To appear in Automatica, 2023

On the informativity of direct identification experiments in dynamical networks

X. Bombois [a,b] K. Colin [c] P.M.J. Van den Hof [d] H. Hjalmarsson [c]

[a] *Laboratoire Ampère, Ecole Centrale de Lyon, Université de Lyon, Ecully, France*

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

[c] *Centre for Advanced Bio Production and Division of Decision and Control Systems, School of Electrical Engineering and Computer Science, KTH, Stockholm, Sweden*

[d] *Control Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands*

Abstract

Data informativity is a crucial property to ensure the consistency of the prediction error estimate. This property has thus been extensively studied in the open-loop and in the closed-loop cases, but has only been briefly touched upon in the dynamic network case. In this paper, we consider the prediction error identification of the modules in a row of a dynamic network using the full input approach. Our main contribution is to propose a number of easily verifiable data informativity conditions for this identification problem. Among these conditions, we distinguish a sufficient data informativity condition that can be verified based on the topology of the network and a necessary and sufficient data informativity condition that can be verified via a rank condition on a matrix of coefficients that are related to a full-order model structure of the network. These data informativity conditions allow to determine different situations (i.e., different excitation patterns) leading to data informativity. In order to be able to distinguish between these different situations, we also propose an optimal experiment design problem that allows to determine the excitation pattern yielding a certain pre-specified accuracy with the least excitation power.

Key words: Dynamic network identification, data informativity, Optimal experiment design

1 Introduction

In prediction error identification, it is crucial that the excitation yields informative data. Indeed, provided that the model structure describes the true system with a unique true parameter vector θ_0, data informativity is a necessary and sufficient condition for the prediction error estimate of θ_0 to be consistent (i.e., to converge to θ_0 with probability one when the number of data tends to infinity) [14]. In a nutshell, data informativity is obtained when the excitation signal is sufficiently rich for the prediction error to distinguish the different models in the chosen model structure. Due to its crucial importance, the literature provides a large number of contributions on data informativity. See e.g., [14,9,6,5]. In these papers, the identification of single-input single-output (SISO) and multiple-input multiple-output (MIMO) systems is treated in both open-loop and closed-loop configurations. This paper addresses data informativity in dynamic networks (i.e., another important configuration).

We will here consider the dynamic network framework introduced in [7]. In this framework, a dynamic network is represented by a number N_{mod} of nodes and each of these nodes is characterized by a measurable scalar signal w_j ($j = 1, ..., N_{mod}$). The signal w_j at Node j is related to the signals w_k at other nodes ($k \neq j$) through causal transfer functions $G_{0,jk}(z)$ (also called modules). If Node l is not connected to Node j, the corresponding module $G_{0,jl}(z)$ is then equal to zero. The signal w_j is also possibly a function of exogenous signals: an unknown process noise v_j and/or a known external excitation signal r_j added for identification purpose. These exogenous signals v_j and r_j are not necessarily present at all nodes j. The dynamic network is thus entirely determined by the matrix $\bar{G}_0(z)$ gathering the transfer functions $G_{0,jk}(z)$ and its noise and excitation pattern i.e., the description of the nodes at which a process noise v_j is present (which is not an user choice) and of the nodes at which an external excitation r_j is present (which is at least partly an user choice). In the literature, different approaches have been developed to derive an (accurate) model of (part of) the matrix $\bar{G}_0(z)$. Like in closed-loop

identification, we can distinguish, on the one hand, direct approaches where the node signals w_k are used as data for the identification of the elements of $\bar{G}_0(z)$ [7] and, on the other hand, indirect approaches where the model(s) of the module(s) of $\bar{G}_0(z)$ are back-computed from an identified model of the *closed-loop* transfer matrix between the external excitation signals and the node signals [12].

Data informativity is as crucial for (prediction error) dynamic network identification as it is for SISO and MIMO identification. Indeed, in order to derive consistent models in dynamic network identification, the noise and excitation pattern must lead to sufficiently rich data for the prediction error to distinguish the different models in the chosen model structure. However, until now, even though some preliminary results can be found in e.g., [7,10,8], data informativity for dynamic network identification has only been briefly touched upon in the literature. As opposed to *data informativity*, the concept of *network identifiability* has been extensively explored (see e.g., [12,17]). In a nutshell, a network is said *identifiable* if we can uniquely retrieve (part of) the open-loop representation of the network (i.e., $\bar{G}_0(z)$) from closed-loop representations of the network (i.e., representations of the tranfer between the exogenous signals and (some of) the node signals). Network identifiability and data informativity are thus clearly different concepts and a deep analysis of data informativity conditions for dynamic network identification is thus absent in the literature. In this paper, we will address the data informativity issue for one particular dynamic network identification method: the so-called *full input approach*. The full input approach pertains to the identification of all the unknown modules of a row of the matrix $\bar{G}_0(z)$ via the node signals w_k (it is thus a *direct* dynamic network method) [7]. The full input approach may be a way to identify one single module in this row (see [7] for more details).

In [7] (where the full input approach was introduced), the issue of data informativity was tackled via a condition on the power spectrum matrix of the data used for the identification. However, it is difficult to determine if a given noise and excitation pattern yields data informativity (or not) using this condition. Moreover, as pointed out in [10], it is also difficult to interpret this condition in order to determine at which node(s) excitation signals r_k have to be added if the current noise and excitation pattern does not lead to data informativity. In the present paper, our first contribution is to reformulate the data informativity condition proposed in [7]. With this reformulation, data informativity can be inferred, in some cases, with the sole excitation of the process noises v_j present in the network and, in other cases, by the addition of filtered white noise excitations r_j at some nodes. The data informativity condition requires that a part of the transfer matrix $\bar{S}_0(z) = (I - \bar{G}_0(z))^{-1}$ is full row rank. Using the results in [12], this rank condition can be verified via the analysis of the graph of the network (i.e., the interconnection structure of the network). Using this graph interpretation, we will see that it is then straightforward to verify if a given noise and excitation pattern yields data informativity and that we can also easily determine at which node(s) excitation signals r_k have to be added if the current noise and excitation pattern does not lead to data informativity. We will also observe that one noise and excitation pattern leading to data informativity is the one proposed in [8], but many other excitation patterns are also possible.

The main characteristic of the data informativity condition proposed in [7] and of its reformulation derived in this paper is that they ensure data informativity for any model structure. This means that, if a noise and excitation pattern respects this condition, it will ensure consistent estimates of modules that can have infinite complexity. While this could be seen as an advantage in some cases, this also means that, when the to-be-identified modules are of restricted complexity, many noise and excitation patterns that would lead to data informativity will not be detected by these conditions. In particular, we will see that multisine excitation signals yielding data informativity will never be detected by these conditions.

In order to be able to more finely detect when data informativity is obtained, a second and important contribution of this paper is to derive a data informativity condition that takes into account the complexity of the to-be-identified modules. This condition is a necessary and sufficient condition to verify whether, for a given noise and excitation pattern and for given multisine or filtered white noise excitation signals, the obtained data are informative for a full-order model structure of the to-be-identified modules. This necessary and sufficient condition has a rather complex form. However, using the framework introduced in [5], we show that its verification boils down to a rank condition on a matrix of coefficients. When the data are not informative for a given choice of external excitations, the insights developed in [5] on this rank condition can be used to determine the necessary measures to increase the informativity of the data.

Using these data informativity conditions, we can determine a set of situations in which the addition at certain nodes of excitation signals with given power spectra leads to consistent estimates of the to-be-identified modules, i.e. to models that converge to the true value of the module when the number of data tends to infinity. All these situations are thus equivalent when the number of data tends to infinity. However, for a finite data set, the respective accuracy can be much different. Another contribution of the paper is to propose to use optimal experiment design to distinguish between these situations,. In particular, we will determine the particular excitation pattern that leads to the desired accuracy of the to-be-identified modules with the smallest excitation power. The use of optimal experiment design for this purpose was first introduced in our paper [3] where a very specific type of network is considered, namely the interconnection of simple closed-loop systems. We here extend this work towards the generic network descrip-

tion in [7]. It is to be noted that, in [15], a similar problem is also considered, but for an indirect identification approach.

The sequel of the paper will be organized as follows. In Section 2, we will describe in more details the considered dynamic network. In Section 3, the full input approach will be presented and we will show that data informativity is indeed crucial to derive consistent estimates of the to-be-identified modules. In Section 4, we will then present the more conservative data informativity conditions while, in Section 5, we will present our necessary and sufficient data informativity condition. Section 6 will pertain to the optimal experiment design approach to detect an optimal excitation pattern.

Notations: In this paper, vectors of discrete-time signals and matrices/vectors of discrete-time transfer functions will be denoted with a bar: $\bar{x}(t)$ and $\bar{X}(z)$ (t represents the sample number and z denotes both the Z-transform variable and the shift operator). We denote by $x_i(t)$ (resp. $X_{ik}(z)$) the i^{th} entry of the vector of signals $\bar{x}(t)$ (resp. the (i,k)-entry of the matrix of transfer functions $\bar{X}(z)$). To define parts of $\bar{x}(t)$ and $\bar{X}(z)$, we will use calligraphic symbols such as $\mathcal{X}$, $\mathcal{Y}$ to denote set of indexes corresponding to the entries of $\bar{x}(t)$ or corresponding to the rows and columns of $\bar{X}(z)$. The cardinality of a set of indexes $\mathcal{X}$ will be denoted by $n_{\mathcal{X}}$. For a vector of signals $\bar{x}(t)$, $\bar{x}_{\mathcal{X}}(t)$ is the vector of dimension $n_{\mathcal{X}}$ obtained by only conserving the entries in $\mathcal{X}$ ($\bar{x}_{\mathcal{X}}(t) = (x_1(t), x_2(t))$ for $\mathcal{X} = \{1,2\}$). For a matrix of transfer functions $\bar{X}(z)$, we will denote by $\bar{X}_{\mathcal{X},\mathcal{Y}}(z)$ the part of $\bar{X}(z)$ obtained by only conserving the rows in $\mathcal{X}$ and the columns in $\mathcal{Y}$. As an example, if $\mathcal{X} = \{1,2\}$ and $\mathcal{Y} = \{2,3\}$, we have:

$$\bar{X}_{\mathcal{X},\mathcal{Y}}(z) = \begin{pmatrix} \bar{X}_{12}(z) & \bar{X}_{13}(z) \\ \bar{X}_{22}(z) & \bar{X}_{23}(z) \end{pmatrix}$$

When $\mathcal{X}$ or $\mathcal{Y}$ are singletons, we use the following shorthand notation for $\bar{X}_{\mathcal{X},\mathcal{Y}}(z)$: $\bar{X}_{i,\mathcal{Y}}(z)$ when $\mathcal{X} = \{i\}$ and $\bar{X}_{\mathcal{X},k}(z)$ when $\mathcal{Y} = \{k\}$. Using these notations, we have that $\bar{w}_{\mathcal{X}}(t) = \bar{X}_{\mathcal{X},\mathcal{Y}}(z)\bar{x}_{\mathcal{Y}}(t)$ for any sets $\mathcal{X}$ and $\mathcal{Y}$ when $\bar{w}(t) = \bar{X}(z)\bar{x}(t)$. Note also that, when the matrix $\bar{X}(z)$ pertains to the true dynamic network, we will use the notation $\bar{X}_0(z)$. Consequently, entries of this matrix will then be logically denoted $\bar{X}_{0,ik}(z)$ and parts of this matrix $\bar{X}_{0,\mathcal{X},\mathcal{Y}}(z)$. In addition, the matrix I_n denotes the identity matrix of dimension n and $diag(a_1, ..., a_n)$ denotes the matrix of dimension $n \times n$:

$$\begin{pmatrix} a_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & a_n \end{pmatrix}$$

For a matrix A, A^T denotes the transpose of A and A^* its conjugate transpose. Finally, for a quasi-stationary signal $x(t)$ [14], $\bar{E}x(t) \triangleq lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} Ex(t)$ (E is the expectation operator).

2 Network description

In this paper, we consider the problem of identifying particular modules of a stable dynamic network. This dynamic network is made up of N_{mod} nodes that are each characterized by a scalar valued measurable signal $w_k(t)$ ($k = 1, ..., N_{mod}$). The vector $\bar{w}(t) = (w_1(t), w_2(t), ..., w_{N_{mod}}(t))^T$ obeys the following equation [7]:

$$\bar{w}(t) = \bar{G}_0(z)\ \bar{w}(t)\ + \bar{r}(t) + \underbrace{\bar{H}_0(z)\bar{e}(t)}_{=\bar{v}(t)} \tag{1}$$

$$\bar{G}_0(z) = \begin{pmatrix} 0 & G_{0,12}(z) & ... & G_{0,1N_{mod}}(z) \\ G_{0,21}(z) & 0 & ... & G_{0,2N_{mod}}(z) \\ ... & ... & ... & ... \\ G_{0,N_{mod}1}(z) & G_{0,N_{mod}2}(z) & ... & 0 \end{pmatrix} \tag{2}$$

$$\bar{H}_0(z) = diag\left(H_{0,1}(z),\ H_{0,2}(z),\ ..., H_{0,N_{mod}}(z)\right) \tag{3}$$

where all the non-zero entries in (2) are proper transfer functions and where $\bar{r}(t) = (r_1(t), r_2(t), ..., r_{N_{mod}}(t))^T$ is a vector of external excitation signals that can be freely chosen by the user, e.g., for identification purposes ($\bar{r}(t) = 0$ in normal operations). In (1), the vector $\bar{v}(t) = (v_1(t), v_2(t), ..., v_{N_{mod}}(t))^T$ represents the process noise acting on the network. This process noise is modeled as $\bar{v}(t) = \bar{H}_0(z)\bar{e}(t)$ where $\bar{H}_0(z)$ is a diagonal transfer matrix with diagonal elements $H_{0,k}(z)$ ($k = 1, ..., N_{mod}$) that are all stable, inversely stable and monic and where $\bar{e}(t) = (e_1(t), e_2(t), ..., e_{N_{mod}}(t))^T$ with $e_k(t)$ ($k = 1, ..., N_{mod}$) being zero-mean white noise signals of variance $\sigma^2_{e_k}$ ($k = 1, ..., N_{mod}$). The covariance matrix $E\bar{e}(t)\bar{e}^T(t)$ of $\bar{e}(t)$ will be denoted by[1] Σ_0. We will not impose any constraint on Σ_0 i.e., Σ_0 is neither required to be diagonal nor strictly positive definite.

Let us also make the following additional standard assumptions on the network:

Assumption 1 *Consider the network described by* (1)-(2)-(3). *We assume that $\bar{e}(t)$ is independent of $\bar{r}(t)$ and is also such that $E\bar{e}(t)\bar{e}^T(t-\tau) = 0$ for all $\tau \neq 0$. We also assume that the network is well-posed with a stable closed-loop description $\bar{S}_0(z) = (I_{N_{mod}} - \bar{G}_0(z))^{-1}$, so that the network can also be expressed as*

$$\bar{w}(t) = \bar{S}_0(z)\left(\bar{r}(t) + \bar{v}(t)\right). \tag{4}$$

The above description of the network allows for some elements $v_k(t)$ of $\bar{v}(t)$ to be identically zero. We indeed just have to choose $\sigma^2_{e_k} = 0$ and $H_{0,k}(z) = 1$ in this case. In this paper, we will suppose that we know which $v_k(t)$ are equal to zero and which $v_k(t)$ are not equal to zero.

[1] The variance $\sigma^2_{e_k}$ ($k = 1, ..., N_{mod}$) of the white noise entries e_k of $\bar{e}(t)$ are the diagonal elements of Σ_0.

For the sequel, let us denote by $\mathcal{V}$ the set of indexes k corresponding to nodes such that $\sigma^2_{e_k} \neq 0$ i.e., the set of nodes where $e_k(t) \neq 0$ and $v_k(t) \neq 0$. Using the notations introduced at the end of Section 1, $\bar{v}_{\mathcal{V}}(t)$ (resp. $\bar{e}_{\mathcal{V}}(t)$) corresponds to the non-zero elements of $\bar{v}(t)$ (resp. $\bar{e}(t)$) and we have that $\bar{v}_{\mathcal{V}}(t) = H_{0,\mathcal{V},\mathcal{V}}(z)\bar{e}_{\mathcal{V}}(t)$. In the sequel, the covariance matrix $E\bar{e}_{\mathcal{V}}(t)\bar{e}^T_{\mathcal{V}}(t)$ of $\bar{e}_{\mathcal{V}}(t)$ will be denoted by $\Sigma_{0,\mathcal{V}} \geq 0$. For the sequel, it is important to note that this covariance matrix (as any other positive semi-definite matrix) can be decomposed as $\Sigma_{0,\mathcal{V}} = \Xi_{0,\mathcal{V}}\Xi^T_{0,\mathcal{V}}$ where $\Xi_{0,\mathcal{V}}$ is a matrix with $n_{\mathcal{V}}$ rows and a number of columns equal to the rank of $\Sigma_{0,\mathcal{V}}$.

Note also that, similarly to $\bar{v}(t)$, the excitation vector $\bar{r}(t)$ that we use for identification purpose can also contain zero elements. Let us thus denote $\mathcal{R}$ the set of indexes k corresponding to nodes such that $r_k \neq 0$.

3 Full input identification approach

As already mentioned, our objective is to use prediction error identification to accurately identify the (unknown) modules in a row of the matrix $\bar{G}_0(z)$, say Row j. We will use for this purpose the Multiple Input Single Output (MISO) approach introduced in [7] i.e., the so-called *full input approach*. Before presenting this identification approach in more detail, let us introduce some concepts related to the j^{th} row of $\bar{G}_0$. This row can contain entries $G_{0,jk}$ that are known to be identically zero, entries that are both known and not equal to zero and, finally, entries that are unknown. For the sequel, we need to define two additional sets of indexes related to these types of elements: $\mathcal{K}$ is the set of indexes k corresponding to entries $G_{0,jk}(z)$ that are both known and not equal to zero, while $\mathcal{D}$ is the set of indexes k corresponding to unknown entries $G_{0,jk}(z)$.

Let us now present the MISO identification problem considered in this paper. For this purpose, let us define, using the notations introduced at the end of Section 1, the signal $y_j(t)$ as follows:

$$y_j(t) \triangleq w_j(t) - r_j(t) - \bar{G}_{0,j,\mathcal{K}}(z)\ \bar{w}_{\mathcal{K}}(t). \tag{5}$$

Since $\bar{G}_{0,j,\mathcal{K}}(z)$ is a row vector containing the known non-zero elements of the j^{th} row of $\bar{G}_0(z)$, the signal $y_j(t)$ is a computable quantity that obeys (see (1)):

$$y_j(t) = \bar{G}_{0,j,\mathcal{D}}(z)\ \bar{w}_{\mathcal{D}}(t) + H_{0,j}(z)\ e_j(t) \tag{6}$$

where $\bar{G}_{0,j,\mathcal{D}}(z)$ is a row vector of dimension $n_{\mathcal{D}}$ containing the unknown elements of the j^{th} row of $\bar{G}_0(z)$. As already mentioned, the identification approach will pertain to the identification of a model of $\bar{G}_{0,j,\mathcal{D}}(z)$ and a model of $H_{0,j}(z)$ and, as evidenced by (6), this identification will be performed using a data set $Z^N = \{y_j(t),\ \bar{w}_{\mathcal{D}}(t) \mid t = 1...N\}$. Note that (6) has the classical form of a data-generating system in MISO (prediction error) identification since the measurable output $y_j(t)$ is made up of the combination of an unknown stochastic disturbance $H_{0,j}(z)e_j(t)$ and of a contribution of the known input $\bar{w}_{\mathcal{D}}(t)$ through an unknown vector of transfer functions $\bar{G}_{0,j,\mathcal{D}}(z)$. Since $\bar{w}_{\mathcal{D}}(t)$ may be correlated with $e_j(t)$, we are moreover in a situation that is very similar to direct closed-loop identification [14]. As in direct closed-loop identification, we will here also need in many cases to require that $\bar{G}_{0,j,\mathcal{D}}(z)$ is stable [14]. We will for simplicity make that assumption in the sequel.

Let us thus suppose that we have collected on the network (1), the data set Z^N and that we have defined a model structure $\mathcal{M} = \{\bar{G}_{j,\mathcal{D}}(z,\theta),\ H_j(z,\theta) \mid \theta \in \Theta\}$ where $\bar{G}_{j,\mathcal{D}}(z,\theta)$ (resp. $H_j(z,\theta)$) is a model for $\bar{G}_{0,j,\mathcal{D}}(z)$ (resp. $H_{0,j}(z)$) and Θ is the set of all parameter vectors θ leading to a stable $\bar{G}_{j,\mathcal{D}}(z,\theta)$ and to a monic, stable and inversely stable $H_j(z,\theta)$. In the sequel, $\mathcal{M}$ is also assumed to have the following property:

Assumption 2 *The model structure* $\mathcal{M} = \{\bar{G}_{j,\mathcal{D}}(z,\theta),\ H_j(z,\theta) \mid \theta \in \Theta\}$ *has the property that there exists a unique parameter vector* $\theta_0 \in \Theta$ *such that* $\bar{G}_{j,\mathcal{D}}(z,\theta_0) = \bar{G}_{0,j,\mathcal{D}}(z)$ *and* $H_j(z,\theta_0) = H_{0,j}(z)$.

Using the data set Z^N and the model structure $\mathcal{M}$, we can then obtain an estimate $\hat{\theta}_N$ of θ_0 using the following prediction error criterion [14]:

$$\hat{\theta}_N = \arg\min_{\theta\in\Theta}\ \frac{1}{N}\sum_{t=1}^{N}\ \epsilon_j^2(t,\theta) \tag{7}$$

$$\epsilon_j(t,\theta) = H_j^{-1}(z,\theta)\left(y_j(t) - \bar{G}_{j,\mathcal{D}}(z,\theta)\bar{w}_{\mathcal{D}}(t)\right). \tag{8}$$

Using this estimate $\hat{\theta}_N$, we obtain a model $\bar{G}_{j,\mathcal{D}}(z,\hat{\theta}_N)$ of $\bar{G}_{0,j,\mathcal{D}}(z)$ and a model $H_j(z,\hat{\theta}_N)$ of $H_{0,j}(z)$.

Remark. The MISO approach presented above can also be used if we are only interested in models of part of $\bar{G}_{0,j,\mathcal{D}}(z)$ [7]. Moreover, this approach can also be repeated for each row in order to get a model of the full matrices $\bar{G}_0(z)$ and $\bar{H}_0(z)$. The data informativity conditions that we will develop in the sequel are thus also relevant for these two particular situations. ■

In the sequel, we will determine the conditions under which $\hat{\theta}_N$ is a consistent estimate of θ_0 which means that $\hat{\theta}_N$ converges to θ_0 with probability one when $N \to \infty$. The consistency of the estimate (7) can also equivalently be established by proving that θ_0 is the unique minimum of $\bar{E}\epsilon_j^2(t,\theta)$ [14]. We will prove this in two steps, i.e. we will first prove that θ_0 minimizes $\bar{E}\epsilon_j^2(t,\theta)$ and then we will determine the conditions under which this minimum is unique. The results will depend on whether there is noise present in Node j. Let us thus suppose that this is indeed the case. The simpler case $v_j(t) = e_j(t) = 0$ is treated in Appendix A of [2].

Assumption 3 *In the network* (1)*, the variance* $\sigma^2_{e_j}$ *of the noise* e_j *at Node* j *is such that* $\sigma^2_{e_j} \neq 0$ *(*$j \in \mathcal{V}$*).*

As shown in the following proposition, θ_0 is a minimum of $\bar{E}\epsilon_j^2(t,\theta)$ under Assumptions 1, 2 and 3 if we add a delay condition similar to the one required for the direct closed-loop identification method [14] (see also [7]).

Proposition 1 *Consider the stable MISO system* (6) *that is an element of a network* (1) *satisfying Assumptions 1 and 3 as well as the sets* $\mathcal{V}$ *and* $\mathcal{D}$ *defined in Sections 2 and 3, respectively. Consider the prediction error* (8) *computed based on data collected on this network and a model structure* $\mathcal{M}$ *satisfying Assumption 2. Then,* θ_0 *is a minimum of* $\bar{E}\epsilon_j^2(t,\theta)$ *if, for all* θ*, all the entries of the vector of transfer functions* $(\bar{G}_{0,j,\mathcal{D}}(z) - \bar{G}_{j,\mathcal{D}}(z,\theta))\bar{S}_{0,\mathcal{D},\mathcal{V}}(z)$ *are either zero or contain at least one delay. Moreover, all (eventual) other minimizers* θ^* *of* $\bar{E}\epsilon_j^2(t,\theta)$ *are such that* $\epsilon_j(t,\theta^*) = \epsilon_j(t,\theta_0) = e_j(t)$. ■

Proof. The proof is relatively straightforward and can be found in [2]. ■

Let us now consider the conditions under which θ_0 is the **unique** minimum of $\bar{E}\epsilon_j^2(t,\theta)$. Due to the property stated at the end of Proposition 1, this will be the case if, for each $\theta \in \Theta$ such that $\bar{E}(\epsilon_j(t,\theta) - \epsilon_j(t,\theta_0))^2 = 0$, we have $\theta = \theta_0$. Due to Assumption 2, this condition will be respected if, for each $(\bar{G}_{j,\mathcal{D}}(z,\theta),\ H_j(z,\theta)) \in \mathcal{M}$ such that $\bar{E}(\epsilon_j(t,\theta) - \epsilon_j(t,\theta_0))^2 = 0$, we have

$$\bar{G}_{0,j,\mathcal{D}}(z) - \bar{G}_{j,\mathcal{D}}(z,\theta) = 0 \quad \textit{and} \quad H_{0,j}(z) - H_j(z,\theta) = 0 \tag{9}$$

This property is generally called *data informativity* in the literature [14]. Let us formally define this notion. For this purpose, let us introduce the following notation using (8):

$$\epsilon_j(t,\theta) = \bar{W}(z,\theta)\bar{x}(t) = \left(W_y(z,\theta), \bar{W}_w(z,\theta)\right)\begin{pmatrix} y_j(t) \\ \bar{w}_{\mathcal{D}}(t) \end{pmatrix} \tag{10}$$

$$W_y(z,\theta) = H_j^{-1}(z,\theta) \ \textit{ and } \ \bar{W}_w(z,\theta) = -H_j^{-1}(z,\theta)\bar{G}_{j,\mathcal{D}}(z,\theta) \tag{11}$$

Definition 1 *Consider the data* $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ *collected on a network* (1) *satisfying Assumptions 1 and 3 and the condition in the statement of Proposition 1. Consider also a model structure* $\mathcal{M}$ *satisfying Assumption 2. Define the set:*

$$\boldsymbol{\Delta}_{\bar{\mathbf{W}}} = \{\Delta\bar{W}(z) = \bar{W}(z,\theta) - \bar{W}(z,\theta_0) \mid \theta \in \Theta\} \tag{12}$$

Then, the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ *are said to be informative wrt.* $\mathcal{M}$ *when, for all* $\Delta\bar{W}(z) \in \boldsymbol{\Delta}_{\bar{\mathbf{W}}}$*, we have:*

$$\bar{E}\left(\Delta\bar{W}(z)\bar{x}(t)\right)^2 = 0 \Longrightarrow \Delta\bar{W}(z) = \mathbf{0} \tag{13}$$

We can summarize the above discussion in the following proposition whose proof is straightforward.

Proposition 2 *Consider the stable MISO system* (6) *that is an element of a network* (1) *satisfying Assumptions 1 and 3. Consider the prediction error* (10) *computed based on data* $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ *collected on this network and a model structure* $\mathcal{M}$ *satisfying Assumption 2. Then,* θ_0 *is the unique minimum of* $\bar{E}\epsilon_j^2(t,\theta)$ *if, in addition to the delay condition in the statement of Proposition 1, the data* $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ *are informative wrt.* $\mathcal{M}$ *(see Definition 1).* ■

We have now all the elements to derive conditions for data informativity. We will see that data informativity can be obtained by adding a quasi-stationary excitation signal $r_k(t)$ at a number of nodes, but also in certain situations, using the sole excitation of the process noises $v_k(t)$ i.e., $\bar{r}(t) = 0$ (the so-called costless identification [4,5]).

4 Simple (but only sufficient) data informativity conditions

4.1 Results

We will start by deriving simple, but only sufficient data informativity conditions. For this purpose, let us first recall the data informativity condition proposed in the paper [7] where the identification method presented in Section 3 has been introduced.

Proposition 3 ([7]) *Consider the data* $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ *collected on a network* (1) *satisfying Assumptions 1 and 3. Consider also the following condition on the power spectrum matrix* $\Phi_{\bar{x}}(\omega)$ *of* $\bar{x}(t)$*:*

$$\Phi_{\bar{x}}(\omega) > 0 \textit{ at almost all } \omega. \tag{14}$$

Then, we have that, irrespectively of the complexity of $\bar{G}_{0,j,\mathcal{D}}(z)$ *and* $H_{0,j}(z)$*, the condition* (14) *ensures data informativity with respect to a model structure* $\mathcal{M}$ *satisfying Assumption 2.* ■

Proof. This result is a direct consequence of Definition 1. Indeed, the left-hand side of (13) can be rewritten as $\frac{1}{2\pi}\int_{-\pi}^{\pi}\Delta\bar{W}(e^{j\omega})\Phi_{\bar{x}}(\omega)\Delta\bar{W}^*(e^{j\omega})d\omega = 0$ and it is clear that, if (14) holds, this relation implies that $\Delta\bar{W}(z) = 0$ and this is true whatever the complexity of the vector $\Delta\bar{W}(z)$ and thus whatever $\mathcal{M}$. ■

Proposition 3 means that, if a noise and excitation pattern yields (14), this noise and excitation pattern will ensure consistent estimates of modules that can have infinite complexity. While this could be seen as an advantage in some cases, this also means that, when the to-be-identified modules are of restricted complexity, many excitation patterns that would lead to data informativity

will not be detected by this condition. This also means that to ensure (14) will require a larger number of excitations v_k and/or r_k than what is strictly necessary to ensure data informativity for a specific model structure $\mathcal{M}$ of restricted complexity.

In other words, (14) is not a necessary condition for (13) to hold and is thus conservative. Moreover, based on (14), it is difficult to determine if a given noise and excitation pattern yields data informativity or not. Moreover, as pointed out in [10], it is also difficult to use (14) in order to determine at which node(s) excitation signals r_k have to be added if the current noise and excitation pattern does not lead to data informativity.

In this section, we will reformulate (14) in such a way that the objectives presented in the previous paragraph can be achieved. Let us first start with the case where we can assume that the network satisfies $\Sigma_{0,\mathcal{V}} > 0$ i.e, the case of networks where $\bar{e}_{\mathcal{V}}(t)$ is a full rank vector of signals. With this additional assumption on the network, we can derive Propositions 4 and 5. Proposition 4 pertains to the case where $\bar{r}(t) = 0$ (costless identification) and Proposition 5 to the case where, besides the costless excitation of $v_k(t)$ ($k \in \mathcal{V}$), we also add, at the nodes $k \in \mathcal{R}$, external excitations r_k.

Proposition 4 *Consider the framework of Proposition 3 for a network* (1) *with* $\Sigma_{0,\mathcal{V}} > 0$. *Consider also the set* $\mathcal{V}$ *defined at the end of Section 2. Then, in the case where the excitation vector* $\bar{r}(t)$ *is equal to zero,* (14) *holds if the following condition is satisfied*

(i) the set $\mathcal{V}$ *describing the nodes where a disturbance* v_k *is present is such that, at (almost) all frequencies* ω, $\bar{S}_{0,\mathcal{D},\mathcal{V}\backslash\{j\}}(e^{j\omega})$ *is full row rank i.e.,* $rank(\bar{S}_{0,\mathcal{D},\mathcal{V}\backslash\{j\}}(e^{j\omega})) = n_{\mathcal{D}}$ ■

Proof. See Appendix A. ■

In the next subsection, we will show how Condition (i) of Proposition 4 can be verified in practice. Since the matrix $\bar{S}_{0,\mathcal{D},\mathcal{V}\backslash\{j\}}$ has dimension $n_{\mathcal{D}} \times (n_{\mathcal{V}} - 1)$, we can nevertheless right away note that a necessary condition for Condition (i) of Proposition 4 to hold is that $n_{\mathcal{V}} - 1 \geq n_{\mathcal{D}}$. Since $j \in \mathcal{V}$, this necessary condition will be satisfied if, in addition to the noise v_j, there are at least as many other noise sources v_k ($k \neq j$) as modules to be identified in $\bar{G}_{0,j,\mathcal{D}}$ (i.e., $n_{\mathcal{D}}$).

If Condition (i) of Proposition 4 is not satisfied, the next proposition shows that we can obtain data $\bar{x}(t)$ satisfying (14) by adding external excitation signals r_k at a certain number of nodes.

Proposition 5 *Consider the framework of Proposition 3 for a network* (1) *with* $\Sigma_{0,\mathcal{V}} > 0$. *Consider also the sets* $\mathcal{V}$ *and* $\mathcal{R}$ *defined at the end of Section 2. Then,* (14) *holds if the following conditions are both satisfied:*

(i) the set $\mathcal{R}$ *describing the nodes where an excitation signal* r_k *is present is chosen in such a way that the set of indexes* $\mathcal{Q} = \mathcal{R} \cup (\mathcal{V}\backslash\{j\})$ *has the property that, at (almost) all frequencies* ω, $rank(\bar{S}_{0,\mathcal{D},\mathcal{Q}}(e^{j\omega})) = n_{\mathcal{D}}$

(ii) the power spectrum matrix $\Phi_{\bar{r}_{\mathcal{R}}}(\omega)$ *of the excitation vector* $\bar{r}_{\mathcal{R}}(t)$ *is such that* $\Phi_{\bar{r}_{\mathcal{R}}}(\omega) > 0$ *at almost all* ω ■

Proof. See Appendix B. ■

As already mentioned, the next subsection will give a simple method to check Condition (i) of Proposition 5. Note here also that a necessary condition for Condition (i) of Proposition 5 to hold is that $n_{\mathcal{Q}} \geq n_{\mathcal{D}}$. Consequently, the sum of the number $n_{\mathcal{R}}$ of external excitation noises and of the number of noise processes v_k (with $k \in \mathcal{V}\backslash\{j\}$ and $k \notin \mathcal{R}$) must at least be equal to the number $n_{\mathcal{D}}$ of modules to be identified in $\bar{G}_{0,j,\mathcal{D}}$. It is also important to note that Condition (ii) in Proposition 5 shows that, if external excitation signals r_k are necessary to ensure data informativity via Proposition 5, these excitation signals r_k must be (filtered) white noises.

If we cannot assume $\Sigma_{0,\mathcal{V}} > 0$, then we cannot rely, as in Propositions 4 and 5, on the noises v_k with $k \in \mathcal{V}\backslash\{j\}$ to ensure (14) (and thus data informativity). This is summarized in the following proposition.

Proposition 6 *Consider the framework of Proposition 5, but for a network with* $\Sigma_{0,\mathcal{V}} \geq 0$. *Then,* (14) *holds if the following conditions are both satisfied:*

(i) the set $\mathcal{R}$ *describing the nodes where an excitation signal* r_k *is present is chosen in such a way that, at (almost) all frequencies* ω, $rank(\bar{S}_{0,\mathcal{D},\mathcal{R}}(e^{j\omega})) = n_{\mathcal{D}}$

(ii) the excitation vector $\bar{r}_{\mathcal{R}}(t)$ *satisfies Condition (ii) of Proposition 5 i.e., its power spectrum matrix* $\Phi_{\bar{r}_{\mathcal{R}}}(\omega)$ *is such that* $\Phi_{\bar{r}_{\mathcal{R}}}(\omega) > 0$ *at almost all* ω. ■

Proof. See Appendix C. ■

In order to satisfy the conditions of Proposition 6, we will thus require at least $n_{\mathcal{D}}$ external excitations r_k under the form of (filtered) white noises.

4.2 Verification of the data informativity conditions and determination of an excitation pattern

In order to use the results in the previous subsection to infer data informativity for a certain noise and excitation pattern, the main difficulty is the verification of Condition (i) in Propositions 4, 5 and 6. The submatrices of $\bar{S}_0(e^{j\omega})$ involved in these rank conditions are indeed functions of the unknown matrix $\bar{G}_0(e^{j\omega})$. This is however not a crucial problem since [12] proposes a simple approach based on the graph of the network to verify rank conditions of this type.

This result of [12] is given in Lemma 1 below. Before presenting this lemma, let us first recall some notions of

graph theory [12]. The graph of the network can be obtained by drawing a directed edge from Node k to Node l if $G_{0,lk}(z) \neq 0$. The graph is therefore a representation of the topology of the network (i.e., its interconnection structure). A path from Node k to Node $l \neq k$ is a series of adjacent edges that starts in Node k and ends in Node l. Vertex-disjoint paths are paths that do not pass through the same nodes/vertexes. Finally, in the framework of this paper (see (1)), there is always a path from Node k to Node k since r_k and v_k have a direct influence on w_k.

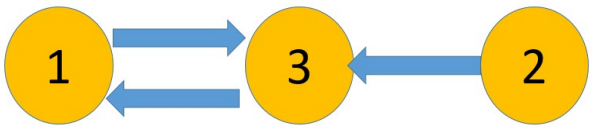

Fig. 1. Graph representation of (15). Each circle represents a node and the edges represent the structure of $\bar{G}_0(z)$.

Example 1 *Let us consider a network described by $N_{mod} = 3$ nodes and the following matrix $\bar{G}_0(z)$:*

$$\bar{G}_0(z) \triangleq \begin{pmatrix} 0 & 0 & G_{0,13}(z) \\ 0 & 0 & 0 \\ G_{0,31}(z) & G_{0,32}(z) & 0 \end{pmatrix} \qquad (15)$$

The graph of this network is represented in Figure 1. In this figure, we see that there is no path from Node 1 to Node 2 and no path from Node 3 to Node 2, but there exists, e.g., paths from Node 3 to Node 1 and from Node 2 to Node 1. Let us now e.g., choose $\mathcal{Y} = \{2, 3\}$ and $\mathcal{X} = \{1, 2\}$ and let us observe that there are two vertex-disjoint paths from the nodes in $\mathcal{Y}$ to the nodes in $\mathcal{X}$, namely the path $2 \to 2$ (since r_2/v_2 has a direct influence on w_2) and the path $3 \to 1$. If $\mathcal{Y} = \{2, 3\}$ and $\mathcal{X} = \{1, 3\}$, there is only one vertex-disjoint path from the nodes in $\mathcal{Y}$ to the nodes in $\mathcal{X}$: e.g., the path $3 \to 3$. The other paths from $\mathcal{Y}$ to $\mathcal{X}$ (i.e., the path $2 \to 3 \to 1$, the path $2 \to 3$ and the path $3 \to 1$) all contain Node 3 and are thus not vertex disjoint with the path $3 \to 3$. ■

Lemma 1 ([12]) *Consider the graph of a network and two arbitrary sets of indexes $\mathcal{X}$ and $\mathcal{Y}$ (of respective cardinality $n_{\mathcal{X}}$ and $n_{\mathcal{Y}}$). Suppose that, in the graph of the network, there are $n_{\mathcal{X}}$ vertex-disjoint paths from the nodes in $\mathcal{Y}$ to the nodes in $\mathcal{X}$. Then, for almost all $\bar{G}_0(z)$ corresponding to the topology of the network, the part $\bar{S}_{0,\mathcal{X},\mathcal{Y}}(e^{j\omega})$ of $\bar{S}_0(e^{j\omega}) = (I_{N_{mod}} - \bar{G}_0(e^{j\omega}))^{-1}$ is full row rank at almost all ω.*

Lemma 1 gives a simple graphical method to check the rank conditions in Propositions 4, 5 and 6. However, as mentioned in this lemma, the verification will be done in a *generic* manner i.e., for almost all $\bar{G}_0(e^{j\omega})$. Indeed, it cannot be excluded that, even if they have the correct topology, very specific values of $\bar{G}_0(e^{j\omega})$ can lead to situations where some submatrices of $\bar{S}_0(e^{j\omega})$ have a rank smaller than the generic rank given by Lemma 1 (see Section 4.5 for an example).

Since, for any k, there is always a path from Node k to Node k, we have also the following useful result:

Lemma 2 *Consider the graph of a network and two arbitrary sets of indexes $\mathcal{X}$ and $\mathcal{Y}$ (of respective cardinality $n_{\mathcal{X}}$ and $n_{\mathcal{Y}}$). Then, if $\mathcal{X} \subseteq \mathcal{Y}$, there is always $n_{\mathcal{X}}$ vertex-disjoint paths from the nodes in $\mathcal{Y}$ to the nodes in $\mathcal{X}$.* ■

Let us come back to our data informativity problem. Using Lemma 1, Condition (i) of Proposition 4 is generically verified if there are $n_{\mathcal{D}}$ vertex-disjoint paths from the nodes in $\mathcal{V}\backslash\{j\}$ to the nodes in $\mathcal{D}$. If this is satisfied, then we have data informativity with the sole excitation of the unknown process noises in the network. If this is not satisfied, data informativity can be obtained by adding external excitations r_k at nodes k making the set $\mathcal{Q} = \mathcal{R} \cup \mathcal{V}\backslash\{j\}$ such that there are $n_{\mathcal{D}}$ vertex disjoint paths from the nodes in $\mathcal{Q}$ to the nodes in $\mathcal{D}$ (Condition (i) of Proposition 5).

Using Lemma 2, we see that Condition (i) of Proposition 5 holds if an external excitation r_k is applied at (at least) all the nodes k such that $k \in \mathcal{D}$ and $k \notin (\mathcal{V}\backslash\{j\})$. This particular choice for $\mathcal{R}$ is equivalent to the data informativity condition proposed in Theorem 2 in [8] (when this result is particularized to the case of a diagonal $\bar{H}_0(z)$). As will be evidenced in the next subsection, note however that, in general, many other choices for $\mathcal{R}$ can lead to the desired number of vertex-disjoint paths.

Let us now briefly turn the attention to Proposition 6. When we cannot make the assumption that $\Sigma_{0,\mathcal{V}} > 0$, data informativity can be verified by checking that there are $n_{\mathcal{D}}$ vertex-disjoint paths from the nodes in $\mathcal{R}$ to the nodes in $\mathcal{D}$. When $\mathcal{V}\backslash\{j\} \neq \emptyset$, verifying this condition will require a larger number of external excitation signals r_k than in the case where we can make the assumption that $\Sigma_{0,\mathcal{V}} > 0$.

Before giving a number of illustrations of these results, let us finally note that the number of vertex-disjoint paths between two sets of nodes in the graph of a network can also be determined algorithmically [12]. This is an important feature for networks with a large number of nodes.

4.3 First illustration

We consider first a network with $\bar{G}_0(z)$ given by (15) (see Figure 1) for which we wish to identify the third row (i.e., $j = 3$). Let us assume that $\mathcal{K} = \emptyset$ and that $\mathcal{V} = \{3\}$. This means that we want to identify consistently the transfer functions $G_{0,31}(z)$, $G_{0,32}(z)$ and $H_{3,0}(z)$ and that the only unknown process noise in the network is $v_3(t)$ (Assumption 3 is thus respected). We thus deduce that $\mathcal{D} = \{1, 2\}$ and that $\Sigma_{0,\mathcal{V}} = \sigma^2_{e_3} > 0$. We can thus use Propositions 4 and 5 to check data informativity.

Since $\mathcal{V}\backslash\{j\} = \emptyset$, Condition (i) of Proposition 4 can-

not hold. Let us thus deduce an excitation pattern $\mathcal{R}$ yielding data informativity using Condition (i) of Proposition 5 (and Lemma 1). Since $n_{\mathcal{D}} = 2$ and $\mathcal{Q} = \mathcal{R}$, the cardinality $n_{\mathcal{R}}$ of $\mathcal{R}$ must be at least equal to two to get two vertex-disjoint paths from the nodes in $\mathcal{Q}$ to the nodes in $\mathcal{D}$. Using Lemma 2, an obvious choice is to choose $\mathcal{D} = \{1, 2\} \subseteq \mathcal{R}$. This leads to two possible choices $\mathcal{R} = \{1, 2\}$ and $\mathcal{R} = \{1, 2, 3\}$. However, this is not the only choices and having other choices can be important in practice. Indeed, it may be impossible to add an excitation at Node 1. In this case, we could also choose $\mathcal{R} = \{2, 3\}$ since there are also two vertex-disjoint paths from the nodes in $\mathcal{R} = \{2, 3\}$ to the nodes in $\mathcal{D} = \{1, 2\}$ (see Example 1 and Figure 1). Finally, let us also note that we cannot infer data informativity with Proposition 5 for the choice $\mathcal{R} = \{1, 3\}$ since there is no path from the nodes in $\mathcal{R}$ to Node $2 \in \mathcal{D}$ (we have thus only one vertex-disjoint path from the nodes in $\mathcal{R}$ to the nodes in $\mathcal{D}$: the path $1 \to 1$). It is obvious that this particular choice for $\mathcal{R}$ cannot lead to data informativity since, in this situation where $r_2(t) = v_2(t) = 0$, we have that $w_2(t) = 0$ and thus it will be impossible to identify $G_{0,32}(z)$.

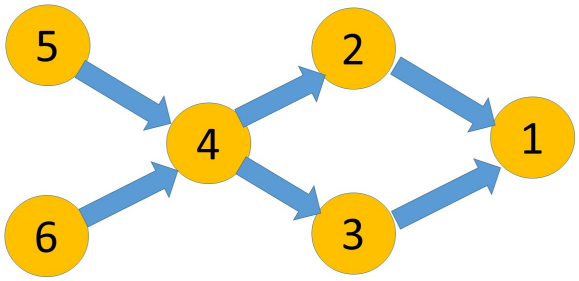

Fig. 2. Graph representation of the network in Section 4.4.

4.4 Second illustration

Let us now consider another example to stress even more the importance of having enough **vertex-disjoint** paths between the nodes in $\mathcal{Q}$ and the nodes in $\mathcal{D}$. Consider for this purpose the network with $N_{mod} = 6$ nodes described by the graph given in Figure 2. Let us assume that we wish to identify Row 1 (i.e., $j = 1$), that $\mathcal{K} = \emptyset$ and that $\mathcal{V} = \{1\}$. Consequently, $\mathcal{D} = \{2, 3\}$ and $\Sigma_{0,\mathcal{V}} = \sigma^2_{e_1} > 0$. Here also, Proposition 4 cannot be used since $\mathcal{V}\backslash\{j\} = \emptyset$. Let us thus consider adding external excitations r_k to the network. According to Proposition 5, the locations $\mathcal{R}$ of these external excitations must be chosen such that there are $n_{\mathcal{D}} = 2$ vertex-disjoint paths from the nodes in $\mathcal{R}$ to the nodes in $\mathcal{D}$. It is clear that $\mathcal{R} = \{2, 3\}$ satisfies this property since, in this case, $\mathcal{R} = \mathcal{D}$ (Lemma 2). The choice $\mathcal{R} = \{2, 6\}$ will also lead to two vertex disjoint-paths (i.e., $2 \to 2$ and $6 \to 4 \to 3$). The same can be said of $\mathcal{R} = \{2, 4\}$. However, when $\mathcal{R} = \{5, 6\}$, if there are paths between the nodes in $\mathcal{R}$ and the ones in $\mathcal{D}$, we have only one vertex-disjoint path since, to go from $\mathcal{R}$ to $\mathcal{D}$, we must always pass through Node 4. To understand why this last choice of $\mathcal{R}$ is problematic, let us notice that, for the identification of the modules $G_{0,12}(z)$ and $G_{0,13}(z)$, these two excitations r_5 and r_6 can be equivalently replaced by a unique excitation $r_4(t) = G_{0,45}(z)r_5(t) + G_{0,46}(z)r_6(t)$ at Node 4.

Let us finally stress that, both in this subsection and in the previous one, the excitation vector $r_{\mathcal{R}}(t)$ must of course also satisfy Condition (ii) of Proposition 5 to effectively lead to data informativity. This can be e.g., achieved by choosing each excitations r_k ($k \in \mathcal{R}$) as independent white noises.

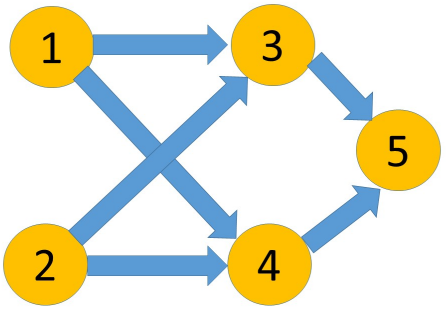

Fig. 3. Graph representation of the network in Section 4.5.

4.5 Third illustration

Let us now illustrate the issue mentioned below Lemma 1 using an example inspired by [12]. Consider for this purpose the network whose graph is represented in Figure 3. We suppose that $j = 5$, that $\mathcal{K} = \emptyset$, that $\mathcal{V} = \{1, 2, 5\}$ and that $\Sigma_{0,\mathcal{V}} > 0$. In this network, we have $w_5(t) = G_{0,53}(z)w_3(t) + G_{0,54}(z)w_4(t) + v_5(t)$. Consequently, $\mathcal{D} = \{3, 4\}$. Let us use Proposition 4 to see whether data informativity can be obtained with the sole excitation of the process noises v_1, v_2 and v_5. We observe that $\mathcal{V}\backslash\{j\} = \{1, 2\}$ and that there are two vertex-disjoint paths from $\mathcal{V}\backslash\{j\}$ to $\mathcal{D}$ ($1 \to 3$ and $2 \to 4$). Consequently, using Proposition 4 and Lemma 1, we can conclude that the noise pattern $\mathcal{V} = \{1, 2, 5\}$ yields data informativity for almost all $\bar{G}_0(z)$ having the topology described in Figure 3. To verify this result, let us derive the matrix $\bar{S}_{0,\mathcal{D},\mathcal{V}\backslash\{j\}}$ involved in Condition (i) of Proposition 4:

$$\bar{S}_{0,\mathcal{D},\mathcal{V}\backslash\{j\}}(z) = \begin{pmatrix} G_{0,31}(z) & G_{0,32}(z) \\ G_{0,41}(z) & G_{0,42}(z) \end{pmatrix}$$

This matrix has clearly a *generic* rank equal to two. However, when $G_{0,31}(z)G_{0,42}(z) = G_{0,41}(z)G_{0,32}(z)$ (which is clearly a singular case), this rank reduces to one and Condition (i) of Proposition 4 does not hold. Indeed, in this case $\bar{w}_{\mathcal{D}}(t) = \bar{S}_{0,\mathcal{D},\mathcal{V}\backslash\{j\}}(z)v_{\mathcal{V}}(t)$ has a power spectrum matrix that is rank-deficient at all ω and (14) can therefore not be respected.

Note that, even when $G_{0,31}(z)G_{0,42}(z) = G_{0,41}(z)G_{0,32}(z)$, data informativity can be inferred via Proposition 4 if, in addition, a process noise v_4 is also present at Node 4 i.e., $\mathcal{V} = \{1, 2, 4, 5\}$. Indeed, in this particular situation,

$$\bar{S}_{0,\mathcal{D},\mathcal{V}\setminus\{j\}}(z) = \begin{pmatrix} G_{0,31}(z) & G_{0,32}(z) & 0 \\ G_{0,41}(z) & G_{0,42}(z) & 1 \end{pmatrix}$$

and, in this case, $\bar{S}_{0,\mathcal{D},\mathcal{V}\setminus\{j\}}$ is full row rank for all $\bar{G}_0(z)$.

Remark. As shown in Section 4.6 of [2], the data informativity conditions developed in this section can be linked to the *network identifiability* condition of [17].

5 Necessary and sufficient condition for data informativity

5.1 Results

As mentioned in the previous section, the data informativity conditions derived in Section 4 are conservative when the model structure $\mathcal{M}$ has a restricted complexity. In this section, we will derive a necessary and sufficient condition for data informativity for a given model structure (of restricted complexity).

For this purpose, we will need to distinguish the multisine and filtered white noise contributions in the excitation vector $\bar{r}(t)$ in more details. For this purpose, let us introduce the set of indexes $\mathcal{R}_s$ as the set of indexes k such that r_k contains a multisine contribution and the set of indexes $\mathcal{R}_n$ as the set of indexes k such that r_k contains a filtered white noise contribution. The vector $\bar{r}_{\mathcal{R}_s}(t)$ corresponding to $\mathcal{R}_s$ is thus a vector for which each entry is a multisine while the vector $\bar{r}_{\mathcal{R}_n}(t)$ corresponding to $\mathcal{R}_n$ can always be expressed as:

$$\bar{r}_{\mathcal{R}_n}(t) = \bar{F}(z)\bar{q}(t) \tag{16}$$

with $\bar{F}(z)$ a known matrix of transfer functions of dimension $n_{\mathcal{R}_n} \times n_q$ and a vector $\bar{q}(t)$ of dimension n_q such that $\Phi_{\bar{q}}(\omega) = I_{n_q}$ (i.e., $\bar{q}(t)$ is a vector of independent white noises of variance 1). As an example, if $N_{mod} = 3$ and $\bar{r}(t) = (0, cos(0.1t), cos(0.2t) + m(t))^T$ with $m(t)$ a filtered white noise, then $\mathcal{R} = \{2, 3\}$, $\mathcal{R}_s = \{2, 3\}$ and $\mathcal{R}_n = \{3\}$. Moreover, $\bar{r}_{\mathcal{R}_s}(t) = (cos(0.1t), cos(0.2t))^T$ and $\bar{r}_{\mathcal{R}_n}(t) = m(t)$ and the filtered white noise $m(t)$ can always be expressed as in (16) (with $n_q = 1$). Note that $\mathcal{R} = \mathcal{R}_s \cup \mathcal{R}_n$.

We will also need to rewrite the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ (see Definition 1) in an appropriate manner. Recall that $y_j(t) = w_j(t) - r_j(t) - \bar{G}_{0,j,\mathcal{K}}(z)\bar{w}_{\mathcal{K}}(t)$ (see (5)). Using (16) and the fact that, for any set of indexes $\mathcal{X}$, $w_{\mathcal{X}}(t) = \bar{S}_{0,\mathcal{X},\mathcal{R}_s}(z)\bar{r}_{\mathcal{R}_s}(t) + \bar{S}_{0,\mathcal{X},\mathcal{R}_n}(z)\bar{r}_{\mathcal{R}_n}(t) + \bar{S}_{0,\mathcal{X},\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t)$, it is clear that we have that:

$$\bar{x}(t) = \bar{T}_{0,\mathcal{V}}(z)\ \bar{v}_{\mathcal{V}}(t) + \bar{X}_{0,\mathcal{R}_n}(z)\ \bar{F}(z)\ \bar{q}(t) + \bar{d}(t) \tag{17}$$

$$\bar{X}_{0,\mathcal{R}_n}(z) = (\bar{T}_{0,\mathcal{R}_n}(z) - M_n)$$

$$\bar{d}(t) = (\bar{T}_{0,\mathcal{R}_s}(z) - M_s)\bar{r}_{\mathcal{R}_s}(t)$$

where M_n and M_s depend on the excitation signal r_j at Node j. When $j \in \mathcal{R}_s$, M_s is a matrix of dimension $(1 + n_{\mathcal{D}}) \times n_{\mathcal{R}_s}$ such that $M_s\ \bar{r}_{\mathcal{R}_s}(t) = (r_j^s(t), 0, ..., 0)^T$ with $r_j^s(t)$ the multisine contribution in $r_j(t)$. When $j \in \mathcal{R}_n$, M_n is a matrix of dimension $(1 + n_{\mathcal{D}}) \times n_{\mathcal{R}_n}$ such that $M_n\ \bar{r}_{\mathcal{R}_n}(t) = (r_j^n(t), 0, ..., 0)^T$ with $r_j^n(t)$ the filtered white noise contribution in $r_j(t)$. When $j \notin \mathcal{R}_s$ (resp. $j \notin \mathcal{R}_n$), we have $M_s = 0$ (resp. $M_n = 0$). In (17), we have also that, for any set $\mathcal{X}$, $\bar{T}_{0,\mathcal{X}}(z)$ is a matrix of transfer functions of dimension $(1 + n_{\mathcal{D}}) \times n_{\mathcal{X}}$ given by

$$\bar{T}_{0,\mathcal{X}}(z) = \begin{pmatrix} \bar{S}_{0,j,\mathcal{X}}(z) - \bar{G}_{0,j,\mathcal{K}}(z)\bar{S}_{0,\mathcal{K},\mathcal{X}}(z) \\ \bar{S}_{0,\mathcal{D},\mathcal{X}}(z) \end{pmatrix} \tag{18}$$

Let us also finally observe that, in (17), $\bar{d}(t)$ is a vector of dimension $1 + n_{\mathcal{D}}$ for which each entry is a multisine.

The notations introduced in (17) seem complex. However, for the very classical case where $\mathcal{K} = \emptyset$ and $r_j = 0$, they are much simplified [2] since $M_s = M_n = 0$ and (18) reduces to the following submatrix of $\bar{S}_0(z)$:

$$\bar{T}_{0,\mathcal{X}}(e^{j\omega}) = \begin{pmatrix} \bar{S}_{0,j,\mathcal{X}}(e^{j\omega}) \\ \bar{S}_{0,\mathcal{D},\mathcal{X}}(e^{j\omega}) \end{pmatrix} \tag{19}$$

We have now all the elements to derive the following proposition that gives a necessary and sufficient condition for data informativity. Like in Section 4, we will have different results in the cases where we can assume $\Sigma_{0,\mathcal{V}} > 0$ and the cases where we cannot. Proposition 7 gives the result for the case $\Sigma_{0,\mathcal{V}} > 0$. In this proposition, we observe that data informativity can be obtained by adding a quasi-stationary excitation signal $r_k(t)$ at a number of nodes, but also in certain situations, using the sole excitation of the process noises $v_k(t)$ i.e., $\bar{r}(t) = 0$.

Proposition 7 *Consider the network* (1) *described in Section 2 and satisfying Assumptions 1 and 3 and for which we can also assume that* $\Sigma_{0,\mathcal{V}} > 0$. *Consider also the sets* $\mathcal{V}$ *and* $\mathcal{R}$ *defined at the end of Section 2. Consider finally Definition 1 and observe that we have expression* (17) *for* $\bar{x}(t)$. *Then, in the case where the excitation vector* $\bar{r}(t)$ *is equal to zero, the data* $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ *are informative wrt.* $\mathcal{M}$ *if and only if, for all* $\Delta\bar{W}(z) \in \mathbf{\Delta}_{\bar{\mathbf{W}}}$,

$$\Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z) \Longrightarrow \Delta\bar{W}(z) = 0 \tag{20}$$

In the case where $\bar{r}(t) \neq 0$, *the data* $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}(t))^T$ *are informative wrt.* $\mathcal{M}$ *if and only if, for all* $\Delta\bar{W}(z) \in \mathbf{\Delta}_{\bar{\mathbf{W}}}$,

[2] This is the main reason why (5) is used to derive (17) instead of (6).

$$\begin{cases} \Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z) = 0 \\ \Delta\bar{W}(z)\ \bar{X}_{0,\mathcal{R}_n}(z)\ \bar{F}(z) = 0 \end{cases} \Longrightarrow \Delta\bar{W}(z) = 0 \qquad (21)$$
$$\bar{E}\left(\Delta\bar{W}(z)\ \bar{d}(t)\right)^2 = 0$$

If $\bar{F}(e^{j\omega})$ is full row rank at almost all ω, the second line of the left hand side of (21) can be equivalently replaced by $\Delta\bar{W}(z)\ \bar{X}_{0,\mathcal{R}_n}(z) = 0$. ■

Proof. See Appendix D. ■

In Proposition 7, we suppose, for the sake of generality, that both $\mathcal{R}_s \neq \emptyset$ and $\mathcal{R}_n \neq \emptyset$. In the classical case where one of these two sets is empty (i.e., $\bar{r}_{\mathcal{R}}$ contains only multisines or only filtered white noises), we have to remove the corresponding term in the left-hand side of (21) (e.g., we have to remove the second line of the left-hand side of (21) when $\mathcal{R}_n = \emptyset$).

Let us now consider the case of networks where we cannot suppose that $\Sigma_{0,\mathcal{V}} > 0$.

Proposition 8 *Consider the framework of Proposition 7, but with $\Sigma_{0,\mathcal{V}} \geq 0$. In this case, when $\bar{r}(t) = 0$, the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are informative wrt. $\mathcal{M}$ if and only if, for all $\Delta\bar{W}(z) \in \boldsymbol{\Delta}_{\bar{\mathbf{W}}}$,*

$$\Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z)\ \ \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\ \Xi_{0,\mathcal{V}} = 0 \Longrightarrow \Delta\bar{W}(z) = 0$$

where $\Xi_{0,\mathcal{V}}$ is such that $\Sigma_{0,\mathcal{V}} = \Xi_{0,\mathcal{V}}\Xi_{0,\mathcal{V}}^T$ (see the end of Section 2). In the case where $\bar{r}(t) \neq 0$, the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are informative wrt. $\mathcal{M}$ if and only if, for all $\Delta\bar{W}(z) \in \boldsymbol{\Delta}_{\bar{\mathbf{W}}}$,

$$\begin{cases} \Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z)\ \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\ \Xi_{0,\mathcal{V}} = 0 \\ \Delta\bar{W}(z)\ \bar{X}_{0,\mathcal{R}_n}(z)\ \bar{F}(z) = 0 \end{cases} \Longrightarrow \Delta\bar{W}(z) = 0$$
$$\bar{E}\left(\Delta\bar{W}(z)\ \bar{d}(t)\right)^2 = 0$$

■

Proof. See Appendix E. ■

As we will see in the next subsection, the framework that we recently developed in [5] allows to verify whether the data informativity conditions of Proposition 7 (and of Proposition 8) are satisfied in a given situation. By *given situation*, we mean a given network configuration, a given $\mathcal{M}$ satisfying Assumption 2, a given $\mathcal{V}$ and given $\bar{r}_{\mathcal{R}_s}(t)$ and $\bar{r}_{\mathcal{R}_n}(t)$. Like in the previous section, the data informativity conditions derived in this section for dynamic network identification are a function of the unknown true transfer matrix $\bar{G}_0(z)$ (and also of $\bar{H}_0(z)$ and $\Sigma_{0,\mathcal{V}}$ in the case where we cannot assume that $\Sigma_{0,\mathcal{V}} > 0$). This is an important difference with the previous results on data informativity developed for open-loop and closed-loop identification (see e.g., [14,9,6,5]). Similarly as in Section 4.2, this drawback will be circumvented by verifying that the conditions of Propositions 7 and 8 are satisfied in a generic manner. In this case, by replacing $\bar{G}_0(z)$, $\bar{H}_0(z)$ and $\Sigma_{0,\mathcal{V}}$ by any full-order models of these unknown quantities (see later).

Remark. In the case where we cannot assume that $e_{\mathcal{V}}$ is a full rank vector of signals (Proposition 8), assuming to know a model of $\Sigma_{0,\mathcal{V}} \geq 0$ can be seen as a strong assumption. If such a model is unknown, we can then decide to disregard the contribution of the process noises v_k ($k \in \mathcal{V}$) in the informativity of the data (like in Proposition 6) and use the following data informativity condition instead of the one given in Proposition 8:

$$\begin{cases} \Delta\bar{W}(z)\ \bar{X}_{0,\mathcal{R}_n}(z)\ \bar{F}(z) = 0 \\ \bar{E}\left(\Delta\bar{W}(z)\ \bar{d}(t)\right)^2 = 0 \end{cases} \Longrightarrow \Delta\bar{W}(z) = 0$$

In this case, the verification of the data informativity condition only requires a full-order model of $\bar{G}_0(z)$ (such as in the case of condition (21)) ■

5.2 Using the necessary and sufficient conditions of Propositions 7 and 8

As mentioned above, we can verify the rather complex necessary and sufficient conditions in Propositions 7 and 8 by following an approach similar to the one in [5]. We will first present the procedure to verify the conditions in Proposition 7.

Let us first recall two straightforward technical results that are used in [5]. The first technical result is the fact that any polynomial matrix $N(z)$ (i.e., a matrix whose entries are polynomials in z^{-1}) can be factorized as $N(z) = \mathcal{A}_N Z_N(z)$ with $\mathcal{A}_N$ a matrix of coefficients and $Z_N(z)$ a matrix having the following properties. If each entry of Column k of $N(z)$ is equal to zero, then each entry of Column k of $Z_N(z)$ is also equal to zero. The remaining columns of $Z_N(z)$ form a block-diagonal matrix for which each diagonal block is a column vector whose entries are equal to z^{-m} with a different integer $m \geq 0$ for each entry of the same column vector. Let us give an example of this first technical result for a $N(z)$ of dimension 2×3:

$$\underbrace{\begin{pmatrix} 2z^{-2} & 0 & 5+3z^{-2} \\ 0 & 0 & 4z^{-2} \end{pmatrix}}_{=N(z)} = \underbrace{\begin{pmatrix} 2 & 5 & 3 \\ 0 & 0 & 4 \end{pmatrix}}_{=\mathcal{A}_N} \underbrace{\begin{pmatrix} z^{-2} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & z^{-2} \end{pmatrix}}_{=Z_N(z)}$$

The second technical result uses Euler formula i.e., $\Lambda\ cos(\omega_1 t + \Psi) = \frac{1}{2}\left(\Lambda e^{j\Psi} e^{j\omega_1 t} + \Lambda e^{-j\Psi} e^{-j\omega_1 t}\right)$. Using this formula, any multisine vector $\bar{s}(t)$ whose k^{th} element $s_k(t)$ can be expressed as $s_k(t) = \sum_{l=1}^{n} \Lambda_{kl}\ cos(\omega_l t + \Psi_{kl})$ can be factorized as $\bar{s}(t) = \mathcal{B}\ \bar{\phi}(t)$ with $\mathcal{B}$ a time-independent and complex matrix (that is a function

of the amplitudes Λ_{kl} and the phase shifts Ψ_{kl}) and $\bar{\phi}(t) = \frac{1}{2}(e^{j\omega_1 t}, e^{-j\omega_1 t}, e^{j\omega_2 t}, ..., e^{-j\omega_n t})^T$.

As shown in Lemma 6 in [5], for the classical model structures $\mathcal{M}$ that are used in prediction error identification, we can always find a left factorization of $\Delta\bar{W}(z)$ i.e., $\Delta\bar{W}(z) = Q^{-1}(z)\Upsilon(z)$ with $\Upsilon(z)$ a row polynomial vector and $Q(z)$ a monic polynomial. Let us apply the first technical result on the row polynomial vector $\Upsilon(z)$:

$$\Upsilon(z) = \delta^T Z_\Upsilon(z) \tag{22}$$

with δ a vector of coefficients that are functions of θ and θ_0 (see Definition 1). Let us illustrate this in the case where the model structure $\mathcal{M}$ satisfying Assumption 2 is given by $H_j(z,\theta) = 1/(1+az^{-1})$ and:

$$\bar{G}_{j,\mathcal{D}}(z,\theta) = \left(\frac{b_1 z^{-1}}{1+az^{-1}} \;\; \frac{b_2 z^{-1}}{1+az^{-1}}\right)$$

We have thus $\theta = (a, b_1, b_2)^T$. Using (11), the left factorization $Q^{-1}(z)\Upsilon(z)$ of $\Delta\bar{W}(z)$ is here given by $Q(z) = 1$ and $\Upsilon(z) = \left(\delta a z^{-1},\ \delta b_1 z^{-1},\ \delta b_2 z^{-1}\right)$ with $\delta a = a - a_0$, $\delta b_1 = b_{1,0} - b_1$ and $\delta b_2 = b_{2,0} - b_2$ ($\theta_0 = (a_0, b_{1,0}, b_{2,0})^T$). Defining $\delta = (\delta a,\ \delta b_1,\ \delta b_2)^T$, we have thus:

$$\Upsilon(z) = \delta^T \underbrace{\begin{pmatrix} z^{-1} & 0 & 0 \\ 0 & z^{-1} & 0 \\ 0 & 0 & z^{-1} \end{pmatrix}}_{=Z_\Upsilon(z)} \tag{23}$$

Let us now consider Conditions (20) and (21) in Proposition 7. Using the matrix $Z_\Upsilon(z)$ derived from the expression of $\Delta\bar{W}(z)$, we derive a right factorization $N_1(z)V_1^{-1}(z)$ of $Z_\Upsilon(z)\bar{T}_{0,\mathcal{V}}(z)$ with $N_1(z)$ and $V_1(z)$ polynomial matrices. Using the first technical result presented in the beginning of this subsection, we subsequently derive the following factorization of $N_1(z)$:

$$N_1(z) = \mathcal{A}_{N_1} Z_{N_1}(z) \tag{24}$$

If $\mathcal{R}_n \neq \emptyset$, we then consider a similar right factorization $N_2(z)V_2^{-1}(z)$ of $Z_\Upsilon(z)\bar{X}_{0,\mathcal{R}_n}(z)\bar{F}(z)$ and we apply the first technical result on the polynomial matrix $N_2(z)$:

$$N_2(z) = \mathcal{A}_{N_2} Z_{N_2}(z) \tag{25}$$

Finally, if $\mathcal{R}_s \neq \emptyset$, we apply the second technical result (presented in the beginning of this subsection) on the multisine vector $Z_\Upsilon(z)\bar{d}(t)$:

$$Z_\Upsilon(z)\bar{d}(t) = \mathcal{B}_{\bar{d}}\ \bar{\phi}(t) \tag{26}$$

Using (22) and (24), Condition (20) can be equivalently rewritten as $\delta^T \mathcal{A}_{N_1} Z_{N_1}(z) = 0 \Longrightarrow \delta^T Z_\Upsilon(z) = 0$ which, due to the characteristic of $Z_{N_1}(z)$ and $Z_\Upsilon(z)$, is also equivalent to [6,5]:

$$\delta^T \mathcal{A}_{N_1} = 0 \Longrightarrow \delta = 0 \tag{27}$$

Using (22), (24), (25) and (26), Condition (21) can be equivalently rewritten as

$$\begin{cases} \delta^T \mathcal{A}_{N_1} Z_{N_1}(z) = 0 \\ \delta^T \mathcal{A}_{N_2} Z_{N_2}(z) = 0 \\ \bar{E}\left(\delta^T\ \mathcal{B}_{\bar{d}}\ \bar{\phi}(t)\right)^2 = 0 \end{cases} \Longrightarrow \delta^T Z_\Upsilon(z) = 0 \tag{28}$$

which, due to the characteristic of $Z_\Upsilon(z), Z_{N_1}(z), Z_{N_2}(z)$ and $\bar{\phi}(t)$, is also equivalent to [6,5]:

$$\begin{cases} \delta^T \mathcal{A}_{N_1} = 0 \\ \delta^T \mathcal{A}_{N_2} = 0 \\ \delta^T\ \mathcal{B}_{\bar{d}} = 0 \end{cases} \Longrightarrow \delta = 0 \tag{29}$$

To sum up, if we define $\mathbf{\Delta}_\delta = \{\delta \mid Q^{-1}(z)\left(\delta^T Z_\Upsilon(z)\right) \in \mathbf{\Delta}_{\bar{\mathbf{W}}}\}$, we can reformulate Proposition 7 as follows. When $\bar{r}(t) = 0$, $\bar{x}(t)$ is informative wrt. $\mathcal{M}$ if and only if (27) holds for all $\delta \in \mathbf{\Delta}_\delta$. When $\bar{r}(t) \neq 0$, $\bar{x}(t)$ is informative wrt. $\mathcal{M}$ if and only if (29) holds for all $\delta \in \mathbf{\Delta}_\delta$. This leads to the following proposition that allows one to verify if data informativity is obtained in a certain situation (i.e., a given network configuration, a given $\mathcal{M}$, a given $\mathcal{V}$ and given $\bar{r}_{\mathcal{R}_s}(t)$ and $\bar{r}_{\mathcal{R}_n}(t)$).

Proposition 9 *Consider the framework of Proposition 7. Then, in the case where $\bar{r}(t) = 0$, we can construct the matrix $\mathcal{A}_{N_1}$ as indicated above this proposition and we can verify whether the data $\bar{x}(t)$ are informative wrt. $\mathcal{M}$ by checking that the matrix $\mathcal{A}_{N_1}$ is a full row rank matrix. In the case where $\bar{r}(t) \neq 0$, besides $\mathcal{A}_{N_1}$, we also construct, as indicated above this proposition, the matrices $\mathcal{A}_{N_2}$ (if $\mathcal{R}_n \neq \emptyset$) and $\mathcal{B}_{\bar{d}}$ (if $\mathcal{R}_s \neq \emptyset$) and we can verify that the data $\bar{x}(t)$ are informative wrt. $\mathcal{M}$ by checking that the matrix $\mathcal{C}$ is a full row rank matrix. The matrix $\mathcal{C} = (\mathcal{A}_{N_1}\ \mathcal{A}_{N_2}\ \mathcal{B}_{\bar{d}})$ if both $\mathcal{R}_n \neq \emptyset$ and $\mathcal{R}_s \neq \emptyset$ while $\mathcal{C} = (\mathcal{A}_{N_1}\ \mathcal{A}_{N_2})$ if $\mathcal{R}_s = \emptyset$ and $\mathcal{C} = (\mathcal{A}_{N_1}\ \mathcal{B}_{\bar{d}})$ if $\mathcal{R}_n = \emptyset$.* ■

Proof. This proposition is a straightforward consequence of the paragraph preceding the proposition [3]. ■

Note that the number of columns in $\mathcal{B}_{\bar{d}}$ is related to the number of entries in $\bar{r}_{\mathcal{R}_s}$ and, more particularly, to

[3] The conditions in Proposition 9 are (generally) only sufficient conditions for (27) and (29) to hold since $\mathbf{\Delta}_\delta$ is (generally) not equal to the whole vectorial space [5]. However, the introduced conservatism is generally much lower than with the sufficient conditions of Section 4 as will be shown in the examples of Section 5.3 and 5.4 [5].

the number of sinusoids at different frequencies in $\bar{r}_{\mathcal{R}_s}$, while the number of columns in $\mathcal{A}_{N_2}$ is related to the complexity of $\bar{F}(z)$ and to how large n_q is in (16). Consequently, if we face a situation where we do not have data informativity, we can easily determine what measures have to be taken in order to increase the informativity of the data (see Section 8 of [5] for more details).

Since the data informativity condition of Proposition 7 is a function of the unknown matrix $\bar{G}_0(z)$, the same can be said for the matrices $\mathcal{A}_{N_1}$ and $\mathcal{C}$ in Proposition 9. As mentioned in the previous section, we can nevertheless check the data informativity in a *generic* manner by replacing $\bar{G}_0(z)$ by any full-order model of this matrix. The models of the entries of $\bar{G}_0(z)$ do not need to be accurate, but they should be of the **correct** order. Consequently, to verify the data informativity using Proposition 9, we need to know the orders of all the entries of $\bar{G}_0(z)$ and not only the entries in its j^{th} row (see Assumption 2).

If we cannot assume that $\Sigma_{0,\mathcal{V}} > 0$, we can use Proposition 8 instead of Proposition 7 (see Section 5.1). Proposition 9 remains valid in this case if we define $\mathcal{A}_{N_1}$ based on a right factorization $N_1(z)V_1^{-1}(z)$ of $Z_\Upsilon(z)\bar{T}_{0,\mathcal{V}}(z)\bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\Xi_{0,\mathcal{V}}$. As proposed in the remark at the end of Section 5.1, if a model of $\Sigma_{0,\mathcal{V}}$ is not available, we can neglect the contribution of the process noises to data informativity and define $\mathcal{C}$ uniquely based on $\mathcal{A}_{N_2}$ (if $\mathcal{R}_n \neq \emptyset$) and $\mathcal{B}_{\bar{d}}$ (if $\mathcal{R}_s \neq \emptyset$).

Let us now give two illustrations of the advantages of the results in this section with respect to the results in Section 4.

5.3 First illustration

We consider the same network as in Section 4.3 i.e., a network with $\bar{G}_0(z)$ given by (15) and whose graph is given in Figure 1. However, let us now define more precisely the non-zero transfer functions in $\bar{G}_0(z)$. We have $G_{0,31}(z) = \frac{0.173z^{-1}}{A_0(z)}$ and $G_{0,32}(z) = \frac{0.259z^{-1}}{A_0(z)}$, $G_{0,13}(z) = 0.3\,G_{0,32}(z)$ ($A_0(z) = 1-0.741z^{-1}$). Moreover, let us also assume that $\bar{v}(t) = diag(A_0^{-1}(z), A_0^{-1}(z), A_0^{-1}(z))\bar{e}(t)$ with $\bar{e}(t)$ a white noise vector of covariance matrix $\Sigma_0 = diag(0,0,0.1)$. Since $\mathcal{V} = \{3\}$, we have that $\Sigma_{0,\mathcal{V}} = 0.1 > 0$ and we are thus in the framework of Proposition 7.

Like in Section 4.3, we take $j = 3$ and $\mathcal{K} = \emptyset$ i.e., we want to identify consistently the transfer functions $G_{0,31}(z)$, $G_{0,32}(z)$ and $H_{3,0}(z)$ ($\mathcal{D} = \{1,2\}$).

Let us first verify whether the sole excitation of $v_3(t)$ could lead to data informativity (enabling costless identification). For this purpose, we must construct the matrix $\mathcal{A}_{N_1}$ and verify whether it is full row rank.

We observe that the model structure $\mathcal{M}$ taken as example in Section 5.2 is a full-order model structure $\mathcal{M}$ for the to-be-identified transfer functions $G_{0,31}(z)$, $G_{0,32}(z)$ and $H_{3,0}(z)$. For this model structure $\mathcal{M}$, we have thus that $Z_\Upsilon(z)$ is the one defined in (23).

Since $\mathcal{K} = \emptyset$, we have, using (19), that $\bar{T}_{0,\mathcal{V}}(z) = (\bar{S}_{0,33}(z),\ \bar{S}_{0,13}(z),\ \bar{S}_{0,23}(z))^T$ i.e. a permutation of the third column of $\bar{S}_0(z)$. The matrix $Z_\Upsilon(z)\bar{T}_{0,\mathcal{V}}(z)$ can be factorized as $N_1(z)V_1^{-1}(z)$ with:

$$N_1(z) = \begin{pmatrix} z^{-1} - 1.482\ z^{-2} + 0.5488\ z^{-3} \\ 0.0778\ z^{-2} - 0.0576\ z^{-3} \\ 0 \end{pmatrix}$$

and $V_1(z) = 1 - 1.482\,z^{-1} + 0.5354\,z^{-2}$. The polynomial matrix $N_1(z)$ can be factorized as:

$$N_1(z) = \mathcal{A}_{N_1} Z_{N_1}(z) = \underbrace{\begin{pmatrix} 1 & -1.482 & 0.5488 \\ 0 & 0.0778 & -0.0576 \\ 0 & 0 & 0 \end{pmatrix}}_{=\mathcal{A}_{N_1}} \begin{pmatrix} z^{-1} \\ z^{-2} \\ z^{-3} \end{pmatrix} \tag{30}$$

Since the matrix $\mathcal{A}_{N_1}$ is not full row rank, we cannot infer data informativity with the sole excitation of $v_3(t)$ (see Proposition 9). This is an expected result since $w_2(t)$ is equal to zero when $r_2(t) = v_2(t) = 0$, and it will therefore be impossible to identify $G_{0,32}(z)$. It makes thus sense to improve the data informativity by adding an external excitation r_2 to Node 2.

Let us try this with a very simple excitation signal $r_2(t)$ i.e., $r_2(t) = cos(\omega_0 t)$ with an arbitrary frequency ω_0, say $\omega_0 = 0.1$. This means that $\mathcal{R} = \mathcal{R}_s = \{2\}$. In order to verify whether we have data informativity with this excitation pattern, we need to construct the matrix $\mathcal{B}_{\bar{d}}$ using the multisine vector $Z_\Upsilon(z)\bar{d}(t)$. For this purpose, we observe that $\bar{d}(t) = \bar{T}_{0,\{2\}}(z)\ r_2(t)$ ($M_s = 0$ in (17) since $r_{j=3}(t) = 0$) and that, using (19), $\bar{T}_{0,\{2\}}(z) = (\bar{S}_{0,32}(z),\ \bar{S}_{0,12}(z),\ \bar{S}_{0,22}(z))^T$. Using the fact that $Z_\Upsilon(e^{j\omega_0})\bar{T}_{0,\{2\}}(e^{j\omega_0}) = (0.88 - 0.62\ j,\ 0.17 - 0.26\ j,\ 0.995 - 0.099\ j\ )^T$ for $\omega_0 = 0.1$, we obtain the following factorization (26) of $Z_\Upsilon(z)\bar{d}(t)$

$$Z_\Upsilon(z)\bar{d}(t) = \underbrace{\begin{pmatrix} 0.88 - 0.62j & 0.88 + 0.62j \\ 0.17 - 0.26j & 0.17 + 0.26j \\ 0.995 - 0.099j & 0.995 + 0.099j \end{pmatrix}}_{=\mathcal{B}_{\bar{d}}} \begin{pmatrix} 0.5\ e^{j\omega_0 t} \\ 0.5\ e^{-j\omega_0 t} \end{pmatrix}$$

Since $\mathcal{R}_n = \emptyset$, the matrix $\mathcal{C}$ in Proposition 9 is given by $\mathcal{C} = (\mathcal{A}_{N_1}\ \mathcal{B}_{\bar{d}})$. Since the rank of this matrix $\mathcal{C}$ is equal to three, $\mathcal{C}$ is full row rank and we can thus conclude that we will get a consistent estimate of $G_{0,31}(z)$, $G_{0,32}(z)$ and $H_{3,0}(z)$ using an excitation $r_2(t) = cos(0.1t)$ and the noise disturbance $v_3(t)$. The data informativity property can be confirmed by performing an identification with a large N in these conditions (i.e. $\bar{r}(t) = (0,\ cos(0.1t), 0)^T$ and $\mathcal{V} = \{3\}$) and by observing that $\hat{\theta}_N$ is indeed a very close estimate of θ_0. It is also clear that, due to the above result, data informativity will also be obtained if r_2 is made up of more than one sinusoid and if r_1 and r_3 are also multisines (at other frequencies than the sinusoids in r_2). This indeed only add more columns to $\mathcal{B}_{\bar{d}}$. Using

the procedure of Section 5.2, we can also prove that we have data informativity when r_2 is filtered white noise.

The above results show the important advantage of Proposition 9 upon the conservative results of Section 4. Indeed, as shown in Section 4.3, with these conservative results, we could only prove data informativity if the cardinality of $\mathcal{R}$ is larger or equal to two and if the excitations are filtered white noises.

In the analysis above, we have used the true $\bar{G}_0(z)$ to deduce $\mathcal{C}$. We have nevertheless also applied the above procedure for different full-order models of $\bar{G}_0(z)$ and the rank of the matrix $\mathcal{C}$ remained equal to three for all these models. The tested models have all the following form $G_{31}(z) = \frac{b_{31}z^{-1}}{1+az^{-1}}$, $G_{32}(z) = \frac{b_{32}z^{-1}}{1+az^{-1}}$, $G_{13}(z) = \frac{b_{13}z^{-1}}{1+az^{-1}}$ with a parameter vector $(b_{31}, b_{32}, b_{13}, a)^T \in \mathbf{R}^4$ generated randomly.

5.4 Second illustration

Let us consider a network (1) with $N_{mod} = 3$ nodes where $\bar{G}_0(z)$ is given by:

$$\bar{G}_0(z) \triangleq \begin{pmatrix} 0 & 0 & 0.5z^{-1} \\ 0.5z^{-1} & 0 & 0.5z^{-1} \\ 0.5z^{-1} & 0.5z^{-1} & 0 \end{pmatrix} \tag{31}$$

and where $\bar{H}_0(z) = I_3$ and $\Sigma_0 = diag(0, 0.1, 0)$. Since $\mathcal{V} = \{2\}$, we have that $\Sigma_{0,\mathcal{V}} = 0.1 > 0$ and we are thus here also in the framework of Proposition 7. The graph of this network is given in Figure 4. We here wish to identify Row 2 of $\bar{G}_0(z)$ (i.e., $j = 2$) and we suppose here also that there is no known element in the second row of $\bar{G}_0(z)$ ($\mathcal{K} = \emptyset$). This means that the identification procedure of Section 3 pertains to the consistent identification of $G_{0,21}(z)$ and $G_{0,23}(z)$ (since $H_{0,j} = 1$). We have thus $\mathcal{D} = \{1, 3\}$. Observe also that Assumption 3 is respected in this setup.

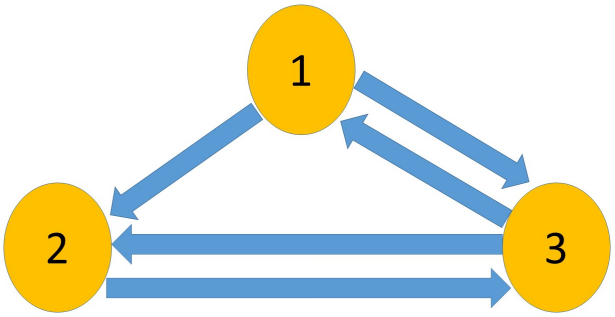

Fig. 4. Graph representation of (31).

For this network, we will prove, via Proposition 9, that a consistent estimate of the transfer functions $G_{0,21}(z)$, $G_{0,23}(z)$ can be obtained via the (costless) excitation of the noise $v_2(t) = e_2(t)$. As said in Proposition 9, we must thus construct the matrix $\mathcal{A}_{N_1}$ and verify whether it is full row rank. For this purpose, following the procedure in Section 5.2, we first factorize $\Delta\bar{W}(z)$. Using (31), a model structure $\mathcal{M}$ satisfying Assumption 2 is $\mathcal{M} = \{\bar{G}_{2,\mathcal{D}}(z, \theta) = \begin{pmatrix}\theta_1 z^{-1} & \theta_2 z^{-1}\end{pmatrix},\ H_2(z, \theta) = 1\}$ ($\theta = (\theta_1, \theta_2)^T$). Using (11), the left factorization $Q^{-1}(z)\Upsilon(z)$ of $\Delta\bar{W}(z)$ is $Q(z) = 1$ and $\Upsilon(z) = \begin{pmatrix}0, & \delta\theta_1 z^{-1}, & \delta\theta_2 z^{-1}\end{pmatrix}$ (with $\delta\theta_1 = \theta_{1,0} - \theta_1$ and $\delta\theta_2 = \theta_{2,0} - \theta_2$). Defining $\delta = (\delta\theta_1,\ \delta\theta_2)^T$, we can write $\Upsilon(z) = \delta^T Z_\Upsilon(z)$ with

$$Z_\Upsilon(z) = \begin{pmatrix} 0 & z^{-1} & 0 \\ 0 & 0 & z^{-1} \end{pmatrix}$$

Since we are here in the conditions of Proposition 7, the matrix $\mathcal{A}_{N_1}$ will be based on a right factorization of $Z_\Upsilon(z)\bar{T}_{0,\mathcal{V}}(z)$. The matrix $\bar{T}_{0,\mathcal{V}}(z)$ is here equal to $(\bar{S}_{0,22}(z),\ \bar{S}_{0,12}(z),\ \bar{S}_{0,32}(z))^T$ i.e., a permutation of the second column of $\bar{S}_0(z)$. The matrix $Z_\Upsilon(z)\bar{T}_{0,\mathcal{V}}(z)$ can be here factorized as $N_1(z)V_1^{-1}(z)$ with: $N_1(z) = \begin{pmatrix}0.25z^{-3} & 0.5z^{-2}\end{pmatrix}^T$ and $V_1(z) = 1 - 0.5z^{-2} - 0.125z^{-3}$. The polynomial matrix $N_1(z)$ can subsequently be factorized as:

$$N_1(z) = \mathcal{A}_{N_1} Z_{N_1}(z) = \underbrace{\begin{pmatrix} 0 & 0.25 \\ 0.5 & 0 \end{pmatrix}}_{=\mathcal{A}_{N_1}} \begin{pmatrix} z^{-2} \\ z^{-3} \end{pmatrix}$$

Since $\mathcal{A}_{N_1}$ is full row rank, we have thus data informativity under the sole excitation of the unknown process noise $v_2(t)$. The data informativity property is confirmed by performing an identification with a large N in these conditions (i.e. $\bar{r}(t) = 0$ and $\mathcal{V} = \{2\}$) and by observing that $\hat{\theta}_N$ is indeed a very close estimate of θ_0. We can thus prove the data informativity using Proposition 9 while it is not possible to do so using Proposition 4 since $\mathcal{V}\backslash\{j\} = \emptyset$. This confirms the usefulness of Proposition 9 to check data informativity.

In the analysis above, we have used the true $\bar{G}_0(z)$ to deduce $\mathcal{A}_{N_1}$. We have nevertheless also applied the above procedure for different full-order models of $\bar{G}_0(z)$ and the rank of the matrix $\mathcal{A}_{N_1}$ remained equal to two for all these models. The tested models have all the topology of the network in Figure 4 and the five non-zero transfer functions have all the form $G_{kl}(z) = b_{kl}z^{-1}$ with a parameter vector $(b_{13}, b_{21}, b_{23}, b_{31}, b_{32})^T \in \mathbf{R}^5$ generated randomly.

6 Optimal experiment design

6.1 Results

Using the data informativity conditions of Proposition 9, we can show that consistent estimates of $(\bar{G}_{0,j,\mathcal{D}}(z), H_{0,j}(z))$ can be obtained for different sets $\mathcal{R}$ and different types of excitation vectors $\bar{r}_\mathcal{R}(t)$ (i.e., $\bar{r}_{\mathcal{R}_s}(t)$ and $\bar{r}_{\mathcal{R}_n}(t)$). This defines different identification options. Since consistency is an asymptotic property, these results do not say anything about the accuracy of the identified parameter vector $\hat{\theta}_N$ (which defines

the model of $(\bar{G}_{0,j,\mathcal{D}}(z), H_{0,j}(z)))$ under these different options. In this section, we will analyze the accuracy of $\hat{\theta}_N$ and determine the signal vector $\bar{r}_{\mathcal{R}}(t)$ leading to the desired accuracy for $\hat{\theta}_N$ with the smallest excitation power.

Since $\hat{\theta}_N$ is a consistent estimate of θ_0 and $\epsilon_j(t,\theta_0) = e_j(t)$, the estimate $\hat{\theta}_N$ is also (asymptotically) normally distributed around θ_0 with a covariance matrix P_θ that is given by $P_\theta = \frac{\sigma^2_{e_j}}{N}\left(\bar{E}\psi_j(t,\theta_0)\psi_j^T(t,\theta_0)\right)^{-1}$ with $\psi_j(t,\theta) = \frac{d\epsilon_j(t,\theta)}{d\theta}$. When we are more particularly interested in a subvector ρ_0 of θ_0, we use the fact that the subvector ρ_0 can always be written as $\rho_0 = S\,\theta_0$ for some matrix S and we define P_ρ as the covariance matrix of $S\,\hat{\theta}_N$. We have then $P_\rho = SP_\theta S^T$. Since $P_\theta = P_\rho$ when S is chosen as $S = I_{n_\theta}$ (n_θ denotes the dimension of θ), let us for the sake of generality continue our analysis for $P_\rho = SP_\theta S^T$.

We want to determine the excitation vector $\bar{r}_{\mathcal{R}}(t)$ that, for an identification experiment of duration N, yields an acceptable covariance matrix $P_\rho = SP_\theta S^T$ with the least excitation power. For this purpose, we will first assume that $n_{\mathcal{R}} = N_{mod}$ i.e., $\mathcal{R} = \{1, 2, ..., N_{mod}\}$ and we will determine the power spectrum matrix $\Phi_{\bar{r}}$ of the excitation vector $\bar{r}(t)$ having the smallest power while guaranteeing that the estimate $\hat{\theta}_N$ obtained via an identification experiment of duration N with this excitation has a covariance covariance matrix P_θ that satisfies the following constraint $SP_\theta S^T \leq R_{adm}$ where R_{adm} specifies the desired accuracy (a diagonal R_{adm} e.g., allows to constrain the standard deviations of each entries of $S\,\hat{\theta}_N$ [11]). We thus require $R_{adm} - SP_\theta S^T \geq 0$ and, using Schur complement, this gives the following optimal experiment design problem:

$$\min_{\Phi_{\bar{r}}(\omega)} \; trace\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_{\bar{r}}(\omega)\;d\omega\right)$$
$$\text{subject to } \begin{pmatrix} R_{adm} & S \\ S^T & P_\theta^{-1}\end{pmatrix} \geq 0 \qquad (32)$$

This optimization problem is convex since, as will be shown in the sequel, P_θ^{-1} is an affine function of $\Phi_{\bar{r}}(\omega)$. As indicated above, in this optimal experiment design problem, the main objective is to determine the least powerful excitation vector $\bar{r}(t)$ such that $SP_\theta S^T \leq R_{adm}$. Note nevertheless that the objective function of the optimization problem (32) has a l_1-norm structure. Since it is frequently observed that such objective functions, when minimized under convex constraints, generate a sparse solution (see e.g., [16]), we can expect that the optimal excitation vector $\bar{r}(t)$ will have some elements r_k equal to zero. In other words, the set $\mathcal{R}_{opt}$ corresponding to the solution of (32) will generally have a cardinality $n_{\mathcal{R}_{opt}}$ that is smaller than N_{mod} (defining in this way the optimal excitation pattern).

Let us now derive the affine relation between $\Phi_{\bar{r}}(\omega)$ and P_θ^{-1}. Using the philosophy introduced in [10,3], we have that: $\psi_j(t,\theta_0) = \Gamma_1(z,\theta_0)\;\bar{w}_{\mathcal{D}}(t) + \Gamma_2(z,\theta_0)\;e_j(t)$ where $\Gamma_1(z,\theta)$ is a matrix of dimension $n_\theta \times n_{\mathcal{D}}$ whose l^{th} row is given by $H_{0,j}^{-1}(z)\frac{d\bar{G}_{j,\mathcal{D}}(z,\theta)}{d\theta_l}$ (θ_l is the l^{th} entry of $\theta \in \mathbf{R}^{n_\theta}$) and $\Gamma_2(z,\theta)$ is a vector of dimension n_θ whose l^{th} entry is given by $H_{0,j}^{-1}(z)\frac{dH_j(z,\theta)}{d\theta_l}$. Using now the fact that, for any set $\mathcal{R}$, $\bar{w}_{\mathcal{D}}(t) = \bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t) + \bar{S}_{0,\mathcal{D},\mathcal{V}}(z)H_{0,\mathcal{V},\mathcal{V}}(z)\bar{e}_{\mathcal{V}}(t)$, we can rewrite the previous equation as follows: $\psi_j(t,\theta_0) = \Gamma_{\bar{r}}(z,\theta_0)\bar{r}_{\mathcal{R}}(t) + \Gamma_{\bar{e}}(z,\theta_0)\bar{e}_{\mathcal{V}}(t)$ with $\Gamma_{\bar{r}}(z,\theta_0) = \Gamma_1(z,\theta_0)\bar{S}_{0,\mathcal{D},\mathcal{R}}(z)$ and with $\Gamma_{\bar{e}}(z,\theta_0) = \Gamma_2(z,\theta_0)\,m_j^T + \Gamma_1(z,\theta_0)\bar{S}_{0,\mathcal{D},\mathcal{V}}(z)H_{0,\mathcal{V},\mathcal{V}}(z)$ where the column vector m_j of dimension $n_{\mathcal{V}}$ is a unit vector such that $m_j^T\bar{e}_{\mathcal{V}}(t) = e_j(t)$. As mentioned above, we here choose $\mathcal{R} = \{1, 2, ..., N_{mod}\}$ for the experiment design and we have thus $\bar{r}_{\mathcal{R}}(t) = \bar{r}(t)$. Consequently,

$$P_\theta^{-1} = \tfrac{N}{\sigma^2_{e_j}}\left(\bar{E}\psi_j(t,\theta_0)\psi_j^T(t,\theta_0)\right) = R_{\bar{r}}(\Phi_{\bar{r}}(\omega),\theta_0) + R_{\bar{e}}(\theta_0)$$
$$R_{\bar{r}}(\Phi_{\bar{r}}(\omega),\theta_0) = \tfrac{N}{\sigma^2_{e_j}}\;\tfrac{1}{2\pi}\int_{-\pi}^{\pi}\Gamma_{\bar{r}}(e^{j\omega},\theta_0)\;\Phi_{\bar{r}}(\omega)\;\Gamma^*_{\bar{r}}(e^{j\omega},\theta_0)\;d\omega$$
$$R_{\bar{e}}(\theta_0) = \tfrac{N}{\sigma^2_{e_j}}\;\tfrac{1}{2\pi}\int_{-\pi}^{\pi}\Gamma_{\bar{e}}(e^{j\omega},\theta_0)\;\Sigma_{0,\mathcal{V}}\;\Gamma^*_{\bar{e}}(e^{j\omega},\theta_0)\;d\omega$$

with $\Sigma_{0,\mathcal{V}}$ the covariance matrix of $\bar{e}_{\mathcal{V}}(t)$.

Since $\Phi_{\bar{r}}(\omega)$ is a variable of infinite dimension, we need to choose a linear parametrization for $\Phi_{\bar{r}}(\omega)$ to solve the convex optimization problem (32) [13,4,1]. We can e.g., choose the parametrization given in [1] and that corresponds to filtered white noise $\bar{r}(t)$. However, in order to simplify this complex optimization problem, we will here restrict attention to a parametrization corresponding to an excitation vector $\bar{r}(t)$ made up of mutually independent white noises: $\Phi_{\bar{r}}(\omega) = diag(c_1, c_2, ..., c_{N_{mod}})\;\forall\omega$ where c_k ($k = 1, .., N_{mod}$) is the to-be-determined variance of r_k.

Remark. Like in all optimal experiment design problems, P_θ depends on the unknown θ_0 (i.e. the true parameter vector describing $H_{0,j}(z)$ and $\bar{G}_{0,j,\mathcal{D}}(z)$) and also on the unknown matrices $\bar{S}_0(z)$ and $\bar{H}_0(z)$. Initial estimates of these unknown quantities are thus necessary to solve the optimization problem (32). ■

6.2 Numerical illustration

Let us consider the same network as in Section 5.4 where we wish to identify $G_{0,21}(z)$ and $G_{0,23}(z)$ (i.e. $j = 2$). For that network, as proposed above, we will solve (32) using $\Phi_{\bar{r}}(\omega) = diag(c_1, c_2, c_3)\;\forall\omega$ where c_k ($k = 1, .., 3$) is to-be-determined variance of r_k. As shown in Section 5.4, the model structure $\mathcal{M}$ considered for this identification is $\mathcal{M} = \{\bar{G}_{2,\mathcal{D}}(z,\theta)\}$ with $\bar{G}_{2,\mathcal{D}}(z,\theta) = (\theta_1 z^{-1},\; \theta_2 z^{-1})$ ($\theta = (\theta_1, \theta_2)^T$). Let us first suppose that we are interested in having an accurate estimate $\hat{\theta}^N$ of the whole true parameter vector $\theta_0 = (0.5, 0.5)^T$. This means that we choose $S = I_2$. Let us also define R_{adm} as $R_{adm} = 10^{-5}I_2$. In Section 5.4, we have proven that we have data informativity when $\bar{r}(t) = 0$. Consequently,

if we choose N sufficiently large, the optimal spectrum $\Phi_{\bar{r}}^{opt}(\omega)$ will be equal to zero ($\mathcal{R}_{opt} = \emptyset$). This is possible because the matrix $R_{\bar{e}}(\theta_0)$ is strictly positive definite and proportional to N. However, for $N = 1000$, we obtain $\Phi_{\bar{r}}^{opt}(\omega) = diag(8.54, 0, 5.82)$ which corresponds to an excitation signal on Node 1 and on Node 3. In other words, $\mathcal{R}_{opt} = \mathcal{D} = \{1, 3\}$.

Let us now suppose that it is impossible to add an excitation signal on Node 1. We can then solve the optimization problem (32) imposing $c_1 = 0$. We then obtain $\Phi_{\bar{r}}^{opt}(\omega) = diag(0, 0, 29.35)$ which corresponds to a unique excitation signal on Node 3. However, we observe that, when we cannot excite Node 1, much more power has to be injected in the network to obtain the desired accuracy. Let us constrain (32) even more by supposing that we can only add an excitation on Node 2 (not on Nodes 1 and 3). We thus impose $c_1 = c_3 = 0$ in (32) and we obtain $\Phi_{\bar{r}}^{opt}(\omega) = diag(0, 117.41, 0)$ where we see that the excitation power (which is now on the only excitable node) is even larger.

Let us now come back to the case where we can add an excitation on each node, but let us now suppose that we are only interested in an accurate estimate of the transfer function $G_{0,21}(z)$. We therefore choose $S = (1\ 0)$ and $R_{adm} = 10^{-5}$. The optimal solution of (32) is then $\Phi_{\bar{r}}^{opt}(\omega) = diag(9.36, 0, 0)$ which corresponds to a unique excitation signal on Node 1 (i.e., the input signal $w_1(t)$ of $G_{0,21}(z)$). If we are only interested in an accurate estimate of the transfer function $G_{0,23}(z)$ ($S = (0\ 1)$), we observe the same phenomenon since the optimal spectrum is given by $\Phi_{\bar{r}}^{opt}(\omega) = diag(0, 0, 7.28)$.

7 Conclusions

In this paper, we have addressed the data informativity problem for the identification of a row of $\bar{G}_0(z)$ using the full input approach. We have derived conditions that ensure data informativity for the identification of modules having an arbitrary complexity and that can be checked by inspecting the topology of the network. We have also developed a necessary and sufficient condition for data informativity that takes into account the complexity of the to-be-identified modules. This particular data informativity condition is verified if a matrix of coefficients is full row rank. The determination of this matrix of coefficients requires a full-order model of the network matrix $\bar{G}_0(z)$. Consequently, the verification of the necessary and sufficient data informativity condition requires more information on the network than the more conservative conditions (that can be verified by inspecting the network topology).

References

[1] M. Barenthin, X. Bombois, H. Hjalmarsson, and G. Scorletti. Identification for control of multivariable systems: controller validation and experiment design via LMIs. *Automatica*, 44(12):3070–3078, 2008.

[2] X. Bombois, K. Colin, P. Van den Hof, and H. Hjalmarsson. On the informativity of direct identification experiments in dynamical network. HAL document, 2022. Available on the HAL repository at https://hal.archives-ouvertes.fr/hal-03329586v2/document .

[3] X. Bombois, A. Korniienko, H. Hjalmarsson, and G. Scorletti. Optimal identification experiment design for the interconnection of locally controlled systems. *Automatica*, 89:169–179, 2018.

[4] X. Bombois, G. Scorletti, M. Gevers, P.M.J. Van den Hof, and R. Hildebrand. Least costly identification experiment for control. *Automatica*, 42(10):1651–1662, 2006.

[5] K. Colin, X. Bombois, L. Bako, and F. Morelli. Closed-loop identification of MIMO systems in the prediction error framework: Data informativity analysis. *Automatica*, 121:109171, 2020.

[6] K. Colin, X. Bombois, L. Bako, and F. Morelli. Data informativity for the open-loop identification of MIMO systems in the prediction error framework. *Automatica*, 117:109000, 2020.

[7] P. Van den Hof, A. Dankers, P. Heuberger, and X. Bombois. Identification of dynamic models in complex networks with prediction error methods - basic methods for consistent module estimates. *Automatica*, 49(10):2994–3006, 2013.

[8] P. Van den Hof and K. Ramaswamy. Path-based data-informativity conditions for single module identification in dynamic networks. In *59th IEEE Conference on Decision and Control (CDC)*, pages 4354–4359, 2020.

[9] M. Gevers, A. S. Bazanella, and L. Miskovic. Informative data: how to get just sufficiently rich? In *Proc. 47th IEEE Conference on Decision and Control*, pages pp. 1962–1967, Cancun, Mexico, 2008.

[10] M. Gevers and A.S. Bazanella. Identification in dynamic networks: identifiability and experiment design issues. In *Proc. 54th IEEE Conference on Decision and Control*, pages 4005–4010, Osaka, Japan, 2015.

[11] D. Ghosh, X. Bombois, J. Huillery, G. Scorletti, and G. Mercère. Optimal identification experiment design for LPV systems using the local approach. *Automatica*, 87:258–266, 2018.

[12] J. Hendrickx, M. Gevers, and A. Bazanella. Identifiability of dynamical networks with partial node measurements. *IEEE Transactions on Automatic Control*, 64(6):2240–2253, 2019.

[13] H. Jansson and H. Hjalmarsson. Input design via LMIs admitting frequency-wise model specifications in confidence regions. *IEEE Transactions on Automatic Control*, 50(10):1534–1549, October 2005.

[14] L. Ljung. *System Identification: Theory for the User, 2nd Edition*. Prentice-Hall, Englewood Cliffs, NJ, 1999.

[15] E. Mapurunga and A. Bazanella. Optimal allocation of excitation and measurement for identification of dynamic networks. 2020. Available via arXiv 2007.09263.

[16] J.A. Tropp. Just relax: Convex programming methods for identifying sparse signals in noise. *IEEE Transactions on Information Theory, Vol 52, pp. 1030-1051*, 2006.

[17] H. Weerts, P. Van den Hof, and A. Dankers. Identifiability of linear dynamic networks. *Automatica*, 89:247–258, 2018.

A Proof of Proposition 4

Using (6), we can rewrite $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ as:

$$\bar{x}(t) = \underbrace{\begin{pmatrix} 1 & \bar{G}_{0,j,\mathcal{D}}(z) \\ 0 & I_{n_{\mathcal{D}}} \end{pmatrix}}_{=J(z)} \underbrace{\begin{pmatrix} v_j(t) \\ \bar{w}_{\mathcal{D}}(t) \end{pmatrix}}_{=\bar{x}_{bis}(t)} \tag{A.1}$$

As also shown in [8], since $J(e^{j\omega})$ is a full-rank square matrix at all ω, we have that (14) is equivalent to

$$\Phi_{\bar{x}_{bis}}(\omega) > 0 \text{ at almost all } \omega. \tag{A.2}$$

In the sequel of the proof, we will show that (A.2) holds under the conditions of Proposition 4. For this purpose, let us observe that, when $\bar{r}(t) = 0$, we have that $\bar{w}_{\mathcal{D}}(t) = \bar{S}_{0,\mathcal{D},\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t)$, which can also be rewritten as: $\bar{w}_{\mathcal{D}}(t) = \bar{S}_{0,\mathcal{D},j}(z)\ v_j(t) + \bar{S}_{0,\mathcal{D},\mathcal{U}}(z)\ v_{\mathcal{U}}(t)$ with $\mathcal{U} = \mathcal{V}\backslash\{j\}$. The vector $\bar{x}_{bis}(t)$ can thus be rewritten as:

$$\bar{x}_{bis}(t) = \underbrace{\begin{pmatrix} 1 & 0 \\ \bar{S}_{0,\mathcal{D},j}(z) & \bar{S}_{0,\mathcal{D},\mathcal{U}}(z) \end{pmatrix}}_{=\tilde{J}(z)} \underbrace{\begin{pmatrix} v_j(t) \\ v_{\mathcal{U}}(t) \end{pmatrix}}_{=\bar{\xi}(t)} \tag{A.3}$$

Observing that $\bar{\xi}(t) = (v_j(t),\ v_{\mathcal{U}}^T(t))^T$ is just a permutation of $v_{\mathcal{V}}(t) = H_{0,\mathcal{V},\mathcal{V}}(z)e_{\mathcal{V}}(t)$ and recalling that Proposition 4 assumes that $\Sigma_{0,\mathcal{V}} > 0$, it is clear that $\Phi_{\bar{\xi}}(\omega) > 0$ at (almost) all ω. The latter and the fact that, when Condition (i) holds, the matrix $\tilde{J}(e^{j\omega})$ in (A.4) is full row rank at almost all ω show that (A.2) is indeed satisfied under the conditions in Proposition 4; concluding the proof.

B Proof of Proposition 5

In the proof of Proposition 4, we have shown that (14) is equivalent to (A.2). In the sequel, we will show that $\bar{x}_{bis}(t) = (v_j(t), \bar{w}_{\mathcal{D}}(t))^T$ satisfies (A.2) under the conditions of Proposition 5. For this purpose, let us observe that $\bar{w}_{\mathcal{D}}(t) = \bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t) + \bar{S}_{0,\mathcal{D},\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t)$ can here be rewritten as: $\bar{w}_{\mathcal{D}}(t) = \bar{S}_{0,\mathcal{D},j}(z)\ v_j(t) + \bar{S}_{0,\mathcal{D},\mathcal{Q}}(z)\ \bar{\rho}(t)$ where $\mathcal{Q} = \mathcal{R} \cup (\mathcal{V}\backslash\{j\})$ and where $\bar{\rho}(t)$ is a vector of dimension $n_{\mathcal{Q}}$ whose elements are equal to v_k (if $k \in \mathcal{V}\backslash\{j\}$ and $k \notin \mathcal{R}$), to $v_k + r_k$ (if $k \in \mathcal{V}\backslash\{j\}$ and $k \in \mathcal{R}$) or to r_k (if $k \notin \mathcal{V}\backslash\{j\}$ and $k \in \mathcal{R}$). The vector $\bar{x}_{bis}(t) = (v_j(t), \bar{w}_{\mathcal{D}}(t))^T$ can thus be rewritten as:

$$\bar{x}_{bis}(t) = \underbrace{\begin{pmatrix} 1 & 0 \\ \bar{S}_{0,\mathcal{D},j}(z) & \bar{S}_{0,\mathcal{D},\mathcal{Q}}(z) \end{pmatrix}}_{=\bar{L}(z)} \underbrace{\begin{pmatrix} v_j(t) \\ \bar{\rho}(t) \end{pmatrix}}_{=\bar{\kappa}(t)} \tag{B.1}$$

Observing that the noise v_j does not appear in $\bar{\rho}(t)$ and recalling that Proposition 5 assumes that $\Sigma_{0,\mathcal{V}} > 0$, that Condition (ii) holds and that $\bar{r}(t)$ and $\bar{e}(t)$ are uncorrelated (see Assumption 1), it is clear that $\Phi_{\bar{\kappa}}(\omega) > 0$ at (almost) all ω. The latter and the fact that, when Condition (i) holds, the matrix $\bar{L}(e^{j\omega})$ in (B.2) is full row rank at almost all ω show that (A.2) is indeed satisfied under the conditions in Proposition 5; concluding the proof.

C Proof of Proposition 6

In the proof of Proposition 4, we have shown that (14) is equivalent to (A.2). In the sequel, we will show that $\bar{x}_{bis}(t) = (v_j(t), \bar{w}_{\mathcal{D}}(t))^T$ satisfies (A.2) under the conditions of Proposition 6. For this purpose, using $\bar{w}_{\mathcal{D}}(t) = \bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t) + \bar{S}_{0,\mathcal{D},\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t)$, we have

$$\bar{x}_{bis}(t) = \begin{pmatrix} v_j(t) \\ \bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t) + \bar{S}_{0,\mathcal{D},\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t) \end{pmatrix} \tag{C.1}$$

Under Conditions (i) and (ii) of Proposition 6, we have that the vector $\bar{\zeta}(t) = \bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t)$ has a strictly positive-definite power spectrum matrix $\Phi_{\bar{\zeta}}(\omega)$ at almost all frequencies. Since $\bar{\zeta}(t)$ is also uncorrelated with $v_j(t)$ (Assumption 1), it is thus clear that (A.2) is indeed satisfied under the conditions in Proposition 6; concluding the proof.

D Proof of Proposition 7

We will first prove the result for the case $\bar{r}(t) \neq 0$. Let us consider Definition 1 for this purpose. Using (17), we can rewrite $\Delta\bar{W}(z)\bar{x}(t)$ (see Definition 1) in the following way:

$$\Delta\bar{W}(z)\bar{x}(t) = \bar{s}_e(t) + \bar{s}_q(t) + \ \bar{s}_d(t) \tag{D.1}$$

$$\bar{s}_e(t) = \Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z)\ \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\ \bar{e}_{\mathcal{V}}(t) \tag{D.2}$$

$$\bar{s}_q(t) = \Delta\bar{W}(z)\ \bar{X}_{0,\mathcal{R}_n}(z)\ \bar{F}(z)\ \bar{q}(t) \tag{D.3}$$

$$\bar{s}_d(t) = \Delta\bar{W}(z)\ \bar{d}(t) \tag{D.4}$$

Since Assumption 1 states that $\bar{e}(t)$ is independent of $\bar{r}(t)$ (and thus of $\bar{q}(t)$ and $\bar{d}(t)$) and since $\bar{q}(t)$ is also independent of $\bar{d}(t)$, the left hand-side of (13) is equivalent to:

$$\begin{cases} \bar{E}\bar{s}_e^2(t) = 0 \\ \bar{E}\bar{s}_q^2(t) = 0 \\ \bar{E}\bar{s}_d^2(t) = 0 \end{cases} \tag{D.5}$$

When $\Sigma_{0,\mathcal{V}} > 0$, we have that the power spectrum matrix $\Phi_{v_{\mathcal{V}}}(\omega)$ of $v_{\mathcal{V}}(t) = H_{0,\mathcal{V},\mathcal{V}}(z)e_{\mathcal{V}}(t)$ is strictly positive-definite at almost all ω. Using this property, we see that $\bar{E}\bar{s}_e^2(t) = 0$ (see (D.2)) is equivalent to $\Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z) = 0$. Using the fact that $\Phi_{\bar{q}}(\omega) = I_{n_q} > 0$, we see also that $\bar{E}\bar{s}_q^2(t) = 0$ (see (D.3)) is equivalent to the second equation of the left hand side of (21). Finally, $\bar{E}\bar{s}_d^2(t) = 0$ (see (D.4)) is equivalent to the third equation of the left hand side of (21). Combining these facts, we conclude that the left hand side of (13) is equivalent to the left hand side of (21). Consequently, using Definition 1, (21) is indeed a necessary and sufficient data informativity condition. The last statement of Proposition 7 can be proven as follows. When $\bar{F}(e^{j\omega})$ is full row rank at almost all ω, the vector $\bar{r}_{\mathcal{R}_n}(t)$ in (16) has the property that $\Phi_{\bar{r}_{\mathcal{R}_n}}(\omega) > 0$ at almost all ω. This in turn means that $\bar{E}\bar{s}_q^2(t) = 0$ is in this case equivalent to $\Delta\bar{W}(z)\ \bar{X}_{0,\mathcal{R}_n}(z) = 0$.

Let us now turn our attention to the costless case. For this purpose, let us observe that, when $\bar{r}(t) = 0$, (D.1) becomes $\Delta\bar{W}(z)\bar{x}(t) = \bar{s}_e(t)$. Consequently, the left

hand side of (13) is equivalent to $\bar{E}\bar{s}_e^2(t) = 0$ and thus to the left hand side of (20) (as shown above). It is thus clear that (20) is the necessary and sufficient data informativity condition in the costless case.

E Proof of Proposition 8

Let us denote by p the rank of $\Sigma_{0,\mathcal{V}}$ ($p \leq n_{\mathcal{V}}$). Since $\Sigma_{0,\mathcal{V}} = \Xi_{0,\mathcal{V}}\Xi_{0,\mathcal{V}}^T$ with $\Xi_{0,\mathcal{V}} \in \mathbf{R}^{n_{\mathcal{V}} \times p}$, we can rewrite $\bar{e}_{\mathcal{V}}(t)$ as $\bar{e}_{\mathcal{V}}(t) = \Xi_{0,\mathcal{V}}\, \bar{e}_{unit}(t)$ where the power spectrum matrix $\Phi_{\bar{e}_{unit}}(\omega)$ of $\bar{e}_{unit}(t)$ is equal to the identity matrix $I_p > 0$ at all ω. Consequently, (D.2) can be rewritten as

$$\bar{s}_e(t) = \Delta\bar{W}(z)\, \bar{T}_{0,\mathcal{V}}(z)\, \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\, \Xi_{0,\mathcal{V}}\, \bar{e}_{unit}(t) \quad \text{(E.1)}$$

Since $\Phi_{\bar{e}_{unit}}(\omega) = I_p > 0$, $\bar{E}\bar{s}_e^2(t) = 0$ is equivalent to $\Delta\bar{W}(z)\, \bar{T}_{0,\mathcal{V}}(z)\, \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\, \Xi_{0,\mathcal{V}} = 0$. Similar arguments as in the proof of Proposition 7 then leads to the desired result.