z, \hat{\theta}_P) = 10^{-3}\dfrac{2.40z^{-1} - 5.50z^{-2}}{1 - 0.743z^{-1} + z^{-2}} \end{cases}$$

Here again, the closeness between the identified model and the true one illustrates the consistency of the estimator.

[4] The given transfer functions $G_0(z)$ and $P_0(z)$ are the ones obtained with the approach in (Colin et al., 2019e).

[5] Recall that the normalized frequencies ω_l are obtained from the true ones $\tilde{\omega}_l$ by $\omega_l = \tilde{\omega}_l/f_s$.

7. CONCLUSION

In this paper, we have studied the data informativity with respect to a particular parallel Hammerstein system with an input monomial nonlinearity. We have considered two commonly-used signals in Prediction Error Identification: white Gaussian noise and multisine. In the white Gaussian noise case, we can identify any model structure. For multisine excitation, we give some advice on the number of cosinusoids to use. A real life example has been considered to illustrate the results. For future works, we want to study the data informativity property for most complex Hammerstein/Wiener systems.

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 Miškovic, 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. (2019a). Closed-loop Identification of MIMO Systems in the Prediction Error Framework: Data informativity Analysis. Provisionally accepted for a publication in Automatica, https://hal.archives-ouvertes.fr/hal-02351669v1.

Colin, K., Bombois, X., Bako, L., and Morelli, F. (2019b). Data Informativity for the Identification of MISO FIR Systems with Filtered White Noise Excitation. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, 1734–1739.

Colin, K., Bombois, X., Bako, L., and Morelli, F. (2019c). Data Informativity for the Open-Loop Identification of Multivariate System in the Prediction Error Framework. To be published in Automatica, https://hal.archives-ouvertes.fr/hal-02305057v1.

Colin, K., Bombois, X., Bako, L., and Morelli, F. (2019d). Informativity: how to get just sufficiently rich for the Identification of MISO FIR Systems with Multisine Excitation? In *2019 18th European Control Conference (ECC)*, 351–356.

Colin, K., Saggin, F., Le Blanc, C., Bombois, X., Korniienko, A., and Scorletti, G. (2019e). Identification-Based

Approach for Electrical Coupling Compensation in a MEMS Gyroscope. In *2019 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, 1–4.

Gevers, M., Bazanella, A., and Miškovic, L. (2008). Informative data: How to get just sufficiently rich? In *2008 47th IEEE Conference on Decision and Control.*

Gevers, M., Bazanella, A.S., Bombois, X., and Miškovic, L. (2009). Identification and the Information Matrix: How to Get Just Sufficiently Rich? *IEEE Transactions on Automatic Control*, 54(12), 2828–2840.

Gevers, M., Bazanella, A.S., and Mišković, L. (2007). Identifiability and informative experiments in open and closed-loop identification. In *Modeling, Estimation and Control*, 151–170. Springer.

Giri, F. and Bai, E.W. (2010). *Block-oriented nonlinear system identification*, volume 1. Springer.

Kempe, V. (2011). *Inertial MEMS: principles and practice.* Cambridge University Press.

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 edition.

Papoulis, A. and Pillai, U. (2002). *Probability, Random Variables and Stochastic Processes.* McGraw-Hill Europe, 4th edition.

Saukoski, M. (2008). System and circuit design for a capacitive MEMS gyroscope. Ph.D. dissertation.

Schoukens, M., Marconato, A., Pintelon, R., Vandersteen, G., and Rolain, Y. (2015). Parametric identification of parallel Wiener-Hammerstein systems. *Automatica*, 51, 111–122.

Schoukens, M., Pintelon, R., and Rolain, Y. (2011). Parametric identification of parallel Hammerstein systems. *IEEE Transactions on Instrumentation and Measurement*, 60(12), 3931–3938.

Schoukens, M. and Tiels, K. (2017). Identification of block-oriented nonlinear systems starting from linear approximations: A survey. *Automatica*, 85, 272–292.

Appendix A. PROOF OF THEOREM 2

First, since u is a white Gaussian noise with a non-zero variance, u^n is also a white noise with a non-zero variance. We will need the expectation value of $u^n(t)$ when $u(t)$ is Gaussian. It is given in the next lemma.

Lemma 4. (Papoulis and Pillai, 2002). Consider a zero-mean Gaussian variable X with variance equal to σ^2. Then,

$$E[X^n] = \begin{cases} \sigma^n (n-1)!! & \text{when } n \text{ is even.} \\ 0 & \text{when } n \text{ is odd.} \end{cases} \tag{A.1}$$

where the operator !! is defined for odd integer p by $p!! = p \times (p-2) \times (p-4) \times \cdots \times 3 \times 1$. □

We are going to prove that the power spectrum $\Phi_{\mathbf{u}}(\omega)$ of $\mathbf{u}$ is positive definite at almost all frequencies ω, which will conclude the proof from Lemma 1. For that, let us calculate $\Phi_{\mathbf{u}}(\omega)$ by taking the Fourier transform of the correlation matrix $R_{\mathbf{u}}(\tau)$ given by

$$R_{\mathbf{u}}(\tau) = \begin{pmatrix} \bar{E}[u(t)u(t+\tau)] & \bar{E}[u(t)u^n(t+\tau)] \\ \bar{E}[u^n(t)u(t+\tau)] & \bar{E}[u^n(t)u^n(t+\tau)] \end{pmatrix}$$

Let us denote $\sigma_u^2 = \bar{E}[u^2(t)]$. By using the fact that u and u^n are white and that u is zero-mean, we have that, from Lemma 4,

$$\bar{E}[u(t)u(t+\tau)] = \sigma_u^2 \delta(\tau)$$
$$\bar{E}[u(t)u^n(t+\tau)] = \begin{cases} 0 & \text{when } n \text{ is even.} \\ n!!\ \sigma_u^{n+1}\delta(\tau) & \text{when } n \text{ is odd.} \end{cases}$$
$$\bar{E}[u(t+\tau)u^n(t)] = \bar{E}[u(t)u^n(t-\tau)]$$

For the calculation of $\bar{E}[u^n(t)u^n(t+\tau)]$, we give here the details since it is not as simple as the previous ones.

First, when $\tau \neq 0$, we have that $\bar{E}[u^n(t)u^n(t+\tau)] = E[u^n(t)]E[u^n(t+\tau)]$ since u^n is a white noise. If n is odd, $\bar{E}[u^n(t)u^n(t+\tau)] = 0$. If n is even, $\bar{E}[u^n(t)u^n(t+\tau)] = (\sigma_u^n(n-1)!!)^2 = \sigma_u^{2n}((n-1)!!)^2$ from Lemma 4. For $\tau = 0$, we have that $\bar{E}[u^n(t)u^n(t+\tau)] = \bar{E}[u^{2n}(t)] = \sigma_u^{2n}(2n-1)!!$.

We deduce the expression for $\bar{E}[u^n(t)u^n(t+\tau)]$:

- when n is even
$$\bar{E}[u^n(t)u^n(t+\tau)] = \sigma_u^{2n}\left[(2n-1)!! - ((n-1)!!)^2\right]\delta(\tau) + \sigma_u^{2n}((n-1)!!)^2$$
- when n is odd
$$\bar{E}[u^n(t)u^n(t+\tau)] = \sigma_u^{2n}(2n-1)!!\delta(\tau)$$

By taking the Fourier transform of $R_{\mathbf{u}}(\tau)$, the power spectrum matrix $\Phi_{\mathbf{u}}(\omega)$ is given by

$$\Phi_{\mathbf{u}}(\omega) = \begin{cases} \begin{pmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_u^{2n}\left[(2n-1)!! - ((n-1)!!)^2\right] \end{pmatrix} \\ + \begin{pmatrix} 0 & 0 \\ 0 & \sigma_u^{2n}(2n-1)!! \end{pmatrix} \delta(\omega) & \text{when } n \text{ is even.} \\ \begin{pmatrix} \sigma_u^2 & \sigma_u^{n+1} n!! \\ \sigma_u^{n+1} n!! & \sigma_u^{2n}(2n-1)!! \end{pmatrix} & \text{when } n \text{ is odd.} \end{cases}$$

Since the power spectrum matrix $\Phi_{\mathbf{u}}(\omega)$ is positive semi-definite at all frequencies ω, let us prove that the determinant of $\Phi_{\mathbf{u}}(\omega)$ is non-zero for almost all ω to prove that it is strictly positive definite at almost all frequencies ω. When

- n is even, for all frequencies $\omega \neq 0$
$$\det(\Phi_{\mathbf{u}}(\omega)) = \sigma_u^{2n+2}\left[(2n-1)!! - ((n-1)!!)^2\right]$$
- n is odd, for all frequencies ω
$$\det(\Phi_{\mathbf{u}}(\omega)) = \sigma_u^{2n+2}\left[(2n-1)!! - (n!!)^2\right]$$

Let us prove that $(2n-1)!! - ((n-1)!!)^2 \neq 0$ by proving that $(2n-1)!!/((n-1)!!)^2 > 1$ when n is non-zero and even. The proof that $(2n-1)!! - (n!!)^2 > 1$ when n is odd is based on the same principle. Let us first observe that

$$\frac{(2n-1)!!}{((n-1)!!)^2} = \frac{(2n-1)\times(2n-3)\times\cdots\times(n+1)\times(n-1)!!}{((n-1)!!)^2}$$
$$= \frac{(2n-1)\times(2n-3)\times\cdots\times(n+1)}{(n-1)\times\cdots\times 3\times 1}$$

The numerator and the denominator of the latter are the product of $n/2$ factors. Since the minimal factor of the numerator is strictly greater than the maximal factor of the denominator for $n \geq 2$, then $(2n-1)!! > ((n-1)!!)^2$. Therefore, $\det(\Phi_{\mathbf{u}}(\omega)) > 0$ for almost all ω. With Lemma 1, the conclusion follows. ■