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. In this paper, we consider data informativity in the case of dynamic network identification. In particular, we derive a number of data informativity conditions for the prediction error identification of a particular module of a dynamic network using a direct identification approach. An optimal experiment design approach is also proposed to distinguish between the different situations leading to data informativity.

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

1 Introduction

We consider in this paper the prediction error identification of a single module of a dynamical network[10,5]. By module, we mean the causal transfer function between two nodes of this network. The considered network is subject to a number of exogenous excitations. Among these exogenous excitations, we distinguish, on the one hand, the (unknown) process noises v_k acting on (some of) the nodes of the network and, on the other hand, known excitation signals r_k that can be added at some nodes to increase the informativity of the data for identification purpose [5]. In order to identify a single module of that network, we can consider, like in closed-loop identification, either a direct identification of the module [5] or an indirect identification of this module (i.e., the model of the module is back-computed from an identified model of a closed-loop representation of the network [10,11]). In this paper, we will consider the direct approach and, in particular, the full input approach introduced in [5]. When the to-be-identified module is the module between Node i and Node j, this full input approach consists in also identifying all the other modules influencing Node j. The latter is done in order to guarantee the consistency of the prediction error estimate of the desired module. To guarantee this consistency, [5] imposes another important condition, namely a condition on the power spectrum matrix of the data used for the identification. This condition ensures that these data are informative with respect to (wrt.) the model structure $\mathcal{M}$ i.e., it ensures that the prediction error is different for different models in $\mathcal{M}$ [13]. There are however two drawbacks with the particular data informativity condition proposed in [5]. First, this condition is independent of the model structure (i.e., of the model order) and is therefore conservative. Moreover, as also pointed out in [8], it is difficult to interpret this condition in order to determine at which node an excitation signal r_k has to be added and what are the conditions on its power spectrum to ensure data informativity. In [8,6], preliminary steps are made to tackle these two drawbacks. In this paper, we extend these preliminary results by giving a necessary and sufficient condition for data informativity (a condition that takes the model order into account). The form of this rather complex data informativity condition allows the use of the framework introduced in [4] to check whether this condition holds in practice and what measures have to be taken in order to increase the informativity of the data when we face a situation where we do not have data informativity. This possibility is the main difference between the necessary and sufficient data informativity condition proposed in this paper and the one introduced in [8]. Based on this necessary and sufficient condition, we also derive a number of simpler data informativity conditions. These conditions are independent of the model order and are thus conservative, but, unlike the one in [5], it is quite easy to interpret these conditions in order to determine at which node an excitation signal has to be added to ensure data informativity.

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 lead to a consistent estimate of the to-be-identified module, i.e. to an estimate that converges 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 in all these situations can be much different. In order to distinguish between these situations, we propose to use optimal experiment design. In particular, we will determine the particular excitation pattern that leads to a given accuracy of the desired module with the smallest excitation power. The use of optimal experiment design for this purpose was first introduced in our paper [2] where a very specific type of networks is considered, namely the interconnection of simple closed-loop systems. We here extend this work towards the generic network description in [5]. It is to be noted that, in [14], a similar problem is also considered, but for an indirect identification approach.

In this paper, we analyze the data informativity for the direct identification of a module in a network. The present paper is also related to the notion of *network identifiability* (see e.g., [10,7,11,16]). In a nutshell, these works derive conditions on the excitation pattern in order to be able to uniquely retrieve certain modules of the network from closed-loop representations of this network. By closed-loop representations of the network, we mean the matrix transfer functions between the exogenous signals v_k and r_k and (some of) the node measurements w_k. In this paper, we show that the conditions for *network identifiability* and the conditions to guarantee the data informativity for a direct identification approach are rather similar when we restrict attention to excitations signals r_k that are filtered white noises.

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 $\bar{w}(t) = \bar{X}(z)\bar{x}(t)$. 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 $x(t)$, $\bar{E}x(t) \triangleq lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} Ex(t)$ (E is the expectation operator) and $\otimes$ denotes the Kronecker product.

2 Network description

In this paper, we consider the problem of identifying a single module 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 [5]:

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

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

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

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 closed-loop description $\bar{S}_0(z) = (I_{N_{mod}} - \bar{G}_0(z))^{-1}$ of the network is well-posed[2] and stable 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[3]. 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. 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 $\Sigma_{0,\mathcal{V}}$ (see e.g. [7]).

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 $r_k \neq 0$.

Remark. Even though the assumption that $\bar{H}_0(z)$ is diagonal is restrictive, the possibility of spatial correlation between the elements of $\bar{e}(t)$ allows to generate a large class of disturbances $\bar{v}(t)$ having as power spectrum $\Phi_{\bar{v}}(\omega) = \bar{H}_0(e^{j\omega})\Sigma_0\bar{H}^*_0(e^{j\omega})$ (recall that $\Sigma_0 \geq 0$ can have non-zero off-diagonal terms). Note also that, since Σ_0 is not assumed strictly positive definite, the matrix (3) is not the unique monic, stable and inversely stable matrix that allows to model the random vector $\bar{v}(t)$ with the aforementioned spectrum [7]. This non-uniqueness is here not a problem since the identification procedure of the next section will only require that, for a given k where $\sigma^2_{e,k} \neq 0$, $v_k(t)$ can be uniquely represented as $v_k(t) = H_{0,k}(z)e_k(t)$ with a stable, inversely stable and monic transfer function $H_{0,k}$ and this property obviously holds.

3 Identification of a single module $G_{0,ji}(z)$ of the network

As already mentioned, our objective is to use prediction error identification to accurately identify a single module of the matrix $\bar{G}_0(z)$, say $G_{0,ji}(z)$. We will use for this purpose the Multiple Input Single Output (MISO) approach introduced in [5] and consisting in identifying a model of all the (unknown) elements in the j^{th} row of $\bar{G}_0(z)$ and a model of $H_{0,j}(z)$. Before presenting this identification approach in more detail, let us introduce some concepts related to the j^{th} row of $\bar{G}_0(z)$. This row can contain entries $G_{0,jk}(z)$ that are known to be identically zero, entries $G_{0,jk}(z)$ that are both known and not equal to zero and, finally, entries $G_{0,jk}(z)$ that are unknown. For the sequel, we need to define two additional set 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 [13]. As in direct closed-loop identification, we will here also need in many cases[4] to require that $\bar{G}_{0,j,\mathcal{D}}(z)$ is stable [13].

[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.

[2] Note that the well-possedness of the square matrix $\bar{S}_0(z)$ implies that $S_0(e^{j\omega})$ is full row rank at (almost) all frequencies ω.

[3] This can e.g. be the case if w_k represents the output of a controller.

[4] This additional assumption is not required if (6) is in the ARX or the ARMAX form.

Let us thus suppose that we have collected on the network (1), the data set $Z^N = \{y_j(t),\ \bar{w}_{\mathcal{D}}(t) \mid t = 1...N\}$ 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 [13]:

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

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

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 the sought model for $G_{0,ji}(z)$ is one of the entries of $\bar{G}_{j,\mathcal{D}}(z,\hat{\theta}_N)$ (since i necessarily lies in $\mathcal{D}$). The MISO identification problem (7)-(8) has been introduced in [5]. In Appendix A, we compare this MISO approach with a Multiple Input Multiple Output (MIMO) approach consisting in identifying $\bar{G}_0(z)$ and $\bar{H}_0(z)$ entirely.

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 $\bar{E}\epsilon_j^2(t,\theta)$ [13]. 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 $v_j(t) = e_j(t) = 0$ will briefly be treated in Appendix B.

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

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 [13] (see also [5]).

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. Using (6), (8) can be rewritten as $\epsilon_j(t,\theta) = H_j^{-1}(z,\theta)\left(\Delta\bar{G}_{j,\mathcal{D}}(z,\theta)\bar{w}_{\mathcal{D}}(t) + H_{0,j}(z)e_j(t)\right)$ where $\Delta\bar{G}_{j,\mathcal{D}}(z,\theta) = \bar{G}_{0,j,\mathcal{D}}(z) - \bar{G}_{j,\mathcal{D}}(z,\theta)$. Using (4) and the sets $\mathcal{R}$ and $\mathcal{V}$ defined at the end of Section 2, we have also that $\bar{w}_{\mathcal{D}}(t) = \bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t) + \bar{S}_{0,\mathcal{D},\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t)$. This leads to:

$$\epsilon_j(t,\theta) = e_j(t) + s_1(t,\theta) + s_2(t,\theta) \tag{9}$$

$$s_1(t,\theta) = \frac{\Delta\bar{G}_{j,\mathcal{D}}(z,\theta)}{H_j(z,\theta)}\bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t)$$

$$s_2(t,\theta) = \frac{\Delta H_j(z,\theta)}{H_j(z,\theta)}e_j(t) + \frac{\Delta\bar{G}_{j,\mathcal{D}}(z,\theta)}{H_j(z,\theta)}\bar{S}_{0,\mathcal{D},\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t)$$

where $\Delta H_j(z,\theta) = H_{0,j}(z) - H_j(z,\theta)$. Recall that $\bar{v}_{\mathcal{V}}(t) = \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\bar{e}_{\mathcal{V}}(t)$. Due to the delay condition in the statement of Proposition 1, to the assumptions on $\bar{e}(t)$ in Assumption 1 and to the fact that $H_{0,j}(z)$ is monic, we have that $\bar{E}(e_j(t)s_1(t,\theta)) = \bar{E}(e_j(t)s_2(t,\theta)) = \bar{E}(s_1(t,\theta)s_2(t,\theta)) = 0$ and thus:

$$\bar{E}\epsilon_j^2(t,\theta) = \sigma_{e_j}^2 + \bar{E}s_1^2(t,\theta) + \bar{E}s_2^2(t,\theta) \tag{10}$$

where $\sigma_{e_j}^2$ is the variance of $e_j(t)$. Since $\bar{E}s_1^2(t,\theta_0) = \bar{E}s_2^2(t,\theta_0) = 0$, it is clear that θ_0 minimizes $\bar{E}\epsilon_j^2(t,\theta)$.

It is also clear that all eventual other minimizers θ^* of $\bar{E}\epsilon_j^2(t,\theta)$ should also be characterized by $\bar{E}s_1^2(t,\theta^*) = \bar{E}s_2^2(t,\theta^*) = 0$. By virtue of (9), we have thus also that all eventual other minimizers θ^* of $\bar{E}\epsilon_j^2(t,\theta)$ should have the property that $\epsilon_j(t,\theta^*) = \epsilon_j(t,\theta_0) = e_j(t)$. ■

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

$$\bar{G}_{0,j,\mathcal{D}}(z) - \bar{G}_{j,\mathcal{D}}(z,\theta) = 0 \quad \textit{and} \quad H_{0,j}(z) - H_j(z,\theta) = 0 \tag{11}$$

This property is generally called *data informativity* in the literature [13]. Let us formally define this notion. For this purpose, let us introduce the following notation using (8):

$$\epsilon_j(t,\theta) = \bar{W}(z,\theta)\bar{x}(t) = \left(W_y(z,\theta), \bar{W}_w(z,\theta)\right)\begin{pmatrix} y_j(t) \\ \bar{w}_{\mathcal{D}}(t) \end{pmatrix} \tag{12}$$

$$W_y(z,\theta) = H_j^{-1}(z,\theta) \;\; and \;\; \bar{W}_w(z,\theta) = -H_j^{-1}(z,\theta)\bar{G}_{j,\mathcal{D}}(z,\theta) \tag{13}$$

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:*

$$\mathbf{\Delta_{\bar{W}}} = \{\Delta\bar{W}(z) = \bar{W}(z,\theta) - \bar{W}(z,\theta_0) \mid \theta \in \Theta\} \tag{14}$$

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 $\Delta\bar{W}(z) \in \mathbf{\Delta_{\bar{W}}}$, we have:

$$\bar{E}\left(\Delta\bar{W}(z)\bar{x}(t)\right)^2 = 0 \Longrightarrow \Delta\bar{W}(z) = \mathbf{0} \tag{15}$$

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 (12) 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).* ∎

In the sequel, we will derive conditions that will allow us to verify that the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are informative wrt. $\mathcal{M}$. For this purpose, we will first rewrite $\bar{x}(t)$ in an appropriate manner. Recall that $y_j(t) = w_j(t) - r_j(t) - \bar{G}_{0,j,\mathcal{K}}(z)\bar{w}_{\mathcal{K}}(t)$. Using the fact that $w_j(t) = \bar{S}_{0,j,\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t) + \bar{S}_{0,j,\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t)$, the fact that $\bar{w}_{\mathcal{K}}(t) = \bar{S}_{0,\mathcal{K},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t) + \bar{S}_{0,\mathcal{K},\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t)$ and the fact that $\bar{w}_{\mathcal{D}}(t) = \bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t) + \bar{S}_{0,\mathcal{D},\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t)$, it is clear that we have that:

$$\bar{x}(t) = \begin{pmatrix} y_j(t) \\ \bar{w}_{\mathcal{D}}(t) \end{pmatrix} = \bar{T}_{0,\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t) + (\bar{T}_{0,\mathcal{R}}(z) - M)\bar{r}_{\mathcal{R}}(t)$$

where M is a matrix of dimension $(1+n_{\mathcal{D}}) \times n_{\mathcal{R}}$ such that $M\ \bar{r}_{\mathcal{R}}(t) = (r_j(t), 0, ..., 0)^T$ and where, 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

$$\bar{T}_{0,\mathcal{X}}(z) = \underbrace{\begin{pmatrix} 1 & -\bar{G}_{0,j,\mathcal{K}}(z) & 0 \\ 0 & 0 & I_{n_{\mathcal{D}}} \end{pmatrix}}_{=\bar{L}(z)} \begin{pmatrix} \bar{S}_{0,j,\mathcal{X}}(z) \\ \bar{S}_{0,\mathcal{K},\mathcal{X}}(z) \\ \bar{S}_{0,\mathcal{D},\mathcal{X}}(z) \end{pmatrix} \tag{16}$$

Using the expressions above, we can rewrite $\Delta\bar{W}(z)\bar{x}(t)$ (see Definition 1) in the following way:

$$\Delta\bar{W}(z)\bar{x}(t) = s_{\bar{e}}(t) + s_{\bar{r}}(t) \tag{17}$$

$$s_{\bar{e}}(t) = \Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z)\ \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\ \bar{e}_{\mathcal{V}}(t) \tag{18}$$

$$s_{\bar{r}}(t) = \Delta\bar{W}(z)\ \left(\bar{T}_{0,\mathcal{R}}(z) - M\right)\ \bar{r}_{\mathcal{R}}(t) \tag{19}$$

4 Necessary and sufficient condition for data informativity

We have now all the elements to derive the following proposition that gives a necessary and sufficient condition 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 [3,4]).

Proposition 3 *Consider the network (1) described in Section 2 and satisfying Assumptions 1 and 3. 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 $\Delta\bar{W}(z)\bar{x}(t)$ with (18) and (19). 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 $\Delta\bar{W}(z) \in \mathbf{\Delta_{\bar{W}}}$,*

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

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 $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}(t))^T$ are informative wrt. $\mathcal{M}$ if and only if, for all $\Delta\bar{W}(z) \in \mathbf{\Delta_{\bar{W}}}$,

$$\begin{cases} \Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z)\ \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\ \Xi_{0,\mathcal{V}} = 0 \\ \bar{E}\left(\Delta\bar{W}(z)\ (\bar{T}_{0,\mathcal{R}}(z) - M)\ \bar{r}_{\mathcal{R}}(t)\right)^2 = 0 \end{cases} \Longrightarrow \Delta\bar{W}(z) = 0 \tag{21}$$

∎

Before proving this proposition, let us give the following corollary that considers the case of networks for which $\Sigma_{0,\mathcal{V}}$ is strictly positive definite i.e., the case of networks where $\bar{e}_{\mathcal{V}}(t)$ is a full rank vector of signals.

Corollary 1 *Consider the framework of Proposition 3 and suppose that $\Sigma_{0,\mathcal{V}} > 0$. In this case, the necessary and sufficient conditions (20) and (21) remain of course valid, but equivalent (and simpler) necessary and sufficient conditions can be obtained by replacing $\Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z)\ \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\ \Xi_{0,\mathcal{V}} = 0$ by $\Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z) = 0$ in both (20) and (21).* ∎

Proof of Proposition 3 and of Corollary 1. Let us consider Definition 1. Since we have (17) with (18) and (19) and since $\bar{e}(t)$ is independent of $\bar{r}(t)$ (see Assumption 1), the left hand-side of (15) is equivalent to

$$\begin{cases} \bar{E}s_{\bar{e}}^2(t) = 0 \\ \bar{E}s_{\bar{r}}^2(t) = 0 \end{cases} \tag{22}$$

Let us first give the proof of Proposition 3 when the rank p of $\Sigma_{0,\mathcal{V}}$ can be smaller than $n_\mathcal{V}$ i.e. $p \leq n_\mathcal{V}$. 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, $\bar{E}s_{\bar{e}}^2(t) = 0$ is equivalent to $\Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z)\ \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\ \Xi_{0,\mathcal{V}} = 0$. In other words, the left hand side of (15) is thus equivalent to the left hand side of (21). Consequently, using Definition 1, (21) is indeed a necessary and sufficient data informativity condition. Finally, observe that (20) is equivalent to (21) for the case $\bar{r}_\mathcal{R}(t) = 0$ and is therefore the necessary and sufficient data informativity condition in the costless case. The proof of Corollary 1 follows from a similar reasoning. Since $\Sigma_{0,\mathcal{V}} > 0$, the power spectrum matrix of $\bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\ \bar{e}_\mathcal{V}(t)$ is guaranteed to be strictly positive definite at (almost) all ω. Consequently, $\bar{E}s_{\bar{e}}^2(t) = 0$ is here equivalent to $\Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z) = 0$ and this leads to the desired result. ∎

As we will see in the sequel, the framework that we recently developed in [4] allows to verify whether the data informativity conditions of Proposition 3 (and of Corollary 1) are satisfied for a given network configuration, for a given $\mathcal{M}$, for given noise and excitation pattern (i.e., for given $\mathcal{V}$ and $\mathcal{R}$) and for a given $r_\mathcal{R}(t)$ that can be a vector of multisines or a vector of filtered white noises. However, let us first derive, from the above necessary and sufficient conditions, some alternative and simpler data informativity conditions. As opposed to the conditions in Proposition 3 and Corollary 1, these conditions will only be sufficient, but they will be very simple to verify in practice.

5 Simple (but only sufficient) data informativity conditions

As was done in Proposition 3, let us first consider the costless identification case i.e., the case where $\bar{r}(t) = 0$. For the sake of simplicity, let us also consider that $\Sigma_{0,\mathcal{V}} > 0$ (like in Corollary 1).

Proposition 4 *Consider the network* (1) *described in Section 2 with* $\Sigma_{0,\mathcal{V}} > 0$ *and satisfying Assumptions 1 and 3. Consider the set* $\mathcal{V}$ *defined at the end of Section 2 and a model structure* $\mathcal{M}$ *satisfying Assumption 2. Then, in the case where the excitation vector* $\bar{r}(t)$ *is chosen equal to zero, the data* $\bar{x}(t) = (y_j,\ \bar{w}_\mathcal{D}^T(t))^T$ *are informative wrt.* $\mathcal{M}$ *if*

(i) the set $\mathcal{V}$ *describing the nodes where a disturbance* v_k *is present is such that, at (almost) all frequencies* ω, $\bar{T}_{0,\mathcal{V}}(e^{j\omega})$ *is full row rank i.e.,* $rank(\bar{T}_{0,\mathcal{V}}(e^{j\omega})) = 1 + n_\mathcal{D}$. ∎

Proof. If $\bar{T}_{0,\mathcal{V}}(e^{j\omega})$ is full row rank at (almost) all frequencies ω, it is clear that the identity $\Delta\bar{W}(z)\,\bar{T}_{0,\mathcal{V}}(z) = 0$ implies that the vector $\Delta\bar{W}(z) = 0$. The result of Proposition 4 thus follows from Corollary 1. ∎

Let us now consider 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 that will be assumed to be filtered white noises (no multisines).

Proposition 5 *Consider the network* (1) *described in Section 2 with* $\Sigma_{0,\mathcal{V}} > 0$ *and satisfying Assumptions 1 and 3. Consider the sets* $\mathcal{V}$ *and* $\mathcal{R}$ *defined at the end of Section 2 and a model structure* $\mathcal{M}$ *satisfying Assumption 2. Then, the data* $\bar{x}(t) = (y_j,\ \bar{w}_D^T(t))^T$ *are informative wrt.* $\mathcal{M}$ *if the following conditions all are satisfied:*

(ii) 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* $\mathcal{U} = \mathcal{V} \cup \mathcal{R}$ *has the property that, at (almost) all frequencies* ω, $rank(\bar{T}_{0,\mathcal{U}}(e^{j\omega})) = 1 + n_\mathcal{D}$

(iii) 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

(iv) If $j \in \mathcal{R}$, *the excitation signal* r_j *is uncorrelated with the other elements of* $\bar{r}_\mathcal{R}(t)$. ∎

Proof. See Appendix C. ∎

Remark. Condition (i) and Condition (ii) involve a rank condition on the matrix $\bar{T}_{0,\mathcal{X}}(e^{j\omega})$ with $\mathcal{X} = \mathcal{V}$ or $\mathcal{X} = \mathcal{U}$. As will be shown in the sequel, it can be useful in practice to reformulate these conditions as rank conditions on a part of the matrix $\bar{S}_0(e^{j\omega})$. Let us first consider the fairly classical case where $\mathcal{K} = \emptyset$ i.e., $\bar{G}_{0,j,\mathcal{K}} = 0$. When $\mathcal{K} = \emptyset$, we have (see (16)):

$$\bar{T}_{0,\mathcal{X}}(e^{j\omega}) = \begin{pmatrix} \bar{S}_{0,j,\mathcal{X}}(e^{j\omega}) \\ \bar{S}_{0,\mathcal{D},\mathcal{X}}(e^{j\omega}) \end{pmatrix} \tag{23}$$

Consequently, $\bar{T}_{0,\mathcal{X}}(e^{j\omega})$ is full row rank at (almost) all ω if and only if $\bar{S}_{0,\{j\}\cup\mathcal{D},\mathcal{X}}(e^{j\omega})$ is full row rank at (almost) all ω. Let us now consider the case where $\bar{G}_{0,j,\mathcal{K}} \neq 0$. Since $\bar{L}(e^{j\omega})$ in (16) is full row rank, $\bar{T}_{0,\mathcal{X}}(e^{j\omega})$ is full row rank at (almost) all ω if $\bar{S}_{0,\{j\}\cup\mathcal{K}\cup\mathcal{D},\mathcal{X}}(e^{j\omega})$ is full row rank at (almost) all ω. A less conservative condition can also be found if we exclude pathological cases. Let us indeed note that the first row of $\bar{T}_{0,\mathcal{X}}(e^{j\omega})$ (see (16)) is a linear combination of $1+n_\mathcal{K}$ row vectors i.e., $\bar{S}_{0,j,\mathcal{X}}(e^{j\omega})$ and $\bar{S}_{0,k,\mathcal{X}}(e^{j\omega})$ with $k \in \mathcal{K}$:

$$\bar{T}_{0,\mathcal{X}}(e^{j\omega}) = \begin{pmatrix} \bar{S}_{0,j,\mathcal{X}}(e^{j\omega}) - \sum_{k\in\mathcal{K}} G_{0,jk}(e^{j\omega})\bar{S}_{0,k,\mathcal{X}}(e^{j\omega}) \\ \bar{S}_{0,\mathcal{D},\mathcal{X}}(e^{j\omega}) \end{pmatrix}$$

Consequently, except in pathological cases, $\bar{T}_{0,\mathcal{X}}(e^{j\omega})$ is full row rank at (almost) all ω if any of the $1 + n_{\mathcal{K}}$ matrices $\bar{S}_{0,\{l\}\cup\mathcal{D},\mathcal{X}}(e^{j\omega})$ with $l \in \{j\}\cup\mathcal{K}$ is full row rank at (almost) all ω. When comparing this result with the one when $\mathcal{K} = \emptyset$, we observe the advantage of having known elements in the j^{th} row of $\bar{G}_0(z)$. ■

Until now, we have used Definition 1 to derive the data informativity conditions. We can also however use the equivalent definition presented above Definition 1. This allows to derive an additional informativity condition where we can relax both the condition on $\Sigma_{0,\mathcal{V}}$ ($\Sigma_{0,\mathcal{V}}$ can be rank deficient as in Section 2) and the conditions on the excitation vector $\bar{r}_{\mathcal{R}}$ (Condition (iv) is no longer required).

Proposition 6 *Consider the framework of Proposition 5, but with $\Sigma_{0,\mathcal{V}} \geq 0$. Then, the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are informative wrt. $\mathcal{M}$ if Conditions (iii) and (v) are both satisfied:*

(v) 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, $rank(\bar{S}_{0,\mathcal{D},\mathcal{R}}(e^{j\omega})) = n_{\mathcal{D}}$. ■

Proof. See Appendix D. ■

The matrices $\bar{T}_{0,\mathcal{V}}(e^{j\omega})$, $\bar{T}_{0,\mathcal{U}}(e^{j\omega})$ and $\bar{S}_{0,\mathcal{D},\mathcal{R}}(e^{j\omega})$ involved in Conditions (i), (ii) and (v) are functions of the unknown matrix $\bar{S}_0(e^{j\omega})$. One could thus consider to check Conditions (i), (ii) and (v) by replacing $\bar{S}_0(z) = (I_{N_{mod}} - \bar{G}_0(z))^{-1}$ by any stable models of $\bar{S}_0(z)$ obtained by replacing the entries of $\bar{G}_0(z)$ by any full-order models of these entries. However, this approach is not necessary here. Indeed, if we recall the remark below Proposition 5, we observe that Conditions (i), (ii) and (v) can all be checked by verifying that a submatrix $\bar{S}_{0,\mathcal{X},\mathcal{Y}}(e^{j\omega})$ of $\bar{S}_0(e^{j\omega})$ ($\mathcal{X}$ and $\mathcal{Y}$ are some sets of indexes) has a rank equal to $n_{\mathcal{X}}$ at (almost) all ω. This is important since [11] proposes a simple approach to verify such a condition. It is indeed shown in [11] that we can verify whether $rank(\bar{S}_{0,\mathcal{X},\mathcal{Y}}(e^{j\omega})) = n_{\mathcal{X}}$ at almost all ω by checking if there are $n_{\mathcal{X}}$ vertex-disjoint paths from the nodes $k \in \mathcal{Y}$ to the nodes $l \in \mathcal{X}$. The latter is a topological property of the graph of the network (see below) and it will therefore be very easy in practice to check and to interpret the data informativity conditions given in this section. It is to be noted that the equivalence between the rank property and the property on the number of vertex-disjoint paths can only be proven in a generic sense. By this, we in a nutshell mean that pathological cases are excluded (see [11] for more details).

As explained in e.g., [11], the graph of the network can be obtained by drawing a directed edge from Node k to Node l if $G_{0,lk}(z) \neq 0$. 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. Since r_k and v_k have a direct influence on w_k, there is always a path from Node k to Node k. Finally, vertex-disjoint paths are paths that do not pass through the same nodes/vertexes.

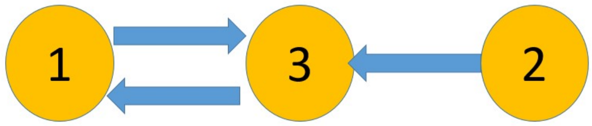

Fig. 1. Graph representation of (24). Each circle represents a node and the edges represent the structure of $\bar{G}_0(z)$.

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

$$\bar{G}_0(z) \triangleq \begin{pmatrix} 0 & 0 & G_{0,13}(z) \\ 0 & 0 & 0 \\ G_{0,31}(z) & G_{0,32}(z) & 0 \end{pmatrix} \quad (24)$$

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 e.g., exists a path 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 that there are two vertex-disjoint paths from the nodes $k \in \mathcal{Y}$ to the nodes $l \in \mathcal{X}$ i.e., 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 $k \in \mathcal{Y}$ to the nodes $l \in \mathcal{X}$ e.g., the path $3 \to 3$. The other paths from the nodes $k \in \mathcal{Y}$ to the nodes $l \in \mathcal{X}$ (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 $3 \to 3$. ■

Since, for any k, there is always a path from Node k to Node k, it is clear that $\mathcal{X} \subseteq \mathcal{Y}$ is a sufficient condition to have $n_{\mathcal{X}}$ vertex-disjoint paths from the nodes $k \in \mathcal{Y}$ to the nodes $l \in \mathcal{X}$ and thus, generically, to have $rank(\bar{S}_{0,\mathcal{X},\mathcal{Y}}(e^{j\omega})) = n_{\mathcal{X}}$ at almost all ω. Let us use this property to interpret Conditions (i) and (ii) and let us for simplicity restrict attention to the classical case where $\mathcal{K} = \emptyset$. As shown in the remark below Proposition 5, when $\mathcal{K} = \emptyset$, Condition (i) is respected if and only if $rank(\bar{S}_{0,\{j\}\cup\mathcal{D},\mathcal{V}}(e^{j\omega})) = 1 + n_{\mathcal{D}}$ at (almost) all ω. Using the property described at the beginning of this paragraph, among all the cases where this condition holds, we have e.g., the case where $(\{j\}\cup\mathcal{D}) \subseteq \mathcal{V}$. Since $j \in \mathcal{V}$ (Assumption 3), $(\{j\}\cup\mathcal{D}) \subseteq \mathcal{V}$ entails that, besides the nonzero v_j, a nonzero v_k has also to be present at at least all the nodes $k \in \mathcal{D}$.

Similarly, as shown in the remark below Proposition 5, when $\mathcal{K} = \emptyset$, Condition (ii) is respected if and only if $rank(\bar{S}_{0,\{j\}\cup\mathcal{D},\mathcal{U}}(e^{j\omega})) = 1 + n_{\mathcal{D}}$ at (almost) all ω. Among all the cases where this condition holds, we have e.g., the case where $(\{j\} \cup \mathcal{D}) \subseteq \mathcal{U}$. The latter entails that a nonzero excitation signal r_k is applied at at least all the nodes k such that $k \in \mathcal{D}$ and $k \notin \mathcal{V}$. Note that this particular choice for $\mathcal{R}$ is equivalent to the data informativity condition proposed in Theorem 2 in [6] (when this result is particularized to the case of a diagonal $\bar{H}_0(z)$), but is only one of many other possible choices to satisfy Condition (ii).

Let us now disgress a bit by comparing the above results with the results of *network identifiability* [16]. It is indeed to be noted that similar rank conditions as the ones in this section are the conditions proposed in [16] to check that a network is *identifiable* or that part of a network is *identifiable*. Let us recall this notion: the part $(\bar{G}_{0,j,\mathcal{D}}(z), H_{0,j}(z))$ of the network (1) is said *identifiable* if we can uniquely retrieve $(\bar{G}_{0,j,\mathcal{D}}(z), H_{0,j}(z))$ from the power spectrum of $\bar{S}_0(z)\bar{v}(t)$ and the closed-loop transfer matrix between $\bar{r}_{\mathcal{R}}(t)$ and $\bar{w}(t)$ (which is a part of $\bar{S}_0(z)$). Using the results of [16], a sufficient condition for this to hold is that $rank(\bar{S}_{0,\mathcal{D},\mathcal{Q}}(e^{j\omega})) = n_{\mathcal{D}}$ at (almost) all ω ($\mathcal{Q} = \mathcal{R} \cup (\mathcal{V}\backslash\{j\})$). It is clear that this condition is very close to Condition (ii) of Proposition 5 (see the remark below this proposition). Moreover, if, besides the other conditions of Proposition 5, we also suppose that $e_j(t)$ is uncorrelated with the other elements of $\bar{e}_{\mathcal{V}}(t)$, we can replace Condition (ii) by the condition that $rank(\bar{S}_{0,\mathcal{D},\mathcal{Q}}(e^{j\omega})) = n_{\mathcal{D}}$ at (almost) all ω (see Appendix E). This thus indicates that, for filtered white noises r_k, the notion of *network identifiability* is closely related to the notion of *data informativity* for the identification method of Section 3. To understand this, it is useful to observe that, when $\bar{r}_{\mathcal{R}}$ is made up of filtered white noises, there is in a way an *equivalence* between, on the one hand, the information in the measurements $\bar{w}(t)$ and, on the other hand, the information in the power spectrum of $\bar{S}_0(z)\bar{v}(t)$ and in the closed-loop transfer matrix between $\bar{r}_{\mathcal{R}}(t)$ and $\bar{w}(t)$ (see (4)). Consequently, the notion of *network identifiability* is very close to the notion of being able to uniquely retrieve $(\bar{G}_{0,j,\mathcal{D}}(z), H_{0,j}(z))$ from the measurements $\bar{w}(t)$ (i.e. the data informativity for the identification method of Section 3). Indicating this close relationship for filtered white noise excitation vectors is a contribution of the present paper.

As a final remark in this section, let us stress that all the data informativity conditions in this section are independent of the model structure $\mathcal{M}$ (i.e. they are independent of the model order) which directly indicates their conservatism. These conditions also exclude the case where the excitation vector $\bar{r}_{\mathcal{R}}(t)$ is e.g., made of multisines. These shortcomings will be alleviated by reconsidering the necessary and sufficient conditions of Proposition 3 (and of Corollary 1).

6 Using the necessary and sufficient conditions of Proposition 3

We can verify the rather complex necessary and sufficient conditions in Proposition 3/Corollary 1 by following an approach similar to the one in [4]. Indeed, Lemma 6 in [4] can be used to derive 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 scalar transfer function (corresponding to the least common factor of the denominator of all the entries of $\Delta\bar{W}(z)$). Focusing first on the condition (20), we can also derive a right factorization $N(z)V^{-1}(z)$ of $\bar{T}_{0,\mathcal{V}}(z)\bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\Xi_{0,\mathcal{V}}$ with $N(z)$ and $V(z)$ polynomial matrices. The condition (20) is thus equivalent to $\Upsilon(z)N(z) = 0 \Longrightarrow \Upsilon(z) = 0$. Using a similar approach as in Section 6 of [4], we can rewrite this condition as $\delta^T\mathcal{A} = 0 \Longrightarrow \delta = 0$ for a given matrix $\mathcal{A}$ that can be constructed with $\Upsilon(z)$ and $N(z)$ and for a real vector of coefficients δ such that $\delta = 0$ is equivalent with $\Upsilon(z) = 0$ (and thus with $\Delta\bar{W}(z) = 0$). We can thus verify the data informativity condition (20) by checking if the matrix $\mathcal{A}$ is full row rank.

Let us now consider the second equation on the left hand side of (21) and let us first suppose that $\bar{r}_{\mathcal{R}}(t)$ is filtered white noise. More particularly, let us suppose that $\bar{r}_{\mathcal{R}}(t)$ is generated as $\bar{r}_{\mathcal{R}}(t) = F(z)\bar{q}(t)$ with $F(z)$ a known matrix of transfer functions of dimension $n_{\mathcal{R}} \times n_q$ and a vector $\bar{q}(t)$ of dimension n_q such that $\Phi_{\bar{q}}(\omega) > 0$ at all frequencies ω. Then, the second equation on the left hand side of (21) can be replaced by

$$\Delta\bar{W}(z)\ \left(\bar{T}_{0,\mathcal{R}}(z) - M\right)\ F(z) = 0 \qquad (25)$$

if $F(e^{j\omega})$ is not full row rank at (almost) all ω, while this second equation can be replaced with $\Delta\bar{W}(z)\ \left(\bar{T}_{0,\mathcal{R}}(z) - M\right) = 0$ if $F(e^{j\omega})$ is full row rank at (almost) all ω. Let us focus on the most general condition (25) in the sequel. Using the factorization $Q^{-1}(z)\Upsilon(z)$ of $\Delta\bar{W}(z)$ and a factorization $N_2(z)V_2^{-1}(z)$ of $\left(\bar{T}_{0,\mathcal{R}}(z) - M\right)\ F(z)$, we have that (25) is equivalent to $\Upsilon(z)N_2(z) = 0$. Using the same approach discussed in the previous paragraph, we can rewrite $\Upsilon(z)N_2(z) = 0$ as $\delta^T\mathcal{B} = 0$ with the same δ as in the previous paragraph and with a matrix $\mathcal{B}$ that depends on $\Upsilon(z)$ and $N_2(z)$.

Let us now suppose that each element of $\bar{r}_{\mathcal{R}}$ is a multisine and let us show that the second equation on the left hand side of (21) can also be rewritten as $\delta^T\mathcal{B} = 0$ for a given matrix $\mathcal{B}$. Using $\bar{T}_{0,\mathcal{R}}(z)$, we can compute the vector of signals $\bar{d}(t) = (\bar{T}_{0,\mathcal{R}}(z) - M)\ \bar{r}_{\mathcal{R}}(t)$. This vector $\bar{d}(t)$ is also made up of multisines and these multisines have the general expression given in equation (3) in [4]. Using the factorization $Q^{-1}(z)\Upsilon(z)$ of $\Delta\bar{W}(z)$, the second equation of the left-hand side of (21) is equivalent to $\bar{E}(\Upsilon(z)\bar{d}(t))^2 = 0$ and, as shown in Section 7 in [4], we can also rewrite $\bar{E}(\Upsilon(z)\bar{d}(t))^2 = 0$ as $\delta^T\mathcal{B} = 0$.

If $\bar{r}_{\mathcal{R}}$ contains both filtered white noises and multisines, we can also rewrite the second equation on the left hand side of (21) as $\delta^T\mathcal{B} = 0$ for a given matrix $\mathcal{B}$. Indeed, using the approaches described in the previous two paragraphs and the fact that the multisines and the filtered white noises are independent, a matrix $\mathcal{B}_1$ can be constructed for the multisine part of $\bar{r}_{\mathcal{R}}$ and a matrix $\mathcal{B}_2$ can be constructed for the filtered white noise part of $\bar{r}_\mathcal{R}$. The matrix $\mathcal{B}$ is then given by $\mathcal{B} = (\mathcal{B}_1,\ \mathcal{B}_2)$.

Based on the above analysis, we have shown that the data informativity condition (21) is equivalent to:

$$\begin{cases} \delta^T \mathcal{A} = 0 \\ \delta^T \mathcal{B} = 0 \end{cases} \Longrightarrow \delta = 0 \tag{26}$$

and we can verify whether (26) holds by checking that the matrix $(\mathcal{A}\ \mathcal{B})$ is full row rank [5].

Note that the number of columns in $\mathcal{B}$ are related to the number of sinusoids in $\bar{r}_\mathcal{R}$ (for multisne $\bar{r}_\mathcal{R}$) or to the complexity of $F(z)$ and to how large n_q is (for filtered white noise $\bar{r}_\mathcal{R}$). 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 [4] for more details).

The matrices $\mathcal{A}$ and $\mathcal{B}$ discussed above are functions of the coefficients of the unknown matrix $\bar{S}_0(z)$ (see (16)). As mentioned in the previous section, we can nevertheless check the data informativity in a *generic* manner by replacing $\bar{S}_0(z) = (I_{N_{mod}} - \bar{G}_0(z))^{-1}$ by any stable models of $\bar{S}_0(z)$ obtained by replacing the entries of $\bar{G}_0(z)$ by any full-order models of these entries. The models of the entries of $\bar{G}_0(z)$ does not need to be accurate, but they should be full-order. Consequently, to verify the data informativity using Proposition 3, 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).

7 First numerical illustration

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

$$\bar{G}_0(z) \triangleq \begin{pmatrix} 0 & 0 & 0.5z^{-1} \\ 0.5z^{-1} & 0 & 0.5z^{-1} \\ 0.5z^{-1} & 0.5z^{-1} & 0 \end{pmatrix} \tag{27}$$

and where $\bar{H}_0(z) = I_3$ and $\Sigma_0 = diag(0, 0.1, 0)$. Let us also assume that $j = 2$ and that there is no known element in the second row of $\bar{G}_0(z)$ ($\mathcal{K} = \emptyset$) i.e., the identification procedure of Section 3 pertains to the consistent identification of $G_{0,21}(z)$ and $G_{0,23}(z)$ (since $H_{0,j} = 1$). This means that $\mathcal{D} = \{1, 3\}$, while $\mathcal{V} = \{2\}$ and $\Sigma_{0,\mathcal{V}} = 0.1$. We also observe that Assumption 3 is respected.

For this network, we will prove, via the approach of Section 6, 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 $v_2(t) = e_2(t)$. To prove this result, we observe that we are here in the conditions of Corollary 1 since $\Sigma_{0,\mathcal{V}} = 0.1 > 0$. We can thus factorize $\Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z)$ instead of $\Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{V}}(z)\ \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\ \Xi_{0,\mathcal{V}}$. For this purpose, let us observe, using (23), that $\bar{T}_{0,\mathcal{V}}(z) = (\bar{S}_{0,22}(z),\ \bar{S}_{0,12}(z),\ \bar{S}_{0,32}(z))^T$. This permutation of the second column of $\bar{S}_0(z)$ can be factorized as $N(z)V^{-1}(z)$ with:

$$N(z) = \begin{pmatrix} 1 - 0.25z^{-2} & 0.25z^{-2} & 0.5z^{-1} \end{pmatrix}^T$$

and $V(z) = (1 + 0.309z^{-1})(1 + 0.5z^{-1})(1 - 0.808z^{-1})$. Recall now that we consider here that $j = 2$. Using (27), 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 (13), the left factorization $Q^{-1}(z)\Upsilon(z)$ of $\Delta\bar{W}(z)$ is here also directly given by $Q(z) = 1$ and $\Upsilon(z) = \begin{pmatrix}0, & \delta\theta_1 z^{-1}, & \delta\theta_2 z^{-1}\end{pmatrix}$ (with $\delta\theta_1 = \theta_1 - \theta_{0,1}$ and $\delta\theta_2 = \theta_2 - \theta_{0,2}$).

Defining $\delta = (\delta\theta_1,\ \delta\theta_2)^T$, we observe, as mentioned in Section 6, that $\delta = 0$ is indeed equivalent to $\Delta\bar{W}(z) = \Upsilon(z) = 0$. Moreover, the term $\Upsilon(z)N(z)$ discussed in Section 6 can be written successively as follows:

$$\Upsilon(z)N(z) = \begin{pmatrix}0 & \delta\theta_1 z^{-1} & \delta\theta_2 z^{-1}\end{pmatrix} \begin{pmatrix} 1 - 0.25z^{-2} \\ 0.25z^{-2} \\ 0.5z^{-1} \end{pmatrix}$$

$$\Upsilon(z)N(z) = \delta^T \begin{pmatrix} 0.25z^{-3} \\ 0.5z^{-2} \end{pmatrix}$$

$$\Upsilon(z)N(z) = \delta^T \underbrace{\begin{pmatrix} 0 & 0.25 \\ 0.5 & 0 \end{pmatrix}}_{=\mathcal{A}} \begin{pmatrix} z^{-2} \\ z^{-3} \end{pmatrix}$$

Consequently, it is clear that $\Upsilon(z)N(z) = 0$ is equivalent to $\delta^T\mathcal{A} = 0$. Since $\mathcal{A}$ is full row rank, we have that $\delta^T\mathcal{A} = 0$ implies $\delta = 0$ and we have thus data informativity under the sole excitation of the unknown disturbance $v_2(t)$. 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 Corollary 1 and the procedure listed in Section 6 while it is not possible to do so using Proposition 4 ($\bar{T}_{0,\mathcal{V}}(z)$ being a permutation of the second column of $\bar{S}_0(z)$, it can therefore not be full row rank). This shows the usefulness of the procedure of Section 6 to check the necessary and sufficient conditions of Section 4.

8 Second numerical illustration

We consider here a network with $\bar{G}_0(z)$ given by (24) with $G_{0,31}(z) = \frac{0.173z^{-1}}{A_0(z)}$ and $G_{0,32}(z) = \frac{0.259z^{-1}}{A_0(z)}$,

[5] Note that this rank condition on $(\mathcal{A}\ \mathcal{B})$ is generally only a sufficient condition for (26) to hold since the vector δ does not always cover the whole vectorial space [4]. However, the introduced conservatism is generally much lower than with the sufficient conditions of Section 5 as shown in the examples of Section 7 and 8.

$G_{0,13}(z) = 0.3\ G_{0,32}(z)$ ($A_0(z) = 1 - 0.741z^{-1}$). Moreover, $\bar{v}(t) = (I_3 \otimes \frac{1}{A_0(z)})\bar{e}(t)$ with $\bar{e}(t)$ a white noise vector of covariance matrix $\Sigma_0 = diag(0, 0, 0.1)$. We observe that the network is in ARX form.

Let us assume that $j = 3$ and that $\mathcal{K} = \emptyset$ i.e., we want to identify consistently the transfer functions $G_{0,31}(z)$, $G_{0,32}(z)$ and $H_{3,0}(z)$. This means that $\mathcal{D} = \{1, 2\}$, while $\mathcal{V} = \{3\}$ and $\Sigma_{0,\mathcal{V}} = 0.1$. We observe that Assumption 3 is respected. In this network, we do not have informativity with the sole excitation of $v_3(t)$. Indeed, $w_2(t)$ is equal to zero when $r_2(t) = v_2(t) = 0$, and it will therefore be impossible to identify $G_{0,32}(z)$. However we will prove that we obtain data informativity by e.g., adding to Node 2 a signal $r_2(t) = cos(\omega_0 t)$ with an arbitrary frequency ω_0, say $\omega_0 = 0.1$. Let us prove this using the procedure of Section 6.

Since $\mathcal{K} = \emptyset$, we have, using (23), that $\bar{T}_{0,\mathcal{V}}(z) = (\bar{S}_{0,33}(z),\ \bar{S}_{0,13}(z),\ \bar{S}_{0,23}(z))^T$. This permutation of the third column of $\bar{S}_0(z)$ can be factorized as $N(z)V^{-1}(z)$ with:

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

and $V(z) = 1 - 1.482z^{-1} + 0.5354z^{-2}$. Since $j = 3$, a model structure $\mathcal{M}$ satisfying Assumption 3 is described by $H_3(z, \theta) = 1/(1 + az^{-1})$ and:

$$\bar{G}_{3,\mathcal{D}}(z, \theta) = \begin{pmatrix} \frac{b_1 z^{-1}}{1 + az^{-1}} & \frac{b_2 z^{-1}}{1 + az^{-1}} \end{pmatrix}$$

We have thus $\theta = (a, b_1, b_2)^T$. Using (13), the left factorization $Q^{-1}(z)\Upsilon(z)$ of $\Delta \bar{W}(z)$ is here given by $Q(z) = 1$ and $\Upsilon(z) = \left(\delta a z^{-1},\ \delta b_1 z^{-1},\ \delta b_2 z^{-1}\right)$. Defining $\delta = (\delta a,\ \delta b_1,\ \delta b_2)^T$, the term $\Upsilon(z)N(z)$ discussed in Section 6 has the following expression:

$$\Upsilon(z)N(z) = \delta^T \underbrace{\begin{pmatrix} 1 & 1.482 & 0.5488 \\ 0 & 0.0778 & -0.0576 \\ 0 & 0 & 0 \end{pmatrix}}_{=\mathcal{A}} \begin{pmatrix} z^{-1} \\ z^{-2} \\ z^{-3} \end{pmatrix} \qquad (28)$$

Since the matrix $\mathcal{A}$ is not full row rank [6], we cannot infer data informativity with the sole excitation of $v_3(t)$, which, as mentioned above, is an expected result.

Let us thus compute the matrix $\mathcal{B}$ when we have $r_2(t) = cos(\omega_0 t)$ and $r_1(t) = r_3(t) = 0$ (i.e. $\mathcal{R} = \{2\}$). Consequently, $\bar{d}(t) = (\bar{T}_{0,\mathcal{R}}(z) - M)\ r_2(t) = \bar{T}_{0,\mathcal{R}}(z)\ r_2(t)$ (since $r_{j=3}(t) = 0$). Using (23), $\bar{T}_{0,\mathcal{R}}(z) = (\bar{S}_{0,32}(z),\ \bar{S}_{0,12}(z),\ \bar{S}_{0,22}(z))^T$ i.e., a permutation of the second column of $\bar{S}_0(z)$. Let us compute the frequency response of $\bar{T}_{0,\mathcal{R}}(z)$ for $\omega_0 = 0.1$: $\bar{T}_{0,\mathcal{R}}(e^{j\omega_0}) = (0.94 - 0.53\,j,\ 0.19 - 0.24\,j,\ 1)^T$ Let us now rewrite the term $\Upsilon(z)\bar{d}(t)$ using $\delta = (\delta a,\ \delta b_1,\ \delta b_2)^T$: $\Upsilon(z)\bar{d}(t) = \delta^T(d_1(t-1),\ d_2(t-1),\ d_3(t-1))^T$. Using Euler formula and the expression of $\bar{T}_{0,\mathcal{R}}(e^{j\omega_0})$, we can write:

$$\Upsilon(z)\bar{d}(t) = \delta^T\ \mathcal{B}\ \begin{pmatrix} 0.5\ e^{j\omega_0 t} \\ 0.5\ e^{-j\omega_0 t} \end{pmatrix}$$

$$\mathcal{B} = \begin{pmatrix} (0.94 - 0.53\ j)\ e^{-j\omega_0} & (0.94 + 0.53\ j)\ e^{j\omega_0} \\ (0.19 - 0.24\ j)\ e^{-j\omega_0} & (0.19 + 0.24\ j)\ e^{j\omega_0} \\ e^{-j\omega_0} & e^{j\omega_0} \end{pmatrix}$$

Using this expression, $\bar{E}(\Upsilon(z)\bar{d}(t))^2 = 0$ is equivalent to $\delta^T\mathcal{B} = 0$ (see also Section 7 of [4]).

Let us now consider $(\mathcal{A}\ \mathcal{B})$ to asses the data informativity. Since the rank of the matrix $(\mathcal{A}\ \mathcal{B})$ is equal to 3, 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 here also 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 $\mathcal{B}$. Using the procedure of Section 6, we can also prove that we have data informativity when r_2 is filtered white noise.

It is to be noted that we cannot use the conditions in Section 5 to derive the above results. Indeed, those results do not pertain to multisines and, even when r_2 is filtered white noise, the data informativity can also not be inferred for $\mathcal{R} = \{2\}$ from Propositions 5 or 6. Indeed, since $1 + n_{\mathcal{D}} = 3$ and $n_{\mathcal{U}} = 2$, $\bar{T}_{0,\mathcal{U}}(z)$ is a matrix of dimension 3×2 and thus Condition (ii) is not satisfied. Moreover, since $n_{\mathcal{D}} = 2 > n_{\mathcal{R}} = 1$, Condition (v) is also not satisfied. We can however infer the data informativity from these sufficient results when $\mathcal{R} = \{1, 2\}$ or $\mathcal{R} = \{2, 3\}$ and when the two excitation signals in $\bar{r}_{\mathcal{R}}$ are mutually independent filtered white noises (to be sure to satisfy both Conditions (iii) and (iv)). Let us first consider $\mathcal{R} = \{1, 2\}$. For this $\mathcal{R}$, we have that $\mathcal{U} = \mathcal{R} \cup \mathcal{V} = \{j\} \cup \mathcal{D}$. Using (23), we have thus that $rank(\bar{T}_{0,\mathcal{U}}(e^{j\omega})) = rank(\bar{S}_{0,\mathcal{U},\mathcal{U}}(e^{j\omega}))$. From the discussion below Example 1 in Section 5, it is clear that $\bar{T}_{0,\mathcal{U}}(e^{j\omega})$ is full row rank at all ω (i.e. Condition (ii) is respected). Let us now consider $\mathcal{R} = \{2, 3\}$. Since $\mathcal{U} = \mathcal{R} \cup \mathcal{V} = \mathcal{R}$, we have that $\bar{T}_{0,\mathcal{U}}(e^{j\omega})$ is a matrix

[6] If we do not wish to identify $G_{0,32}(z)$ (e.g., because this transfer function is known i.e., $\mathcal{K} = \{2\}$), we then have informativity with the sole excitation of $v_3(t)$. Indeed, it can be proven that, in this case, $\mathcal{A}$ reduces to the first two rows of the matrix $\mathcal{A}$ given in (28) and this matrix is full row rank.

of dimension 3×2 and Condition (ii) cannot be satisfied. However, we can nevertheless infer the data informativity for $\mathcal{R} = \{2, 3\}$ from Proposition 6 since $rank(\bar{S}_{0,\mathcal{D},\mathcal{R}}(e^{j\omega})) = n_{\mathcal{D}} = 2$ at all frequencies. The latter can be checked by computing the rank of $\bar{S}_{0,\mathcal{D},\mathcal{R}}(e^{j\omega})$ at each frequency (we have here access to $\bar{S}_0$) or by observing that there are indeed two vertex-disjoint pathes from $\mathcal{R} = \{2, 3\}$ to $\mathcal{D} = \{1, 2\}$ (see Example 1 in Section 5).

Note that the choice $\mathcal{R} = \{1, 3\}$ does not lead to data informativity. Indeed, for this choice of $\mathcal{R}$, $w_2(t)$ is equal to zero ($r_2(t) = v_2(t) = 0$) and it will be impossible to identify $G_{0,32}(z)$.

9 Optimal experiment design

Using the data informativity conditions of Proposition 3, 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 $\bar{r}_{\mathcal{R}}(t)$ (multisines, filtered white noises, ...). 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 use optimal experiment design to select an optimal $\mathcal{R}$ as well as an optimal signal vector $\bar{r}_{\mathcal{R}}(t)$.

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 $\psi_j(t, \theta) = \frac{d\epsilon_j(t,\theta)}{d\theta}$. Our initial goal was to identify a model of $G_{0,ji}(z) = G_{ji}(z, \theta_0) = G_{ji}(z, \theta_{0,ji})$ where $\theta_{0,ji}$ is a part of θ_0. This means that we can always write $\theta_{0,ji} = S\ \theta_0$ for some matrix S. Let us also define define $P_{\theta,ji}$ as the covariance matrix of $S\ \hat{\theta}_N$. We have then $P_{\theta,ji} = SP_\theta S^T$.

We want to determine the excitation pattern $\mathcal{R}$ that, for an identification experiment of duration N, yields an acceptable covariance matrix $P_{\theta,ji}$ 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 $P_{\theta,ji} \leq R_{adm}$ where R_{adm} specifies the desired accuracy (a diagonal R_{adm} e.g., allows to constrain the standard deviations of each entries of $S\ \hat{\theta}_N$ [9]). We thus require $R_{adm} - SP_\theta S^T \geq 0$ and, using Schur complement, this gives the following optimal experiment design problem:

$$\min_{\Phi_{\bar{r}}(\omega)} \ trace \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{\bar{r}}(\omega)\ d\omega \right)$$
$$\text{subject to} \begin{pmatrix} R_{adm} & S \\ S^T & P_\theta^{-1} \end{pmatrix} \geq 0 \tag{29}$$

This optimization problem is convex since, as will be shown in the sequel, P_θ^{-1} is an affine function of $\Phi_{\bar{r}}(\omega)$. Note also that the objective function of the optimization problem (29) has a l_1-norm structure and it is frequently observed that such objective functions, when minimized under convex constraints, generate a sparse solution (see e.g., [15]). Consequently, we can expect that the optimal excitation vector $\bar{r}(t)$ will have some elements r_k equal to zero, determining in this way the optimal excitation pattern $\mathcal{R}_{opt}$.

Let us now derive the affine relation between $\Phi_{\bar{r}}(\omega)$ and P_θ^{-1}. Using the philosophy introduced in [8,2] and denoting the dimension of θ by n_θ, 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 $H^{-1}_{0,j}(z)\frac{d\bar{G}_{j,\mathcal{D}}(z,\theta)}{d\theta_l}$ (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 $H^{-1}_{0,j}(z)\frac{dH_j(z,\theta)}{d\theta_l}$. 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,

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

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 (29) [12,3,1]. 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 (29). ■

Numerical illustration. Let us consider the same network as in Section 7. For that network, as proposed above, we will solve (29) 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. Suppose that we are interested in identifying $G_{0,21}(z) = 0.5z^{-1}$ (i.e., $j = 2$ and $i = 1$). As shown in Section 7, the model structure $\mathcal{M}$ considered for this identification is $\mathcal{M} = \{\bar{G}_{2,\mathcal{D}}(z, \theta)\}$ with $\bar{G}_{2,\mathcal{D}}(z, \theta) = (\theta_1 z^{-1},\ \theta_2 z^{-1})$ $(\theta = (\theta_1, \theta_2)^T)$. We suppose that the maximal allowed variance for the estimate $\hat{\theta}_1^N$ of the coefficient $\theta_{0,1} = 0.5$ is $R_{adm} = 10^{-5}$. In Section 7, 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 $\Phi_{\bar{r}}^{opt}(\omega) = diag(9.36, 0, 0)$ which corresponds to a unique excitation signal on Node $i = 1$. Consequently, the optimal excitation pattern is $\mathcal{R}_{opt} = \{i\} = \{1\}$.

If we are interested in identifying $G_{0,23}(z) = 0.5z^{-1}$ ($j = 2$ and $i = 3$), we observe the same phenomenon since the optimal spectrum is given by $\Phi_{\bar{r}}^{opt}(\omega) = diag(0, 0, 7.27)$ i.e., $\mathcal{R}_{opt} = \{i\} = \{3\}$.

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

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] M. Gevers, A. Bazanella, and G. Pimentel. Identifiability of dynamical networks with singular noise spectra. *IEEE Transactions on Automatic Control*, 64(6):2473–2479, 2018.

[8] 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.

[9] 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.

[10] J. Gonçalves and S. Warnick. Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks. *IEEE Transactions on Automatic Control*, 53(7):1670–1674, 2008.

[11] 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.

[12] 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.

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

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

[15] 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.

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

A Comparison of the MISO approach with the MIMO approach

Let $\bar{G}(z)$ denotes a model of the matrix $\bar{G}_0(z)$ (see (2)) and $\bar{H}(z)$ a model for the diagonal matrix $\bar{H}_0(z)$ (see (3)). Then, the prediction error $\bar{\epsilon}(t)$ for the one-step ahead predictor corresponding to the MIMO data generating system (1) is given by [13]:

$$\bar{\epsilon}(t) = \bar{H}^{-1}(z)\ \left(\bar{w}(t) - \bar{r}(t) - \bar{G}(z)\bar{w}(t)\right) \qquad \text{(A.1)}$$

Indeed, when $\bar{G}(z) = \bar{G}_0(z)$ and $\bar{H}(z) = \bar{H}_0(z)$, $\bar{\epsilon}(t) = \bar{e}(t)$. Now, under our assumption of a diagonal $\bar{H}_0$ and thus of a diagonal $\bar{H}(z)$, we observe that the j^{th} element of $\bar{\epsilon}(t)$ is equivalent to (8).

Let us first suppose that $Ee_j(t)e_k(t) = 0$ for all $k \neq j$. Then, the standard MIMO prediction error criterion [13] with the prediction error (A.1) is the sum of the criterion $\frac{1}{N}\sum_{t=1}^{N} \epsilon_j^2(t)$ in (7)-(8) and of a criterion pertaining to the models of the rows $k \neq j$ of $\bar{G}_0(z)$ and $\bar{H}_0(z)$. Consequently, if we suppose that all entries of $\bar{G}$ and $\bar{H}$ are independently parametrized, besides the fact that (7)-(8) is a much simpler identification problem, there is also no disadvantage from an accuracy point-of-view to use (7)-(8) instead of the MIMO criterion to obtain a model of $\bar{G}_{0,j,\mathcal{D}}(z)$ and of $H_{0,j}(z)$.

When e_j is correlated with at least one of the other e_k $(k \neq l)$ in $\bar{e}(t)$, the low complexity of the MISO approach (7)-(8) is still preferred, although then some accuracy will be lost with respect to the MIMO criterion.

B Data informativity when $j \notin \mathcal{V}$

The case $j \notin \mathcal{V}$ corresponds to $v_j(t) = e_j(t) = 0$. Consequently, (6) is here given by $y_j(t) = \bar{G}_{0,j,\mathcal{D}}(z)\bar{w}_{\mathcal{D}}(t)$. The identification criterion (7) can thus be considered with $\epsilon_j(t,\theta) = y_j(t) - \bar{G}_{j,\mathcal{D}}(z,\theta)\bar{w}_{\mathcal{D}}(t)$ $(H_j(z,\theta) = 1)$. The true parameter vector is obviously a minimum of $\bar{E}\epsilon_j^2(t,\theta)$ since $\bar{E}\epsilon_j^2(t,\theta_0) = 0$. It is to be noted that, as opposed to the case where $j \in \mathcal{V}$, the latter property holds even if there exist τ such that $E\bar{e}(t)\bar{e}^T(t-\tau) \neq 0$ and even if the delay condition in the statement of Proposition 1 is not satisfied.

On the other hand, like in the case where $j \in \mathcal{V}$, the minimum θ_0 of $\bar{E}\epsilon_j^2(t,\theta)$ is unique if the data $\bar{x}(t)$ are informative with respect to $\mathcal{M} = \{\bar{G}_{j,\mathcal{D}}(z,\theta)\}$ (see Definition 1). In the left hand side of (15), $\Delta\bar{W}(z)\bar{x}(t)$ is here given by $(\bar{G}_{j,\mathcal{D}}(z,\theta) - \bar{G}_{j,\mathcal{D}}(z,\theta_0))\bar{w}_{\mathcal{D}}(t)$ with $\bar{w}_{\mathcal{D}}(t) = \bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t) + \bar{S}_{0,\mathcal{D},\mathcal{V}}(z)\bar{v}_{\mathcal{V}}(t)$. Consequently, using the shorthand notation $\Delta\bar{G}_{j,\mathcal{D}}(z) = \bar{G}_{j,\mathcal{D}}(z,\theta) - \bar{G}_{j,\mathcal{D}}(z,\theta_0)$, the necessary and sufficient condition (21) for data informativity becomes:

$$\begin{cases} \Delta\bar{G}_{j,\mathcal{D}}(z)\ \bar{S}_{0,\mathcal{D},\mathcal{V}}(z)\ \bar{H}_{0,\mathcal{V},\mathcal{V}}(z)\ \Xi_{0,\mathcal{V}} = 0 \\ \bar{E}\left(\Delta\bar{G}_{j,\mathcal{D}}(z)\ \bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\ \bar{r}_{\mathcal{R}}(t)\right)^2 = 0 \end{cases} \Longrightarrow \Delta\bar{G}_{j,\mathcal{D}}(z) = 0$$

C Proof of Proposition 5

Let us construct, using $\bar{r}(t)$, the vector of signals $\bar{r}^{bis}(t)$ as: $\bar{r}^{bis}(t) = (r_1(t), ..., r_{j-1}(t), 0, r_{j+1}(t), ..., r_{N_{mod}}(t))^T$. We can then rewrite (17) as $\Delta\bar{W}(z)\bar{x}(t) = s_{r_j}(t) + s_{comp}(t)$ with

$$s_{r_j}(t) = \Delta\bar{W}(z)\left(\bar{T}_{0,\{j\}}(z) - (1, 0..., 0)^T\right) r_j(t)$$

$$s_{comp}(t) = \Delta\bar{W}(z)\ \bar{T}_{0,\mathcal{U}}(z)\ \bar{\xi}(t) \tag{C.1}$$

where $\mathcal{U} = (\mathcal{R}\backslash\{j\}) \cup \mathcal{V} = \mathcal{R} \cup \mathcal{V}$ (recall that $j \in \mathcal{V}$ due to Assumption 3) and where $\bar{\xi}(t)$ is a vector of dimension $n_{\mathcal{U}}$ containing the non-zero elements of $\bar{r}^{bis}(t) + \bar{v}(t)$. Since $s_{r_j}(t)$ and $s_{comp}(t)$ are independent due to Condition (iv), the left hand side of (15) is equivalent to:

$$\begin{cases} \bar{E}s_{r_j}^2(t) = 0 \\ \bar{E}s_{comp}^2(t) = 0 \end{cases} \tag{C.2}$$

and, similarly as in the proof of Proposition 4, we have thus to prove that, under the conditions of Proposition 5, (C.2) implies that $\Delta\bar{W}(z) = 0$. For this purpose, let us analyze the properties of the power spectrum matrix $\Phi_{\bar{\xi}}(w)$ of $\bar{\xi}(t)$. The elements of the vector $\bar{\xi}(t)$ are equal to v_k (if $k \in \mathcal{V}$ and $k \notin \mathcal{R}\backslash\{j\}$), to $v_k + r_k$ (if $k \in \mathcal{V}$ and $k \in \mathcal{R}\backslash\{j\}$) or to r_k (if $k \notin \mathcal{V}$ and $k \in \mathcal{R}\backslash\{j\}$). Due to $\Sigma_{0,\mathcal{V}} > 0$, Condition (iii) and the fact that $\bar{r}(t)$ and $\bar{e}(t)$ are uncorrelated (see Assumption 1), the elements of $\bar{\xi}(t)$ are thus linearly independent and we have thus that $\Phi_{\bar{\xi}}(\omega) > 0$ at (almost) all ω. This together with Condition (ii) also implies that the power spectrum matrix of the vector $\bar{T}_{0,\mathcal{U}}(z)\bar{\xi}(t)$ in (C.1) is strictly positive definite at (almost) all ω. Consequently, under the conditions of Proposition 5, we have that the second equation of (C.2) implies that $\Delta\bar{W}(z) = 0$. This concludes the proof.

D Proof of Proposition 6

An alternative formulation of Definition 1 is that the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are informative wrt. $\mathcal{M}$ if and only if we have (11) whenever $\bar{E}(\epsilon_j(t,\theta) - \epsilon_j(t,\theta_0))^2 = 0$ for a given $(\bar{G}_{j,\mathcal{D}}(z,\theta),\ \ H_j(z,\theta)) \in \mathcal{M}$. Due to (9), $\epsilon_j(t,\theta) - \epsilon_j(t,\theta_0) = s_1(t,\theta) + s_2(t,\theta)$ and $\bar{E}(\epsilon_j(t,\theta) - \epsilon_j(t,\theta_0))^2 = \bar{E}s_1^2(t,\theta) + \bar{E}s_2^2(t,\theta)$. Consequently, the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are informative wrt. $\mathcal{M}$ if and only if we have (11) whenever $\bar{E}s_1^2(t,\theta) = \bar{E}s_2^2(t,\theta) = 0$ for a given $(\bar{G}_{j,\mathcal{D}}(z,\theta),\ \ H_j(z,\theta)) \in \mathcal{M}$. Let us consider the expression for $s_1(t,\theta)$. Due to Conditions (iii) and (v), the power spectrum matrix of $\bar{S}_{0,\mathcal{D},\mathcal{R}}(z)\bar{r}_{\mathcal{R}}(t)$ is strictly positive definite at (almost) all frequencies. Consequently, under the conditions of Proposition 6, we have that $\bar{E}s_1^2(t,\theta) = 0$ implies $\bar{G}_{0,j,\mathcal{D}}(z) - \bar{G}_{j,\mathcal{D}}(z,\theta) = 0$. This identity and the fact that $\bar{E}s_2^2(t,\theta) = 0$ leads to $H_{0,j}(z) - H_j(z,\theta) = 0$ show that the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are indeed informative wrt. $\mathcal{M}$ under the conditions of Proposition 6.

E Corollary to Proposition 5

Corollary 2 *Consider the framework of Proposition 5 and suppose that $e_j(t)$ is uncorrelated with the other elements of $\bar{e}_{\mathcal{V}}(t)$. Then, the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are informative wrt. $\mathcal{M}$ if Conditions (iii) and (vi) are both satisfied:*

(vi) the set $\mathcal{R}$ describing the nodes where a 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, $rank(\bar{S}_{0,\mathcal{D},\mathcal{Q}}(e^{j\omega})) = n_{\mathcal{D}}$. ■

Proof. Let us recall that the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are informative wrt. $\mathcal{M}$ if and only if we have (11) whenever $\bar{E}(\epsilon_j(t,\theta) - \epsilon_j(t,\theta_0))^2 = 0$ for a given $(\bar{G}_{j,\mathcal{D}}(z,\theta),\ \ H_j(z,\theta)) \in \mathcal{M}$. Let us also construct, using $\bar{v}(t)$, the vector of signals $\bar{v}^{bis}(t)$ as: $\bar{v}^{bis}(t) = (v_1(t), ..., v_{j-1}(t), 0, v_{j+1}(t), ..., v_{N_{mod}}(t))^T$. If we use (9), we can rewrite $\epsilon_j(t,\theta)$ as $\epsilon_j(t,\theta) - \epsilon_j(t,\theta_0) = s_3(t,\theta) + s_4(t,\theta)$ with:

$$s_3(t,\theta) = \frac{\Delta\bar{G}_{j,\mathcal{D}}(z,\theta)}{H_j(z,\theta)}\ \bar{S}_{0,\mathcal{D},\mathcal{Q}}(z)\ \bar{\rho}(t)$$

$$s_4(t,\theta) = \frac{\Delta H_j(z,\theta)}{H_j(z,\theta)} e_j(t) + \frac{\Delta\bar{G}_{j,\mathcal{D}}(z,\theta)}{H_j(z,\theta)} \bar{S}_{0,\mathcal{D},j}(z) v_j(t)$$

where $\mathcal{Q} = \mathcal{R} \cup (\mathcal{V}\backslash\{j\})$ and where $\bar{\rho}(t)$ is a vector of dimension $n_{\mathcal{Q}}$ containing the non-zero elements of $\bar{r}(t) + \bar{v}^{bis}(t)$. Since e_j is independent of the other elements of

$\bar{e}_{\mathcal{V}}(t)$, we have that $s_3(t,\theta)$ and $s_4(t,\theta)$ are independent and therefore that $\bar{E}(\epsilon_j(t,\theta)-\epsilon_j(t,\theta_0))^2 = \bar{E}s_3^2(t,\theta) + \bar{E}s_4^2(t,\theta)$. Consequently, the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are informative wrt. $\mathcal{M}$ if and only if we have (11) whenever $\bar{E}s_3^2(t,\theta) = \bar{E}s_4^2(t,\theta) = 0$ for a given $(\bar{G}_{j,\mathcal{D}}(z,\theta), H_j(z,\theta)) \in \mathcal{M}$. To prove that this holds under the conditions of Corollary 2, let us first analyze the properties of the power spectrum matrix $\Phi_{\bar{\rho}}(w)$ of $\bar{\rho}(t)$. The elements of the vector $\bar{\rho}(t)$ 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 $k \in \mathcal{R}$). Due to $\Sigma_{0,\mathcal{V}} > 0$, Condition (iii) and the fact that $\bar{r}(t)$ and $\bar{e}(t)$ are uncorrelated (see Assumption 1), the elements of $\bar{\rho}(t)$ are thus linearly independent and we have thus that $\Phi_{\bar{\rho}}(\omega) > 0$ at (almost) all ω. This together with Condition (vi) also implies that the power spectrum matrix of the vector $\bar{S}_{0,\mathcal{D},\mathcal{Q}}(z)\bar{\rho}(t)$ in the expression of $s_3(t,\theta)$ is strictly positive definite at (almost) all ω. Consequently, under the conditions of Corollary 2, we have that $\bar{E}s_3^2(t,\theta) = 0$ implies $\bar{G}_{0,j,\mathcal{D}}(z) - \bar{G}_{j,\mathcal{D}}(z,\theta) = 0$. This identity and the fact that $\bar{E}s_4^2(t,\theta) = 0$ leads to $H_{0,j}(z) - H_j(z,\theta) = 0$ show that the data $\bar{x}(t) = (y_j,\ \bar{w}_{\mathcal{D}}^T(t))^T$ are indeed informative wrt. $\mathcal{M}$ under the conditions of Corollary 2. ■