



HAL
open science

Optimal identification experiment design for the interconnection of locally controlled systems

Xavier Bombois, Anton Korniienko, Håkan Hjalmarsson, Gérard Scorletti

► **To cite this version:**

Xavier Bombois, Anton Korniienko, Håkan Hjalmarsson, Gérard Scorletti. Optimal identification experiment design for the interconnection of locally controlled systems. *Automatica*, 2018, 89 (March), pp.169-179. 10.1016/j.automatica.2017.12.014 . hal-01492050

HAL Id: hal-01492050

<https://hal.science/hal-01492050>

Submitted on 20 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimal identification experiment design for the interconnection of locally controlled systems^{*}

X. Bombois^a, A. Korniienko^a, H. Hjalmarsson^b, G. Scorletti^a

^a*Laboratoire Ampère, Ecole Centrale de Lyon, 36 avenue Guy de Collongue, 69134 Ecully Cedex, France*

^b*Automatic Control, School of Electrical Engineering, KTH, 100 44 Stockholm, Sweden*

Abstract

This paper considers the identification of the modules of a network of locally controlled systems (multi-agent systems). Its main contribution is to determine the least perturbing identification experiment that will nevertheless lead to sufficiently accurate models of each module for the global performance of the network to be improved by a redesign of the decentralized controllers. Another contribution is to determine the experimental conditions under which sufficiently informative data (i.e. data leading to a consistent estimate) can be collected for the identification of any module in such a network.

Key words: Experiment Design, Identification for Control, Interconnected systems.

1 Introduction

In this paper, we consider the problem of designing an identification experiment that will allow to improve the global performance of a network made up of the interconnection of locally controlled systems. The identification experiment will be designed in such a way that we obtain a sufficiently accurate model of each module in the network to be able to improve the global performance of the network by redesigning the local controllers. The type of networks considered in this paper is usual in the literature on multi-agent systems (see e.g. [10,19]).

This paper contributes to the efforts of developing techniques for the identification of large-scale or interconnected systems when the topology of the network is known. In many papers, the problem is seen as a multivariable identification problem and structural properties of the system are then used to simplify this complex problem (see e.g. [15]). The identifiability of the multivariable structure is studied in a prediction error context in [26] while this multivariable structure is exploited in other papers to reduce the variance of a given module in the network (see [16,14,9]). Unlike most of these papers, we consider here a network whose interconnection is realized by exchanging the measured (and thus noisy) output of neighbouring modules. Another important difference is that, in our setting, all modules can be identified independently using single-input single-output identification. Consequently, we are close to the situation considered in our preceding papers on dynamic network identification (see e.g. [6]). In these contributions, we have developed conditions for consistent estimation of one given module in a dynamic network. Since general networks were considered in these contributions, the data informativity was tackled with a classical condition on the positivity of the spectral density matrix [21]. The first contribution of this paper is to extend these results for the considered type of networks by giving specific conditions for data informativity. In particular, we show that it is not necessary to excite a specific module i to consistently identify it as long as there exists at least one path from another module j to that particular module i . In this case, the noise present in the noisy output measurement y_j will give sufficient excitation for consistent estimation.

However, the main contribution of this paper is to tackle the problem of optimal experiment design for (decentralized) control in a network context. More precisely, our contribution lies in the design of the identification experiment that will lead to sufficiently accurate models of each module of the network to guarantee a certain level of global

^{*} This paper was not presented at any IFAC meeting. Corresponding author X. Bombois.

Email addresses: xavier.bombois@ec-lyon.fr (X. Bombois), anton.korniienko@ec-lyon.fr (A. Korniienko), hakan.hjalmarsson@ee.kth.se (H. Hjalmarsson), gerard.scorletti@ec-lyon.fr (G. Scorletti).

performance via the design of local controllers. The identification experiment consists of simultaneously applying an excitation signal in each module (i.e. in each closed-loop system) and our objective is to design the spectra of each of these excitations signals in such a way that the global control objective is achieved with the least total injected power. In this sense, we extend the results in [4,1] considering one local loop with a local performance objective to the case of network of closed-loop systems with (both a local and) a global performance objectives. Like in [4,1], the uncertainty of an identified model will be represented via its covariance matrix. The difference is that this covariance matrix will here be a function of the excitation signals injected in each module that has a path to the considered module and of course that there will be a covariance matrix per identified module. Like in [4,1], the maximal allowed uncertainty will be determined using tools from robustness analysis. To avoid heavy computational loads linked to a high number of modules N_{mod} and to structured uncertainties characterized by N_{mod} uncertain parameter vectors, the uncertainty is first projected into an unstructured uncertainty on the complementary sensitivity describing each connected closed-loop system and then the robustness analysis is based on the interconnection of these unstructured uncertainties. This approach (called hierarchical approach) to analyze the robustness of large-scale (interconnected) systems has been introduced in [23] and further developed in [7]. A technical contribution of this paper is to develop a methodology that allows the use of the hierarchical approach in the presence of the nonstandard uncertainty delivered by system identification.

Note that the framework considered here is much different than the frameworks of [25,17] which is, to our knowledge, the only other papers treating the optimal experiment design problem in a network. In [25], the authors consider input design for nonparametric identification of static nonlinearities embedded in a network. The main purpose of [17] lies in the use of measurable disturbances in optimal experiment design.

Notations. The matrix

$$\begin{pmatrix} X_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & X_N \end{pmatrix}$$

will be denoted $diag(X_1, \dots, X_N)$ if the elements X_i ($i = 1, \dots, N$) are scalar quantities while it will be denoted $bdiag(X_1, \dots, X_N)$ if the elements X_i ($i = 1, \dots, N$) are vectors or matrices.

2 Identification of interconnected systems

2.1 Description of the network configuration

We consider a network made up of N_{mod} single-input single-output (SISO) systems \mathcal{S}_i ($i = 1 \dots N_{mod}$) operated in closed loop with a SISO decentralized controller K_i ($i = 1 \dots N_{mod}$):

$$\mathcal{S}_i : y_i(t) = G_i(z, \theta_{i,0})u_i(t) + v_i(t) \quad (1)$$

$$u_i(t) = K_i(z)(y_{ref,i} - y_i(t)) \quad (2)$$

$$\bar{y}_{ref}(t) = \mathcal{A} \bar{y}(t) + \mathcal{B} ref_{ext}(t) \quad (3)$$

Let us describe these equations in details. The signal u_i is the input applied to the system \mathcal{S}_i and y_i is the measured output. This output is made up of a contribution of the input u_i and of a disturbance term $v_i(t) = H_i(z, \theta_{i,0})e_i(t)$ that represents both process and measurement noises. The different systems are thus described by two stable transfer functions $G_i(z, \theta_{i,0})$ and $H_i(z, \theta_{i,0})$, the later being also minimum-phase and monic. The signals e_i ($i = 1 \dots N_{mod}$) defining v_i are all white noise signals. Moreover, the vector $\bar{e} \triangleq (e_1, e_2, \dots, e_{N_{mod}})^T$ has the following property:

$$\begin{aligned} E\bar{e}(t)\bar{e}^T(t) &= \Lambda \\ E\bar{e}(t)\bar{e}^T(t - \tau) &= 0 \quad \text{for } \tau \neq 0 \end{aligned} \quad (4)$$

with E the expectation operator and with Λ a strictly positive definite matrix. With (4), the power spectrum $\Phi_{\bar{e}}(\omega)$ of \bar{e} is given by $\Phi_{\bar{e}}(\omega) = \Lambda$ for all ω . We will further assume that the signals e_i ($i = 1 \dots N_{mod}$) are mutually independent. The matrix Λ is then diagonal¹ i.e. $\Lambda = diag(\Lambda_{1,1}, \Lambda_{2,2}, \dots, \Lambda_{N_{mod}, N_{mod}}) > 0$.

The systems \mathcal{S}_i in (1) may all represent the same type of systems (e.g. drones). However, due to industrial dispersion, the unknown parameter vectors $\theta_{i,0} \in \mathbf{R}^{n_{\theta_i}}$ can of course be different for each i , as well as the order of the transfer functions G_i and H_i . Consequently, it will be necessary to identify a model for each of the systems \mathcal{S}_i in the sequel.

¹ We will nevertheless see in the sequel that many of the results of this paper also apply to the case of spatially-correlated noises e_i i.e. to the case where (4) holds with a matrix $\Lambda = \Lambda^T > 0$ that is not necessarily diagonal.

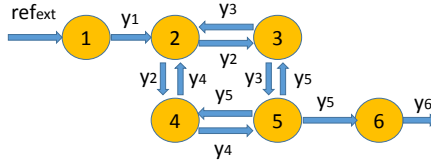


Fig. 1. Example of a network

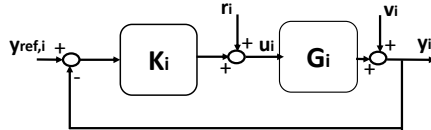


Fig. 2. Local closed-loop system (see (1)-(2)). The signal r_i is used for identification purpose (see (5)).

In this paper, we consider the type of interconnections used in formation control or multi-agent systems (see e.g. [10,19]). As shown in (2), each system \mathcal{S}_i is operated with a decentralized controller $K_i(z)$. In (2), the signal $y_{ref,i}$ is a reference signal that will be computed via (3). The matrix \mathcal{A} and the vector \mathcal{B} in (3) represent the interconnection (flow of information) in the network and we have $\bar{y}_{ref} = (y_{ref,1}, y_{ref,2}, \dots, y_{ref,N_{mod}})^T$ and $\bar{y} = (y_1, y_2, \dots, y_{N_{mod}})^T$. The signal ref_{ext} is a (scalar) external reference signal that should be followed by all outputs y_i and that is generally only available at one node of the network.

As an example, let us consider the network in Figure 1. In this network, we have $N_{mod} = 6$ systems/modules, all of the form (1) and all operated as in (2) with a decentralized controller K_i . These local closed loops are represented by a circle/node in Figure 1 and are further detailed in Figure 2 (consider $r_i = 0$ for the moment in this figure). The objective of this network is that the outputs y_i of all modules follow the external reference ref_{ext} even though this reference is only available at Node 1. For this purpose, a number of nodes are allowed to exchange information (i.e. their measured output) with some other neighbouring nodes. The arrows between the nodes in Figure 1 indicate the flow of information. For example, Node 5 receives the output of two nodes (i.e. Nodes 3 and 4) and sends its output (i.e. y_5) to three nodes (Nodes 3, 4 and 6). The reference signal $y_{ref,i}$ of Node i will be computed as a linear combination of the received information at Node i . For Node 5, $y_{ref,5}$ will thus be a linear combination of y_3 and y_4 . More precisely, for all outputs y_i to be able to follow the external reference ref_{ext} , \mathcal{A} and \mathcal{B} in (3) are chosen as [10,19]:

$$\mathcal{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathcal{B} = (1, 0, \dots, 0)^T.$$

The matrix \mathcal{A} is called the normalized adjacency matrix in the literature [10]. Using (3), we, e.g., see that the tracking error signals $y_{ref,1} - y_1$ and $y_{ref,2} - y_2$ of Nodes 1 and 2 are respectively given by $ref_{ext} - y_1$ and $1/3((y_1 - y_2) + (y_3 - y_2) + (y_4 - y_2))$. Similar relations can be found for all the other nodes. If the different loops $[K_i \ G_i]$ are designed to make the tracking error $y_{ref,i} - y_i$ as small as possible, it can be proven that such an interconnection allows good tracking of ref_{ext} at all nodes. A normalized adjacency matrix can be defined for any information flow using the following rules. Row i of \mathcal{A} is zero if no output is sent to node i . If y_j is sent to node i , entry (i, j) of \mathcal{A} will be nonzero. Finally, all nonzero entries in a row are equal and sum up to one.

2.2 Network identification procedure

In this paper, our objective is to redesign the local controllers K_i ($i = 1 \dots N_{mod}$) in order to improve the performance of the network by identifying sufficiently accurate models of each interconnected system \mathcal{S}_i ($i = 1 \dots N_{mod}$). Let us first consider the identification procedure. An identification experiment is performed by adding an excitation signal $r_i(t)$ ($t = 1 \dots N$) having spectrum Φ_{r_i} at the output of each decentralized controller (see Figure 2). Since the external reference signal ref_{ext} is, as we will see, not required for identification purpose, ref_{ext} is put to zero during the identification experiment. This transforms Equations (2) and (3) into:

$$u_i(t) = r_i(t) + K_i(z)(y_{ref,i} - y_i(t)) \quad (5)$$

$$\bar{y}_{ref}(t) = \mathcal{A} \bar{y}(t) \quad (6)$$

This experiment allows to collect the input-output data sets $Z_i^N = \{u_i(t), y_i(t) \mid t = 1 \dots N\}$ ($i = 1 \dots N_{mod}$) corresponding to each of the N_{mod} modules.

Instead of using one single global MIMO identification criterion to identify in one step all systems \mathcal{S}_i ($i = 1 \dots N_{mod}$) with all data sets Z_i^N ($i = 1 \dots N_{mod}$), we will here use a simpler, but equivalent identification procedure. Indeed, we will show that a consistent estimate $\hat{\theta}_i$ of the true parameter vector $\theta_{i,0}$ of system \mathcal{S}_i can be identified using only the data set Z_i^N . For this purpose, we use the classical SISO prediction-error identification criterion [21] with a full-order model structure $\mathcal{M}_i = \{G_i(z, \theta_i), H_i(z, \theta_i)\}$ (i.e. a model structure such that $\mathcal{S}_i \in \mathcal{M}_i$):

$$\hat{\theta}_i = \arg \min_{\theta_i} \frac{1}{N} \sum_{t=1}^N \epsilon_i^2(t, \theta_i) \quad (7)$$

$$\epsilon_i(t, \theta_i) = H_i^{-1}(z, \theta_i) (y_i(t) - G_i(z, \theta_i)u_i(t)) \quad (8)$$

This approach is simple since it only involves N_{mod} individual SISO prediction-error identification criteria. This can be an important advantage when N_{mod} is large. Moreover, we will show in Subsection 2.3 that this simple approach is in fact equivalent to the global MIMO prediction error identification criterion.

Before going further, one should verify that $\hat{\theta}_i$ obtained in this way is indeed a consistent estimate of $\theta_{i,0}$ or, in other words that $\theta_{i,0}$ is the unique solution of the asymptotic identification criterion $\theta_i^* = \arg \min_{\theta} \bar{E} \epsilon_i^2(t, \theta_i)$ where $\bar{E} \epsilon_i^2(t, \theta_i)$ is defined as $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E \epsilon_i^2(t, \theta_i)$. For this purpose, we will need to make a number of classical structural assumptions; assumptions that have also to be made when we consider classical (direct) closed-loop identification (see e.g. [12]).

Assumptions. (A.1) For all i , $\theta_i = \theta_{i,0}$ is the only parameter vector for which the models $G_i(z, \theta_i)$ and $H_i(z, \theta_i)$ correspond to the true system \mathcal{S}_i . (A.2) For all i , the product $K_i(z)G_i(z, \theta_{i,0})$ contains (at least) one delay. (A.3) For all i , the excitation signal r_i is statistically independent of $\bar{e} = (e_1, e_2, \dots, e_{N_{mod}})^T$.

Conditions for consistent identification of a module in an arbitrary network are given in [6]. Because arbitrary networks are considered in [6], more attention is given to the selection of the right predictor inputs for the identification criterion. Other important aspects as the informativity of the input-output data are dealt by the classical condition of the positivity of the spectral density matrix. However, as mentioned in [13], there is very little information in [6] on how to obtain informative data by an appropriate design of the experiment (i.e. the application of the excitation signals r_i)². In particular, this condition is very far away from the very detailed conditions deduced in e.g. [12] to guarantee the consistency of (7) when the data Z_i^N are collected in a simple closed loop i.e. when the system \mathcal{S}_i in (1) is operated with $u_i(t) = r_i(t) - K_i(z)y_i(t)$ ($y_{ref,i} = 0$). Indeed, under the assumptions (A.1), (A.2) and (A.3), [12] shows that (7) is a consistent estimate of $\theta_{i,0}$ if and only if the number of frequencies, at which the spectrum $\Phi_{r_i}(\omega)$ is nonzero, is larger than a given threshold. This threshold uniquely depends on the order of the controller $K_i(z)$ and on the respective parametrization and the orders of $G_i(z, \theta_i)$ and $H_i(z, \theta_i)$. If the order of $K_i(z)$ is large, this threshold can be negative and consistency is then also guaranteed when $r_i = 0$. Moreover, choosing r_i as a filtered white noise is always a sufficient choice to guarantee consistency when considering a single closed-loop system.

In Theorem 1, we will extend these conditions to the case where the closed-loop systems are interconnected in the manner presented above. Before presenting this result, let us observe that, due to the interconnection (6), the input

² The same observation can be made for the conditions derived in the paper [27] which considers the identification of all the modules in an arbitrary network.

signals u_i ($i = 1 \dots N_{mod}$) during the identification experiment can be expressed as:

$$u_i(t) = \sum_{j=1}^{N_{mod}} (R_{ij}(z) r_j(t) + S_{ij}(z) e_j(t)) \quad (9)$$

for given transfer functions R_{ij} and S_{ij} ($i, j = 1 \dots N_{mod}$) that can easily be computed using (1), (5), (6) and LFT algebra [8]. The transfer functions R_{ii} and S_{ii} are always nonzero. Excluding pathological cases, the transfer functions R_{ij} and S_{ij} for $i \neq j$ are both nonzero if and only if there exists a path from node j to node i . In the example of Figure 1, we can e.g. say that R_{31} and S_{31} will be nonzero. Indeed, there exists a path from Node 1 to Node 3 since y_1 is sent to Node 2 and y_2 is in turn sent to Node 3. Consequently, u_3 will be, via $y_{ref,3}$, a function of y_1 which is in turn a function of r_1 and e_1 . As another example, R_{56} and S_{56} will be zero because there is no path from Node 6 to Node 5. Indeed, y_6 is not sent to any node and can therefore not influence u_5 .

Theorem 1 Consider the data set $Z_i^N = \{u_i(t), y_i(t) \mid t = 1 \dots N\}$ collected in one arbitrary node i of a network made up N_{mod} modules (1). The modules in this network are operated as in (5) and the interconnection of the network is defined via an adjacency matrix (see (6)). Consider furthermore (4) with a matrix $\Lambda = \Lambda^T$ that is strictly positive definite and consider also the assumptions (A.1), (A.2) and (A.3). Then, the estimate $\hat{\theta}_i$ obtained via (7) is a consistent estimate of $\theta_{i,0}$ if and only if one of the following two conditions are satisfied

- (i) there exists at least one path from one node $j \neq i$ to the considered node i (i.e. $R_{ij}(z) \neq 0$ and $S_{ij}(z) \neq 0$)
- (ii) the excitation signal r_i satisfies the signal richness condition of [12] that guarantees consistency of (7) in a simple closed loop (i.e. when the system \mathcal{S}_i (see (1)) is operated with $u_i(t) = r_i(t) - K_i(z)y_i(t)$). This condition uniquely depends on the order of the controller $K_i(z)$ and on the respective parametrization and the orders of $G_i(z, \theta_i)$ and $H_i(z, \theta_i)$.

Proof. Without loss of generality, let us suppose that $i = 1$. Moreover, for conciseness, let us drop the argument z in the transfer functions. Using (1) and (8), the prediction error $\epsilon_1(t, \theta_1)$ is given by:

$$\epsilon_1(t, \theta_1) = e_1(t) + \frac{\Delta H_1(\theta_1)}{H_1(\theta_1)} e_1(t) + \frac{\Delta G_1(\theta_1)}{H_1(\theta_1)} u_1(t) \quad (10)$$

with $\Delta H_1(\theta_1) = H_1(\theta_{1,0}) - H_1(\theta_1)$ and $\Delta G_1(\theta_1) = G_1(\theta_{1,0}) - G_1(\theta_1)$. Inserting (9) into (10) and using the notation $\bar{e} = (e_1, e_2, \dots, e_{N_{mod}})^T$ and $\bar{r} = (r_1, r_2, \dots, r_{N_{mod}})^T$, $\epsilon_1(t, \theta_1)$ can be rewritten as:

$$\epsilon_1(t, \theta_1) = e_1(t) + \mathcal{L}_1(\theta_1) \bar{e}(t) + \mathcal{R}_1(\theta_1) \bar{r}(t) \quad (11)$$

$$\mathcal{L}_1(z, \theta_1) = \left(\mathcal{V}_1(\theta_1), \frac{\Delta G_1(\theta_1)}{H_1(\theta_1)} S_{12}, \dots, \frac{\Delta G_1(\theta_1)}{H_1(\theta_1)} S_{1N_{mod}} \right)$$

$$\mathcal{R}_1(z, \theta_1) = \frac{\Delta G_1(\theta_1)}{H_1(\theta_1)} (R_{11}, R_{12}, \dots, R_{1N_{mod}})$$

with $\mathcal{V}_1(\theta_1) \triangleq \frac{\Delta H_1(\theta_1)}{H_1(\theta_1)} + \frac{\Delta G_1(\theta_1)}{H_1(\theta_1)} S_{11}$. (A.2) implies that, when nonzero, the transfer functions $\frac{\Delta G_1(\theta_1)}{H_1(\theta_1)} S_{1j} \forall j$ all contain at least one delay. Moreover, when nonzero, $\Delta H_1(\theta_1)$ also contains one delay since H_1 is monic. Combining these two facts and (A.3) and recalling that $\Phi_{\bar{e}}(\omega) = \Lambda$, the power $\bar{E}\epsilon_1^2(t, \theta_1)$ of $\epsilon_1(t, \theta_1)$ is equal to:

$$\begin{aligned} \Lambda_{1,1} + \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{L}_1(e^{j\omega}, \theta_1) \Lambda \mathcal{L}_1^*(e^{j\omega}, \theta_1) d\omega + \dots \\ \dots + \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{R}_1(e^{j\omega}, \theta_1) \Phi_{\bar{r}}(\omega) \mathcal{R}_1^*(e^{j\omega}, \theta_1) d\omega \end{aligned}$$

with $\Lambda_{1,1}$ the (1,1) entry of Λ (i.e. the variance of e_1) and $\Phi_{\bar{r}}(\omega) \geq 0$ the power spectrum of \bar{r} .

Note that, for all θ_1 , $\bar{E}\epsilon_1^2(t, \theta_1) \geq \Lambda_{1,1}$. Consequently, since we have that $\epsilon_1(t, \theta_{1,0}) = e_1(t)$ and thus that $\bar{E}\epsilon_1^2(t, \theta_{1,0}) = \Lambda_{1,1}$, all minimizers θ_1^* of $\bar{E}\epsilon_1^2(t, \theta_1)$ must be such that $\bar{E}\epsilon_1^2(t, \theta_1^*) = \Lambda_{1,1}$.

To prove this theorem, we will first show that Condition (i) is sufficient to guarantee that $\theta_1^* = \theta_{1,0}$ is the unique minimizer of $\bar{E}\epsilon_1^2(t, \theta_1)$ i.e. that $\theta_1^* = \theta_{1,0}$ is the only parameter vector such that $\bar{E}\epsilon_1^2(t, \theta_1^*) = \Lambda_{1,1}$. For this purpose, let us start by the following observation. Due to our assumption that $\Lambda > 0$, any parameter vector θ_1^* such that $\bar{E}\epsilon_1^2(t, \theta_1^*) = \Lambda_{1,1}$ must satisfy:

$$\mathcal{L}_1(\theta_1^*) \equiv 0 \quad (12)$$

If Condition (i) holds (i.e. if there is a path from a node $j \neq 1$ to node 1), we have that $S_{1j} \neq 0$. To satisfy (12), it must in particular hold that $\frac{\Delta G_1(\theta_1^*)}{H_1(\theta_1^*)} S_{1j} \equiv 0$. Since $S_{1j} \neq 0$, this yields $\Delta G_1(\theta_1^*) = 0$. Using $\Delta G_1(\theta_1^*) = 0$ and the fact that the first entry $\mathcal{V}_1(\theta_1^*)$ of $\mathcal{L}_1(\theta_1^*)$ must also be equal to zero to satisfy (12), we obtain $\Delta H_1(\theta_1^*) = 0$. Since, by virtue of (A.1), $\theta_{1,0}$ is the unique parameter vector making both ΔG_1 and ΔH_1 equal to 0, we have thus proven the consistency under Condition (i). Note that this result is irrespective of $\Phi_{\bar{r}}$ i.e. it holds for any choice of $\Phi_{\bar{r}}$.

In order to conclude this proof, it remains to be proven that, if Condition (i) does not hold, Condition (ii) is a necessary and sufficient condition for consistency. This part is straightforward. Indeed, if Condition (i) does not hold, $y_{ref,1} = 0$ and the data set $Z_1^N = \{u_1(t), y_1(t) \mid t = 1 \dots N\}$ is generated as in an isolated closed-loop system. The result is thus proven since the paper [12] gives necessary and sufficient conditions on the richness of r_1 to guarantee consistency in an isolated closed-loop system. ■

Theorem 1 shows that the network configuration considered in this paper is in fact beneficial for identification. Indeed, consistency of (7) is not only guaranteed in all situations where consistency is guaranteed in the simple closed-loop case (i.e. when $y_{ref,i} = 0$), but also in many other cases (via Condition (i)). Indeed, Condition (i) shows that, due to the interconnection, disturbances v_j in other nodes connected via a path to node i are sufficient to lead to consistency of (7) and this even in the extreme case where all excitations signals r_j ($j = 1 \dots N_{mod}$) are set to zero. For example, in the network of Figure 1, Condition (i) applies to Nodes 2, 3, 4, 5 and 6. Consequently, the richness condition of [12] has only to be respected to identify a consistent estimate $\hat{\theta}_1$ of the module \mathcal{S}_1 which is the only isolated module in this network.

Remark 1. Note that the result of Theorem 1 only requires that $\Lambda = \Lambda^T > 0$. Consequently, it not only applies to signals e_i that are mutually independent, but also to signals e_i that are spatially correlated. This is an interesting observation since the independence of the disturbances having a path to Node i is an assumption in [6].

2.3 SISO criterion vs. MIMO criterion

Let us define $\hat{\theta} = (\hat{\theta}_1^T, \hat{\theta}_2^T, \dots, \hat{\theta}_{N_{mod}}^T)^T$ using the different estimates $\hat{\theta}_i$ ($i = 1 \dots N_{mod}$) obtained using the individual SISO criteria (7) for all modules in the network. In this subsection, we show that this estimate $\hat{\theta}$ of $\theta_0 = (\theta_{1,0}^T, \dots, \theta_{N_{mod},0}^T)^T$ is equal to the estimate $\hat{\theta}^{mimo}$ obtained with the prediction error criterion using all data sets Z_i^N ($i = 1 \dots N_{mod}$). Using (1), the optimally weighted MIMO prediction error identification criterion [21] is $\hat{\theta}^{mimo} = \arg \min_{\theta} V(\theta)$ with

$$V(\theta) = \frac{1}{N} \sum_{t=1}^N \bar{\epsilon}^T(t, \theta) \Lambda^{-1} \bar{\epsilon}(t, \theta) \quad (13)$$

$$\bar{\epsilon}(t, \theta) = H^{-1}(z, \theta) (\bar{y}(t) - G(z, \theta) \bar{u}(t)) \quad (14)$$

with $\theta = (\theta_1^T, \theta_2^T, \dots, \theta_{N_{mod}}^T)^T$, $G(z, \theta)$ a diagonal transfer matrix equal to $diag(G_1(\theta_1), G_2(\theta_2), \dots, G_{N_{mod}}(\theta_{N_{mod}}))$ and $H(z, \theta)$ defined similarly as $G(z, \theta)$. The data $\bar{y} \triangleq (y_1, \dots, y_{N_{mod}})^T$ and $\bar{u} \triangleq (u_1, \dots, u_{N_{mod}})^T$ are the data collected in Z_i^N ($i = 1 \dots N_{mod}$).

Let us observe that $\bar{\epsilon}(t, \theta) = (\epsilon_1(t, \theta_1), \dots, \epsilon_{N_{mod}}(t, \theta_{N_{mod}}))^T$ with $\epsilon_i(t, \theta_i)$ as in (8). Using this expression for $\bar{\epsilon}(t, \theta)$ and the assumption that Λ is diagonal, we can rewrite $V(\theta)$ as:

$$V(\theta) = \sum_{i=1}^{N_{mod}} \frac{1}{\Lambda_{i,i}} V_i(\theta_i) \quad (15)$$

with $\Lambda_{i,i}$ the (i, i) -entry of Λ (i.e. the variance of e_i) and $V_i(\theta_i) = \frac{1}{N} \sum_{t=1}^N \epsilon_i^2(t, \theta_i)$ the cost function used in the SISO criterion (7). This last expression shows that minimizing the individual cost functions $V_i(\theta_i)$ (as done in (7)) is equivalent to minimizing $V(\theta)$ and thus that $\hat{\theta} = \hat{\theta}^{mimo}$. Consequently, when Λ is diagonal, there is no disadvantage whatsoever in using the individual SISO criteria (7) instead of the MIMO criterion (13).

If the global experiment is designed such that consistency is guaranteed for each Module i (see Theorem 1), the estimate $\hat{\theta} = \hat{\theta}^{mimo}$ of θ_0 has the property that $\sqrt{N}(\hat{\theta} - \theta_0)$ is asymptotically normally distributed around zero [21]. The covariance matrix of $\hat{\theta}$ is moreover given by $P_{\hat{\theta}} = \frac{1}{N} (\bar{E} \Psi(t, \theta_0) \Lambda^{-1} \Psi^T(t, \theta_0))^{-1}$ where $\Psi(t, \theta) = \frac{-\partial \bar{\epsilon}^T(t, \theta)}{\partial \theta}$ [21]. Let us observe that $\Psi(t, \theta)$ is a block-diagonal matrix:

$$\Psi(t, \theta) = bdiag(\psi_1(t, \theta_1), \psi_2(t, \theta_2), \dots, \psi_{N_{mod}}(t, \theta_{N_{mod}}))$$

with $\psi_i(t, \theta_i) = \frac{-\partial \epsilon_i(t, \theta_i)}{\partial \theta_i}$. Consequently, P_θ has the following block-diagonal structure:

$$P_\theta = \text{bdiag}(P_{\theta_1}, P_{\theta_2}, \dots, P_{\theta_{N_{mod}}}) \quad (16)$$

$$P_{\theta_i} = \frac{\Lambda_{i,i}}{N} (\bar{E}\psi_i(t, \theta_{i,0})\psi_i^T(t, \theta_{i,0}))^{-1} \quad i = 1 \dots N_{mod}$$

The covariance matrices P_{θ_i} in P_θ are the covariance matrices of the individual estimates $\hat{\theta}_i$ for each i . Note that P_{θ_i} can be estimated from the data Z_i^N and $\hat{\theta}_i$ [21]. However, for further use, we also derive an expression of P_{θ_i} as a function of the experimental conditions. For this purpose, we recall (see e.g. [4]) that $\psi_i(t, \theta_{i,0}) = F_i(z, \theta_{i,0})u_i(t) + L_i(z, \theta_{i,0})e_i(t)$ with $F_i(\theta_i) = H_i^{-1}(\theta_i) \frac{\partial G(\theta_i)}{\partial \theta_i}$ and $L_i(\theta_i) = H_i^{-1}(\theta_i) \frac{\partial H(\theta_i)}{\partial \theta_i}$. Using (9) and assuming that the excitation signals r_j ($j = 1 \dots N_{mod}$) are all mutually independent, we obtain:

$$P_{\theta_i}^{-1} = \frac{N}{2\pi\Lambda_{i,i}} \int_{-\pi}^{\pi} \mathcal{Z}_i(e^{j\omega}) \Lambda \mathcal{Z}_i^*(e^{j\omega}) d\omega + \dots \quad (17)$$

$$\dots \frac{N}{2\pi\Lambda_{i,i}} \int_{-\pi}^{\pi} F_i(e^{j\omega}) F_i^*(e^{j\omega}) \left(\sum_{j=1}^{N_{mod}} |R_{ij}(e^{j\omega})|^2 \Phi_{r_j}(\omega) \right) d\omega$$

with $\mathcal{Z}_i(z)$ a matrix of transfer functions of dimension $n_{\theta_i} \times N_{mod}$ whose i^{th} column is $L_i + F_i S_{ii}$ and whose j^{th} -column ($j \neq i$) is equal to $F_i S_{ij}$. Note that this expression depends not only on $\theta_{i,0}$, but also, via the nonzero transfer functions S_{ij} , R_{ij} , on the true parameter vector $\theta_{j,0}$ of the systems \mathcal{S}_j ($j \neq i$) having a path to node i .

It is also important to note that not only the excitation signal r_i but also all r_j in nodes having a path to i contributes to the accuracy $P_{\theta_i}^{-1}$ of $\hat{\theta}_i$. In the network of Figure 1, the accuracy $P_{\theta_6}^{-1}$ of the model of \mathcal{S}_6 will thus be influenced by the excitations signals r_j ($j = 1 \dots 6$) in all nodes. Moreover, due to the structure of (17), we could also theoretically obtain any accuracy $P_{\theta_6}^{-1}$ for that model by e.g. only exciting at Node 1 (i.e. $r_1 \neq 0$ and $r_j = 0$ for $j = 2 \dots 6$). It is nevertheless to be noted that a larger excitation power (or a longer experiment) will then be typically necessary to guarantee this accuracy because of the attenuation of the network.

2.4 Uncertainty bounding

Suppose we have made an experiment leading to informative data Z_i^N for all modules \mathcal{S}_i ($i = 1 \dots N_{mod}$). We can thus obtain the estimate $\hat{\theta}$ of θ_0 using the individual identification criteria (7). Given the properties of $\hat{\theta}$ given above, the ellipsoid $U = \{\theta \mid (\theta - \hat{\theta})^T P_\theta^{-1} (\theta - \hat{\theta}) < \chi\}$ with $Pr(\chi^2(n_\theta) < \chi) = \beta$ (n_θ is the dimension of θ) will for sufficiently large sample size N be a $\beta\%$ -confidence region for the unknown parameter vector θ_0 (say $\beta = 95\%$). For the robustness analysis via the hierarchical approach that will be presented in the next section, we will also need the projections U_i of U into the parameter space θ_i of each of the modules $i = 1 \dots N_{mod}$. Using (16) and the fact that $\theta = (\theta_1^T, \theta_2^T, \dots, \theta_{N_{mod}}^T)^T$, the projection $U_i = \{\theta_i \mid \theta \in U\}$ ($i = 1 \dots N_{mod}$) is an ellipsoid given by (see e.g. [2]):

$$U_i = \{\theta_i \mid (\theta_i - \hat{\theta}_i)^T P_{\theta_i}^{-1} (\theta_i - \hat{\theta}_i) < \chi\} \quad (18)$$

3 Controller validation

In the previous section, we have seen that models $G_i(z, \hat{\theta}_i)$ of the systems \mathcal{S}_i ($i = 1 \dots N_{mod}$) can be obtained using a global identification experiment on an interconnected network. The identified models can now be used to design improved decentralized controllers \hat{K}_i for each module (see e.g. [24]). Note that, if the different systems \mathcal{S}_i are homogeneous i.e. they are identical up to industrial dispersion, one could design a common controller $\hat{K}_i = \hat{K} \forall i$ using an average of the identified models (see e.g. [19]).

In any case, the decentralized controllers \hat{K}_i are designed to guarantee both a nominal local performance (performance of the loop $[\hat{K}_i G_i(z, \hat{\theta}_i)]$) and a nominal global performance (performance of the network). In [24,19], the H_∞ framework is used to measure both the local and global performance. As usual in classical H_∞ control design, a sufficient level of local performance is ensured by imposing, for all i , a frequency-dependent threshold on the modulus of the frequency responses of transfer functions such as $1/(1 + \hat{K}_i G_i(z, \hat{\theta}_i))$ and $\hat{K}_i/(1 + \hat{K}_i G_i(z, \hat{\theta}_i))$. This indeed allows e.g. to guarantee a certain tracking ability for each local loop (since the first transfer function is the one between $y_{ref,i}$ and $y_i - y_{ref,i}$) and to limit the control efforts (since the second transfer function is the one between $y_{ref,i}$ and u_i). Since the loops $[\hat{K}_i G_i(z, \hat{\theta}_i)]$ are not isolated, but interconnected as in (3), the control design method in [24,19] also imposes specific constraints on the global performance by imposing a frequency-dependent threshold on the modulus of the frequency responses of transfer functions P describing the behaviour of the network

as a whole. Examples of such global transfer functions P are the transfer functions between the external reference ref_{ext} and the tracking error $y_i - ref_{ext}$ at each node of the network. Other examples are the transfer functions between ref_{ext} and the input signal u_i at each node of the network. Finally, we can also consider the transfer functions between a disturbance v_i in one node and an output signal y_j in the same or another node. It is clear that these transfer functions reflect the performance of the whole network with respect to tracking, control efforts and disturbance rejection, respectively.

For the design of \hat{K}_i , the thresholds (weightings) corresponding to each transfer function described in the previous paragraph must be chosen in such a way that they improve the performance of the original network. The performance of the original network can be evaluated by computing the frequency responses of these transfer functions in the network made up of the interconnection of the loops $[K_i G_i(z, \hat{\theta}_i)]$.

Since the decentralized controllers \hat{K}_i are designed based on the models $G_i(z, \hat{\theta}_i)$ of the true systems S_i ($i = 1 \dots N_{mod}$), it is important to verify whether these decentralized controllers will also lead to a satisfactory level of performance both at the local level and at the global level when they will be applied to the true systems S_i ($i = 1 \dots N_{mod}$). Since both the local and the global performance may be described by different transfer functions, the verification that will be presented below must be done for each of these transfer functions. For the transfer functions describing the local performance, the robustness analysis results of [3] can be used. In the sequel, we will thus restrict attention to the global performance and show how to perform this verification for one arbitrary transfer function P describing the global performance. The input of this transfer function will be denoted by w (e.g. $w = ref_{ext}$) and its output will be denoted by s (e.g. $s = y_i - ref_{ext}$ for some i). Let us thus denote by $P(z, \hat{\theta})$ the value of this transfer function for the network made up of the interconnection of the loops $[\hat{K}_i G_i(z, \hat{\theta}_i)]$. Similarly, let us denote by $P(z, \theta_0)$ its value for the network made up of the loops $[\hat{K}_i G_i(z, \theta_{i,0})]$. In order to verify the robustness of the designed controllers \hat{K}_i with respect to this global transfer function P , we have thus to verify whether, for all ω ,

$$|P(e^{j\omega}, \theta_0)| < W(\omega) \quad (19)$$

where $W(\omega)$ is the threshold that defines the desired global performance (wrt. P) in the redesigned network. Since the unknown θ_0 lies in U (modulo the confidence level β), (19) will be deemed verified if $\sup_{\theta \in U} |P(e^{j\omega}, \theta)| < W(\omega)$ at each ω . Note that a necessary condition for the latter to hold is obviously that the designed transfer function $P(z, \hat{\theta})$ satisfies $|P(e^{j\omega}, \hat{\theta})| < W_{nom}(\omega)$ with $W_{nom}(\omega) < W(\omega)$ for all ω . For a successful redesign of the network, the controller \hat{K}_i must thus be designed with a nominal performance that is (at least slightly) better than the desired performance.

Computing $\sup_{\theta \in U} |P(e^{j\omega}, \theta)|$ exactly is not possible. However, we can deduce upper bounds for this quantity. One possible approach to do this is to use μ -analysis based on the parametric uncertainty set U [28,1]. However, since networks can in practice be made up of a large number N_{mod} of modules yielding a parametric uncertainty set of large dimension, this direct approach could reveal impractical from a computational point-of-view [23,7] and the computed upper bound could possibly turn out to be relatively conservative. Consequently, we will here consider the two-step hierarchical approach proposed in [23,7] and show how this two-step approach can be applied for the type of parametric uncertainties delivered by network identification.

As a first step, we will give a convenient expression of the global transfer function $P(z, \theta_0)$. For this purpose, note that $P(z, \theta_0)$ pertains to a network characterized by the interconnection equation (3) and the following equation describing the local loop $[\hat{K}_i G_i(z, \theta_{i,0})]$:

$$y_i(t) = v_i(t) + T_i(z, \theta_{i,0})(y_{ref,i}(t) - v_i(t)) \quad (20)$$

with $T_i(z, \theta_{i,0}) = \frac{\hat{K}_i G_i(z, \theta_{i,0})}{1 + \hat{K}_i G_i(z, \theta_{i,0})}$. Consequently, the usually used global performance transfer functions $P(z, \theta_0)$ can be written as an LFT of $T(z, \theta_0) = \text{diag}(T_1(z, \theta_{1,0}), T_2(z, \theta_{1,0}), \dots, T_{N_{mod}}(z, \theta_{N_{mod},0}))$ i.e. we can determine vectors of signals p and q such that the transfer function $P(z, \theta_0)$ between w and s can be written as:

$$p = T(z, \theta_0)q \quad \text{and} \quad \begin{pmatrix} q \\ s \end{pmatrix} = \mathcal{I}(z) \begin{pmatrix} p \\ w \end{pmatrix} \quad (21)$$

for a given matrix of transfer functions $\mathcal{I}(z)$ that does not depend on θ_0 (i.e. that does not depend on $G_i(z, \theta_{i,0})$ ($i = 1 \dots N_{mod}$)). For this LFT representation, we will use the shorthand notation: $P(z, \theta_0) = \mathcal{F}(\mathcal{I}(z), T(z, \theta_0))$.

As an example, in the network of Figure 1, let us consider the transfer function between $w = ref_{ext}$ and $s = y_6 - ref_{ext}$ and let us consequently pose $v_i = 0$ ($i = 1 \dots N_{mod}$). This transfer function can be described as in (21)

with $q = \bar{y}_{ref}$, $p = \bar{y}$ and the following constant matrix \mathcal{I} :

$$\mathcal{I} = \begin{pmatrix} & \mathbf{A} & \mathbf{B} \\ (0, 0, 0, 0, 0, 1) & & -1 \end{pmatrix}$$

The matrix $T(z, \theta_0)$ depends on the unknown true parameter vector θ_0 . Since $\theta_0 \in U$ (modulo the confidence level β), $T(z, \theta_0)$ obviously lies in the parametric uncertainty set $\mathcal{T}^s = \{T(z, \theta) \mid \theta \in U\}$. Note that $\sup_{\theta \in U} |P(e^{j\omega}, \theta)| = \sup_{T \in \mathcal{T}^s} |\mathcal{F}(\mathcal{I}(e^{j\omega}), T(e^{j\omega}))|$. Computing an upper bound of the latter expression remains as complicated as with the former since \mathcal{T}^s is still a parametric uncertainty set of large dimension (if N_{mod} is large). However, this LFT representation of $P(z, \theta_0)$ enables the use of the hierarchical approach. The idea of the hierarchical approach is to embed \mathcal{T}^s into an uncertainty set \mathcal{T} having a structure for which the robustness analysis of the global performance is tractable even if N_{mod} is large. For this purpose, we can choose the following structure for \mathcal{T} :

$$\mathcal{T} = \{T(z) \mid T(z) = (I_{N_{mod}} + \Delta(z)) T(z, \hat{\theta}) \text{ with} \\ \dots \Delta(e^{j\omega}) \in \mathbf{\Delta}(\omega) \forall \omega\} \quad (22)$$

where $\Delta(z) = \text{diag}(\Delta_1(z), \Delta_2(z), \dots, \Delta_{N_{mod}}(z))$ is the (stable) uncertainty made up of N_{mod} scalar transfer functions $\Delta_i(z)$ ($i = 1 \dots N_{mod}$). The set $\mathbf{\Delta}(\omega)$ will have a similar diagonal structure: $\mathbf{\Delta}(\omega) = \text{diag}(\mathbf{\Delta}_1(\omega), \mathbf{\Delta}_2(\omega), \dots, \mathbf{\Delta}_{N_{mod}}(\omega))$. The elements $\mathbf{\Delta}_i(\omega)$ constrain the frequency response $\Delta_i(e^{j\omega})$ of $\Delta_i(z)$ as follows:

$$\mathbf{\Delta}_i(\omega) = \{\Delta_i(e^{j\omega}) \mid |\Delta_i(e^{j\omega}) - c_i(\omega)| < \rho_i(\omega)\} \quad (23)$$

i.e. $\mathbf{\Delta}_i(\omega)$ is a disk (in the complex plane) of radius $\rho_i(\omega)$ and of (complex) center $c_i(\omega)$.

Since, as mentioned above, \mathcal{T} will be determined in such a way that $\mathcal{T}^s \subseteq \mathcal{T}$, $T(z, \theta_0)$ will also lie in \mathcal{T} and we will thus be able to verify (19) by verifying at each ω that $\mathcal{P}_{wc}(\omega, \mathcal{T}) < W(\omega)$ with

$$\mathcal{P}_{wc}(\omega, \mathcal{T}) \triangleq \sup_{T(z) \in \mathcal{T}} |\mathcal{F}(\mathcal{I}(e^{j\omega}), T(e^{j\omega}))| \quad (24)$$

Since $T(z) = (I_{N_{mod}} + \Delta(z)) T(z, \hat{\theta})$ is an LFT in $\Delta(z)$, $\mathcal{F}(\mathcal{I}(z), T(z))$ can be rewritten in an LFT in $\Delta(z)$ i.e. $\mathcal{F}(\mathcal{I}(z), T(z)) = \mathcal{F}(M(z), \Delta(z))$ with $M(z)$ a function of $\mathcal{I}(z)$ and $T(z, \hat{\theta})$. Consequently, (24) is also given by:

$$\mathcal{P}_{wc}(\omega, \mathcal{T}) = \sup_{\Delta(e^{j\omega}) \in \mathbf{\Delta}(\omega)} |\mathcal{F}(M(e^{j\omega}), \Delta(e^{j\omega}))| \quad (25)$$

Before presenting how we can evaluate (25), we will first present how the uncertainty set \mathcal{T} can be determined in practice. First note that we can decompose the uncertainty set \mathcal{T} into N_{mod} SISO (unstructured) uncertainty sets \mathcal{T}_i defined as follows:

$$\mathcal{T}_i = \{T_i(z) \mid T_i(z) = (1 + \Delta_i(z)) T_i(z, \hat{\theta}_i) \text{ with} \\ \dots \Delta_i(e^{j\omega}) \in \mathbf{\Delta}_i(\omega) \forall \omega\} \quad (26)$$

with $\mathbf{\Delta}_i(\omega)$ as defined in (23). Ensuring $\mathcal{T}^s \subseteq \mathcal{T}$ can thus be obtained by determining the sets \mathcal{T}_i in such a way that, for all $i = 1 \dots N_{mod}$, $\mathcal{T}_i^s \subseteq \mathcal{T}_i$ with \mathcal{T}_i^s defined as follows:

$$\mathcal{T}_i^s = \{T_i(z, \theta_i) \mid T_i(z, \theta_i) = \frac{\hat{K}_i(z)G_i(z, \theta_i)}{1 + \hat{K}_i(z)G_i(z, \theta_i)} \text{ with } \theta_i \in U_i\}. \quad (27)$$

with U_i as in (18). In order to achieve this in an optimal way, the frequency functions $\rho_i(\omega)$ and $c_i(\omega)$ defining, via (23), the size of \mathcal{T}_i will be determined in such a way that \mathcal{T}_i is the *smallest* set for which it still holds that $\mathcal{T}_i^s \subseteq \mathcal{T}_i$. By doing so, we indeed reduce as much as possible the conservatism linked to the embedding of the uncertainty set \mathcal{T}_i^s (that follows from the identification experiment) into the unstructured uncertainty set \mathcal{T}_i . Consequently, for a given i and for a given ω , $\rho_i(\omega)$ and $c_i(\omega)$ have to be chosen as the solution of the following optimization problem:

$$\begin{aligned} & \min \rho_i(\omega) \\ & \text{s.t. } |\tilde{T}_i(e^{j\omega}, \theta_i) - c_i(\omega)| < \rho_i(\omega) \quad \forall \theta_i \in U_i \\ & \text{with } \tilde{T}_i(e^{j\omega}, \theta_i) = \frac{T_i(e^{j\omega}, \theta_i) - T_i(e^{j\omega}, \hat{\theta}_i)}{T_i(e^{j\omega}, \hat{\theta}_i)} \end{aligned} \quad (28)$$

with $T_i(z, \theta_i)$ as defined in (27). As shown in the following theorem, the solution of the optimization problem (28) can be efficiently determined using LMI optimization [5]. Before presenting this result, let us first give an expression of $\tilde{T}_i(e^{j\omega}, \theta_i)$ as a function of θ_i using the following notation for $G_i(e^{j\omega}, \theta_i) = \frac{Z_{1,i}(e^{j\omega})\theta_i}{1+Z_{2,i}(e^{j\omega})\theta_i}$. In the last expression, $Z_{1,i}(z)$ and $Z_{2,i}(z)$ are row vectors containing only delays or zeros (see [3]). This yields

$$\tilde{T}_i(e^{j\omega}, \theta_i) = \frac{-1 + Z_{N,i}(e^{j\omega})\theta_i}{1 + Z_{D,i}(e^{j\omega})\theta_i} \quad (29)$$

with $Z_{D,i} = Z_{2,i} + \hat{K}_i Z_{1,i}$ and $Z_{N,i} = \frac{\hat{K}_i Z_{1,i}}{T_i(e^{j\omega}, \theta_i)} - Z_{D,i}$.

Theorem 2 Consider the notation $\tilde{T}_i(e^{j\omega}, \theta_i) = \frac{-1+Z_{N,i}(e^{j\omega})\theta_i}{1+Z_{D,i}(e^{j\omega})\theta_i}$ given in (29). The optimization problem (28) at a given ω and at a given i is equivalent to the following LMI optimization problem having as decision variable a positive real scalar $\alpha_i(\omega)$, a complex scalar $c_i(\omega)$, a positive real scalar $\xi_i(\omega)$ and a skew-symmetric matrix $\mathcal{X}_i(\omega) \in \mathbf{R}^{(n_{\theta_i}+1) \times (n_{\theta_i}+1)}$.

min $\alpha_i(\omega)$ subject to

$$\left(\begin{array}{c|c} -\alpha_i(\omega) & \lambda_i(\omega) \\ \hline \lambda_i^*(\omega) & -A_i(\omega) - \xi_i(\omega)B_i + j\mathcal{X}_i(\omega) \end{array} \right) < 0 \quad (30)$$

with $\lambda_i(\omega) = \left(Z_{N,i} - Z_{D,i}c_i \mid -1 - c_i \right)$ and

$$A_i(\omega) = \begin{pmatrix} Z_{D,i}^* Z_{D,i} & Z_{D,i}^* \\ Z_{D,i} & 1 \end{pmatrix} \quad B_i = \begin{pmatrix} P_{\theta_i}^{-1} & -P_{\theta_i}^{-1}\hat{\theta}_i \\ -\hat{\theta}_i^T P_{\theta_i}^{-1} & \hat{\theta}_i^T P_{\theta_i}^{-1}\hat{\theta}_i - \chi_i \end{pmatrix}$$

The above optimization problem is not explicitly function of $\rho_i(\omega)$. However, the optimal $\rho_i(\omega)$ can be obtained by taking the square root of the optimal $\alpha_i(\omega)$.

Proof. For conciseness, we will drop the frequency argument ω in the variables. Using the notations $\rho_i^2 = \alpha_i$ and $\bar{\theta}_i = (\theta_i^T \ 1)^T$ and using (29), we can rewrite the constraint in (28) as:

$$\bar{\theta}_i^T \left(A_i - \lambda_i^* \frac{-1}{\alpha_i} \lambda_i \right) \bar{\theta}_i < 0 \quad \forall \theta_i \in U_i \quad (31)$$

while the constraint $\theta_i \in U_i$ is equivalent to $\bar{\theta}_i^T B_i \bar{\theta}_i < 0$. Consequently, by virtue of the S-procedure [5] and Lemma 2 in [4], (31) holds if and only if there exist $\xi_i > 0$ and $\mathcal{X}_i = -\mathcal{X}_i^T$ such that

$$A_i - \lambda_i^* \frac{-1}{\alpha_i} \lambda_i - \xi_i B_i + j\mathcal{X}_i < 0 \quad (32)$$

Since $\alpha_i > 0$, an application of the Schur complement [5] shows that (32) is equivalent to (30). This concludes the proof. \blacksquare

Remark 2. The novelty of this theorem resides in the determination of the (complex) center c_i of the unstructured uncertainty. Given the definition of $\tilde{T}_i(e^{j\omega}, \theta_i)$, one could of course consider that this center is zero and just compute the radius $\rho_i(\omega)$. In this case, the LMI optimization above is the same as the one in [3]. However, since the mapping between θ_i and $\tilde{T}_i(e^{j\omega}, \theta_i)$ is in general not linear, this could lead to larger embedding sets \mathcal{T}_i and thus to more conservative results for the robustness analysis (as will be illustrated in Section 5).

Using Theorem 2, we can determine $\rho_i(\omega)$ and $c_i(\omega)$ for any value of i ($i = 1 \dots N_{mod}$) and for any value of ω . In this way, we fully determine the sets \mathcal{T}_i (see (26)) for all values of i and therefore also the set \mathcal{T} (see (22)). With this information, we will be able, as shown in the following theorem, to evaluate the worst case global performance $\mathcal{P}_{wc}(\omega, \mathcal{T})$ defined in (25) for any value of ω .

Theorem 3 Consider a given frequency ω and the set \mathcal{T} (see (22)) with a diagonal uncertainty $\Delta(e^{j\omega})$ whose elements $\Delta_i(e^{j\omega})$ are constrained to lie in a disk $\mathbf{\Delta}_i(\omega)$ of radius $\rho_i(\omega)$ and of center $c_i(\omega)$ (see (23)). Define R_ω (resp. C_ω) as a diagonal matrix of dimension N_{mod} whose elements are $\rho_i^2(\omega)$ (resp. $c_i(\omega)$) ($i = 1 \dots N_{mod}$). Then, an upper bound $\mathcal{P}_{wc,ub}(\omega, \mathcal{T})$ of the worst case global performance $\mathcal{P}_{wc}(\omega, \mathcal{T})$ defined in (25) is given by $\sqrt{\gamma_{opt}(\omega)}$ where

$\gamma_{opt}(\omega)$ is the solution of the following LMI optimization problem. This LMI optimization problem has as decision variables a real scalar $\gamma(\omega) > 0$ and a strictly positive definite diagonal matrix $T_\omega \in \mathbf{R}^{N_{mod} \times N_{mod}}$.

$$\begin{aligned} & \min \quad \gamma(\omega) \\ \text{s.t.} \quad & \begin{pmatrix} M(e^{j\omega}) \\ I \end{pmatrix}^* \mathcal{N}(\gamma(\omega)) \begin{pmatrix} M(e^{j\omega}) \\ I \end{pmatrix} < 0 \end{aligned} \quad (33)$$

$$\text{with } \mathcal{N}(\gamma(\omega)) \triangleq \left(\begin{array}{c|c} \left(\begin{array}{c|c} T_\omega(R_\omega - C_\omega^* C_\omega) & 0 \\ \hline 0 & 1 \end{array} \right) & \left(\begin{array}{c|c} T_\omega C_\omega^* & 0 \\ \hline 0 & 0 \end{array} \right) \\ \hline \left(\begin{array}{c|c} T_\omega C_\omega & 0 \\ \hline 0 & 0 \end{array} \right) & \left(\begin{array}{c|c} -T_\omega & 0 \\ \hline 0 & -\gamma(\omega) \end{array} \right) \end{array} \right) \quad (34)$$

This theorem can be straightforwardly deduced from the separation of graph theorem [22] and from the results in [7]. It follows from the fact that the constraint (33) is a sufficient condition for $|\mathcal{F}(M(e^{j\omega}), \Delta(e^{j\omega}))|^2 < \gamma(\omega) \forall \Delta(e^{j\omega}) \in \mathbf{\Delta}(\omega)$ to hold.

Remark 3. The so-called hierarchical approach for robustness analysis presented above is preferred upon a direct μ -analysis approach based on the structured uncertainty U when the number N_{mod} of modules is large. Indeed, even in this case, its computational complexity remains low while the μ -analysis approach would involve very complex multipliers of large dimensions. The radius $\rho_i(\omega)$ and the center $c_i(\omega)$ are indeed computed at the local level and the computation of (the upper bound of) the worst case performance $\mathcal{P}_{wc}(\omega, \mathcal{T})$ in Theorem 3 uses a very simple multiplier T_ω . This approach of course only yields an upper bound of $\sup_{\theta \in U} |P(e^{j\omega}, \theta)|$. However, the μ -analysis approach that would be used to compute $\sup_{\theta \in U} |P(e^{j\omega}, \theta)|$ would also only lead to an upper bound of this quantity which can turn out to be conservative for large N_{mod} . In the simulation example, we will show that the conservatism linked to the hierarchical approach remains limited.

Remark 4. Like the μ -analysis approach, the proposed robustness analysis approach is a frequency-wise approach (i.e. the upper bound $\mathcal{P}_{wc,ub}(\omega, \mathcal{T})$ on the worst case performance is computed frequency by frequency). Consequently, the performance constraint (19) can only be verified for a finite amount of frequencies of the frequency interval $[0 \ \pi]$.

4 Optimal experiment design for networks

Given an arbitrary identification experiment, the obtained confidence ellipsoid U and the corresponding uncertainty set \mathcal{T} can be such that the upper bound $\mathcal{P}_{wc,ub}(\omega, \mathcal{T})$ (computed using Theorem 3) is larger than the threshold $W(\omega)$ for some frequencies. In this section, we will design the experiment in order to avoid such a situation. More precisely, we will design the spectra Φ_{r_i} of the signals r_i ($i = 1 \dots N_{mod}$) for an identification experiment of (fixed) duration N in such a way that the total injected power is minimized while guaranteeing that the obtained uncertainty region \mathcal{T} is small enough to guarantee

$$\mathcal{P}_{wc,ub}(\omega, \mathcal{T}) < W(\omega) \quad (35)$$

at each frequency ω . Note that, for conciseness, we restrict attention to one single global performance objective (see also Remark 5 below). Note also that we here suppose that the signals r_i ($i = 1 \dots N_{mod}$) are all mutually independent. Consequently, the experiment is thus indeed entirely described by the spectra Φ_{r_i} ($i = 1 \dots N_{mod}$).

An important step is to parametrize these spectra $\Phi_{r_i}(\omega)$. Here, we will use the parametrization [18] i.e.: $\Phi_{r_i}(\omega) = \sigma_{i,0} + 2 \sum_{l=1}^M \sigma_{i,l} \cos(l\omega)$ ($i = 1 \dots N_{mod}$) for which $\Phi_{r_i}(\omega) > 0 \forall \omega$ can be guaranteed using an extra LMI constraint on the decision variables $\sigma_{i,l}$ ($i = 1 \dots N_{mod}, l = 0 \dots M$) [18]. With this parametrization, the cost function in our optimization problem is a linear function of the decision variables:

$$\mathcal{J} = \sum_{i=1}^{N_{mod}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{r_i}(\omega) d\omega = \sum_{i=1}^{N_{mod}} \sigma_{i,0} \quad (36)$$

Using (17), we see also that the matrices $P_{\theta_i}^{-1}$ that determines the size $\rho_i(\omega)$ of the elements \mathcal{T}_i of \mathcal{T} via the LMI (30) are all affine functions of the to-be-designed spectra and thus of the decisions variables $\sigma_{j,l}$ ($j = 1 \dots N_{mod}, l = 0 \dots M$).

By reducing \mathcal{J} , we increase the size of U and we thus increase $\rho_i(\omega) \forall i$. By increasing ρ_i , we increase $\mathcal{P}_{wc,ub}(\omega, \mathcal{T})$. Our goal is thus to find the spectra $\Phi_{r_i}(\omega)$ minimizing \mathcal{J} while leading to $\rho_i(\omega)$ that are sufficient small for (35) to hold for each ω in an user-chosen grid Ω of the frequency range. For this purpose, we propose the following optimal identification experiment design problem. This optimization problem has the following decision variables: $\sigma_{i,l}, T_\omega, \alpha_i(\omega) = \rho_i^2(\omega), c_i(\omega), \xi_i(\omega), \mathcal{X}_i(\omega)$ with the same structure as in Theorems 2 and 3 and defined for all $i = 1 \dots N_{mod}$, for all $l = 0 \dots M$ and for all $\omega \in \Omega$.

$\min \mathcal{J}$ under the constraints that, for all $\omega \in \Omega$,

$$\begin{pmatrix} M(e^{j\omega}) \\ I \end{pmatrix}^* \mathcal{N}(W^2(\omega)) \begin{pmatrix} M(e^{j\omega}) \\ I \end{pmatrix} < 0 \quad (37)$$

$$\mathcal{E}(\rho_i^2(\omega), c_i(\omega), \xi_i(\omega), \mathcal{X}_i(\omega), P_{\theta_i}^{-1}) < 0 \quad i = 1 \dots N_{mod}$$

The constraints of the above optimization have to be completed by the N_{mod} constraints on $\sigma_{i,l}$ guaranteeing that $\Phi_{r_i}(\omega) > 0 \forall \omega$ ($i = 1 \dots N_{mod}$). The matrix $\mathcal{E}(\rho_i^2(\omega), c_i(\omega), \xi_i(\omega), \mathcal{X}_i(\omega), P_{\theta_i}^{-1})$ above is the matrix defined in (30) in which $P_{\theta_i}^{-1}$ is replaced by its affine expression in the decision variables $\sigma_{j,l}$ defining the identification spectra $\Phi_{r_j} \forall j$. This optimization problem will lead to the least costly identification experiment (according to \mathcal{J}) that nevertheless guarantees that the identified models will be sufficiently accurate to deliver an acceptable global performance. Indeed, the combination of the set of constraints in (37) at a given ω guarantee that (35) holds for the uncertainty \mathcal{T} defined based on the uncertainty U delivered via the identification.

There are two issues with this optimization problem. First, as in all optimal identification experiment design problem, it depends on the unknown parameter vectors $\theta_{i,0}$ via $P_{\theta_i}^{-1}$, on the to-be identified parameter vectors $\hat{\theta}_i$ (via e.g. A_i) and on the to-be-designed controllers \hat{K}_i (via e.g. B_i). Those unknown variables will be replaced by initial guesses such as in [4,1]. In particular, we need to pre-select a control design method which, based on a model of the modules, leads to decentralized controllers achieving a nominal performance that is (slightly) better than the desired one. This control design method will be used to determine the initial guesses for \hat{K}_i based on the initial guesses for the modules.

The second issue is that the optimization problem is a bilinear problem via the products $T_\omega(R_\omega - C_\omega C_w^*)$, $T_\omega C_\omega$ in \mathcal{N} (see (34)) and $\xi_i B_i$ (and thus $\xi_i P_{\theta_i}^{-1}$) in the constraint $\mathcal{E} < 0$ (see (30)). Indeed, here, unlike in Section 3, all these variables are together decision variables in the same optimization problem. To tackle this issue, we propose the following iterative algorithm inspired by the so called D-K iterations [28]. Before presenting the algorithm, we note that, if we arbitrarily choose spectra $\Phi_{r_i}(\omega)$ ($i = 1 \dots N_{mod}$), we can compute the corresponding P_{θ_i} ($i = 1 \dots N_{mod}$) via (17). With these P_{θ_i} , we can compute $\mathcal{P}_{wc,ub}(\omega, \mathcal{T})$ via Theorems 2 and 3 and we can therefore verify whether (35) holds at each $\omega \in \Omega$. If that is the case, we will say that the spectra $\Phi_{r_i}(\omega)$ ($i = 1 \dots N_{mod}$) are validated.

Algorithm 1. *The algorithm is made up of an initialization step (step 0) and each iteration consists of three steps.*

- S.0. *We initialize the algorithm by arbitrarily choosing the spectra Φ_{r_i} ($i = 1 \dots N_{mod}$) (e.g. $\Phi_{r_i}(\omega) = 1$).*
- S.1. *Using a subdivision algorithm, we determine, using the notion of validation defined above, the minimal positive scalar $\gamma \in \mathbf{R}$ such that the spectra $\gamma \Phi_{r_i}(\omega)$ ($i = 1 \dots N_{mod}$) remain validated. Denote this minimal γ by γ_{min} .*
- S.2. *To validate $\gamma_{min} \Phi_{r_i}(\omega)$, the optimization problems in Theorems 2 and 3 have been used. The corresponding decision variables are $\alpha_i = \rho_i^2(\omega)$, $c_i(\omega)$, $\xi_i(\omega)$, $\mathcal{X}_i(\omega)$ and T_ω ($i = 1 \dots N_{mod}$, $\omega \in \Omega$). From those decisions variables, $\xi_i(\omega)$, $c_i(\omega)$ and T_ω ($i = 1 \dots N_{mod}$, $\omega \in \Omega$) are conserved for Step 3.*
- S.3. *The optimal experiment design problem (37) is transformed into an LMI optimization problem by fixing the decision variables $\xi_i(\omega)$, $c_i(\omega)$ and T_ω ($i = 1 \dots N_{mod}$, $\omega \in \Omega$) to the ones determined in Step 2. The solution of this transformed optimization problem define, via $\sigma_{i,l}$, new spectra $\Phi_{r_i}(\omega)$. These new spectra $\Phi_{r_i}(\omega)$ can then be used in Step 1 for a new iteration.*

The algorithm is stopped when the optimal cost \mathcal{J}_{opt} in Step 3 no longer decreases significantly after each iteration. The optimal spectra $\Phi_{r_i}(\omega)$ are then the ones corresponding to this last iteration (Step 1 can be used a last time to further refine these spectra).

Remark 5. The optimal experiment design problem has been presented in the case where the objective is to guarantee a certain level of global performance described by one transfer function P . However, it is straightforward to extend it to the case where different transfer functions P are considered and, using the results in [4], also to the case where, in addition, a certain level of local performance has to be guaranteed.

Remark 6. Let us now consider briefly the case where the noises e_i in (1) are not independent, but spatially correlated with a strictly positive covariance matrix Λ . We had already observed in Remark 1 that Theorem 1 also holds in that case. As opposed to this, the individual SISO identification criteria (7) are, for e_i having that property, no longer equivalent to the global MIMO identification criterion (13) and will therefore not lead to the smallest variance. However, since $\epsilon_i(t, \theta_{i,0})$ remains equal to $e_i(t)$, $\sqrt{N}(\hat{\theta}_i - \theta_{i,0})$ will in that case too be asymptotically normally distributed around zero and the covariance matrix P_{θ_i} of $\hat{\theta}_i$ will be still given by (17). We also observe that the hierarchical approach only uses the ellipsoids U_i (and not U). Consequently, the controller validation and experiment design results can also be applied when the white noises e_i ($i = 1 \dots N_{mod}$) are spatially correlated. Indeed, even though the covariance matrix P_θ of $\hat{\theta}$ is no longer block-diagonal in that case, the projections $U_i = \{\theta_i \mid \theta \in U\}$ are still given by (18) [2].

5 Numerical illustration

In this numerical illustration, we consider the network of Figure 1 made up of six nodes ($N_{mod} = 6$). We consider here the case of nodes made up of homogenous systems and, for simplicity, the true systems \mathcal{S}_i that will be identified will all be identical and given by the following ARX system [20]: $y_i(t) = (z^{-3}B_0(z))/(A_0(z))u_i(t) + (1)/(A_0(z))e_i(t)$ with $B_0(z) = 0.10276 + 0.18123z^{-1}$, $A_0(z) = 1 - 1.99185z^{-1} + 2.20265z^{-2} - 1.84083z^{-3} + 0.89413z^{-4}$. The variances $\Lambda_{i,i}$ of the white noises $e_i(t)$ are all equal to one i.e. $\Lambda = I_6$. We suppose that these true systems are all controlled by the same local controller K that is designed using the local method in [11]. The global performance $P(z, \theta_0)$ that we consider in this example is the transfer function between $w = ref_{ext}$ and $s = y_6 - ref_{ext}$. Our objective is to determine the identification experiment leading to models whose uncertainty U is small enough to guarantee (35) with the threshold W represented in Figure 3. This threshold requires a global bandwidth that is higher than the one achieved with the controller K present in the network. We suppose that, based on the identified models, the method of [19] will be used to design a unique controller \hat{K} that satisfies the global performance defined by W (but also a certain level of local performance).

Using the methodology presented in Section 4, we design the spectra Φ_{r_i} ($M = 10$) that have to be applied to each module for the identification. For this design, we fix the experiment length to $N = 2000$ and we need to have initial guesses of the true systems, the identified models and \hat{K} . We have here used for this purpose an ARX system having the same structure as the true systems, but described by $B_{init}(z) = 0.1192 + 0.1651z^{-1}$ and $A_{init}(z) = 1 - 2.009z^{-1} + 2.23z^{-2} - 1.859z^{-3} + 0.9004z^{-4}$. With this initial guess \mathcal{S}_{init} for the system, we have designed a controller K_{init} using the same global method that will be used with the identified models. This controller K_{init} has been used as initial guess for the controller \hat{K} . The worst-case performance $\mathcal{P}_{wc,ub}(\omega, \mathcal{T})$ corresponding to these designed spectra and these initial guesses is given in Figure 3 (black dashed) and we see that the performance objective is satisfied. The optimal cost \mathcal{J}_{opt} is equal to 108. This result has been obtained by considering the centers $c_i(\omega)$ as decision variables. This is important since the optimal cost would be equal to 133 if we would force these centers to be zero for all i and for all frequencies. We have also verified that the conservatism linked to the chosen hierarchical robustness analysis approach remains limited. For this purpose, we have computed a lower bound of the exact worst-case performance $\sup_{\theta \in U} |P(e^{j\omega}, \theta)|$ by randomly generating parameter vectors θ on the contour of the obtained uncertainty ellipsoid U and by computing, for each ω , the largest value of $|P(e^{j\omega}, \theta)|$ for those random values of θ . When generating 10000 random θ , the relative error between this lower bound and $\mathcal{P}_{wc,ub}(\omega, \mathcal{T})$ is, for each ω , smaller than 20% (the maximal relative error is attained at $\omega \approx 1.5$) and has a mean of 3% over the frequencies.

In order to further verify the validity of our results, we realize the optimal spectra and apply the corresponding excitation signals r_i ($i = 1 \dots 6$) of length $N = 2000$ to the network. The models $G(z, \hat{\theta}_i)$ are identified using the procedure of Section 2 and a controller \hat{K} is designed with the average of these six models using the chosen global control design method. Finally, using the covariance matrices P_{θ_i} determined along with the models, we use the procedure of Section 3 to determine the worst-case global performance. This worst-case global performance $\mathcal{P}_{wc,ub}(\omega, \mathcal{T})$ is given in Figure 3 (red solid). We observe that, even if the optimal spectra have been designed with initial guesses for the identified parameter vectors and for \hat{K} as well as with the asymptotic formula (9) for the covariance matrices P_{θ_i} , the worst-case performance actually obtained after the identification experiment satisfies (35).

i	$\sigma_{i,0}$	V_i
$i = 1$	16.2	481
$i = 2$	19.5	238
$i = 3$	18.1	179
$i = 4$	18.6	181
$i = 5$	22.6	154
$i = 6$	13.8	101

Table 1
Distribution of power and of uncertainty among the nodes

Consequently, our methodology leads to models that are sufficiently accurate to guarantee a certain level of global performance in this example. Let us analyze how the excitation power and the uncertainty is distributed among the nodes. For this purpose, we give in Table 1 the power injected for each node (i.e. $\sigma_{i,0}$) and a normalized image of the volume of U_i i.e. $V_i = \det(10^5 P_{\theta_i})$. Since V_i is only one of the possible measures to evaluate the size of the uncertainty, we also represent, in Figure 4, the radius $\rho_i(\omega)$ (representing the size of \mathcal{T}_i and computed using Theorem 2) for Nodes $i = 1$ (black dotted), $i = 5$ (black circles) and $i = 6$ (black solid). We observe that both the excitation powers and

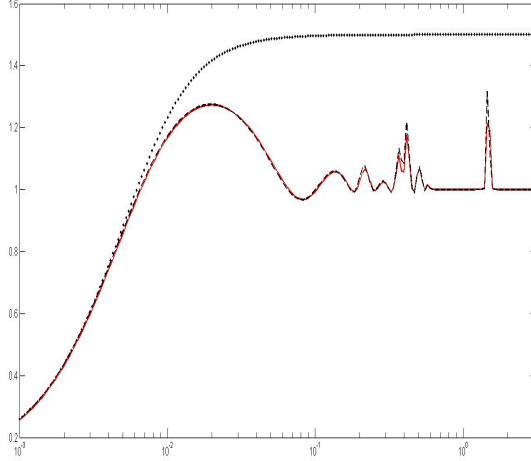


Fig. 3. Desired global performance $W(\omega)$ (black dotted), expected worst case performance $\mathcal{P}_{wc,ub}(\omega, \mathcal{T})$ after optimal experiment design and using the initial guesses (black dashed), obtained worst case performance $\mathcal{P}_{wc,ub}(\omega, \mathcal{T})$ after identification and redesign of the controllers (red solid)

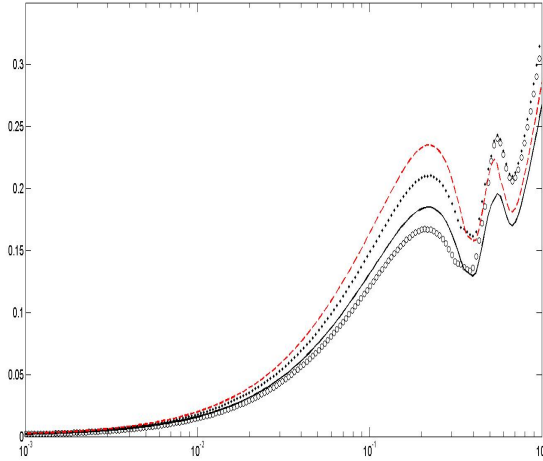


Fig. 4. $\rho_1(\omega)$ (black dotted), $\rho_5(\omega)$ (black circles), $\rho_6(\omega)$ (black solid) and $\rho_6(\omega)$ if Node 6 would be isolated (red dashed) in the frequency range $[0, 1]$

the sizes of the obtained uncertainties are different at different nodes. As an example, more uncertainty is allowed in Node 1 than in Node 6. Interestingly, the smaller uncertainty in Node 6 is obtained with less power than the larger uncertainty in Node 1: $\sigma_{6,0} < \sigma_{1,0}$. This is a consequence of the network configuration. Indeed, the uncertainty of Node 6 is not only function of $\Phi_{r,6}$, but also of all the other spectra applied during the identification experiment (see (17)). To illustrate this, we have also represented, in red dashed in Figure 4, the radius $\rho_6(\omega)$ that would have been obtained if Node 6 would have been isolated as Node 1 (and thus if P_{θ_6} would have been determined uniquely with $\Phi_{r,6}$). We observe that the obtained uncertainty would have been much larger.

6 Concluding remarks

This paper is one of the first contributions on optimal experiment design in a network context. The type of networks considered in this paper is usual in the literature on multi-agent systems. We have seen that many results of this paper not only apply to systems \mathcal{S}_i with independent white noises e_i , but also to spatially correlated ones. However, as mentioned in Remark 6, if the white noises e_i are spatially correlated, our identification procedure using individual SISO criteria is no longer optimal since it is no longer equivalent to the global MIMO prediction error identification criterion. Future work will therefore consider the question of how to deal with this MIMO criterion in

the case of spatially correlated noises without increasing too much the computational complexity. This complexity is indeed an important feature when the number N_{mod} of modules is large.

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, B. Anderson, and M. Gevers. Quantification of frequency domain error bounds with guaranteed confidence level in prediction error identification. *Systems and Control Letters*, 54(5):471–482, 2005.
- [3] X. Bombois, M. Gevers, G. Scorletti, and B.D.O. Anderson. Robustness analysis tools for an uncertainty set obtained by prediction error identification. *Automatica*, 37(10):1629–1636, 2001.
- [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] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in Systems and Control Theory*, volume 15 of *Studies in Appl. Math.* SIAM, Philadelphia, June 1994.
- [6] A. Dankers, P. Van den Hof, X. Bombois, and P. Heuberger. Identification of dynamic models in complex networks with prediction error methods - predictor input selection. *IEEE Transactions on Automatic Control*, 61(4):937–952, 2016.
- [7] M. Dinh, A. Kornienko, and G. Scorletti. Convex hierarchical analysis for the performances of uncertain large-scale systems. In *Proc. 53rd IEEE Conference on Decision and Control*, pages 5979–5984, Los Angeles, 2014.
- [8] J. Doyle, A. Packard, and K. Zhou. Review of LFT’s, LMI’s and μ . In IEEE, editor, *Proc. IEEE Conf. on Decision and Control*, volume 2, pages 1227–1232, Brighton, England, December 1991.
- [9] N. Everitt, G. Bottegal, C. Rojas, and H. Hjalmarsson. On the variance analysis of identified linear MIMO models. In *Proc. 54th IEEE Conference on Decision and Control*, pages 1447–1452, Osaka, Japan, 2015.
- [10] J.A. Fax and R.H. Murray. Information flow and cooperative control of vehicle formations. *IEEE Transactions on Automatic Control*, 49(9):1465–1476, 2004.
- [11] G. Ferreres and V. Fromion. H_∞ control for a flexible transmission system. In *CD-ROM Proc. European Control Conference*, Brussels, Belgium, 1997.
- [12] 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.
- [13] 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.
- [14] B. Gunes, A. Dankers, and P. Van den Hof. A variance reduction technique for identification in dynamic networks. In *Proc. IFAC World Congress*, pages 2842–2847, Cape Town, South Africa, 2014.
- [15] A. Haber and M. Verhaegen. Moving horizon estimation for large-scale interconnected systems. *IEEE Transactions on Automatic Control*, 58(11):2834–2847, 2013.
- [16] P. Hägg and B. Wahlberg. On identification of parallel cascade serial systems. In *Proc. IFAC World Congress*, pages 9978–9983, Cape Town, South Africa, 2014.
- [17] P. Hägg and B. Wahlberg. On optimal input design for networked systems. *Automatica*, 53:275 – 281, 2015.
- [18] 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.
- [19] A. Kornienko, G. Scorletti, E. Colinet, and E. Blanco. Performance control for interconnection of identical systems: Application to pll network design. *International Journal of Robust and Nonlinear Control*, 2014.
- [20] I.D. Landau, D. Rey, A. Karimi, A. Voda, and A. Franco. A flexible transmission system as a benchmark for robust digital control. *European Journal of Control*, 1(2):77–96, 1995.
- [21] L. Ljung. *System Identification: Theory for the User, 2nd Edition*. Prentice-Hall, Englewood Cliffs, NJ, 1999.
- [22] M. G. Safonov. *Stability and Robustness of Multivariable Feedback Systems*. MIT Press, Cambridge, 1980.
- [23] M. G. Safonov. Propagation of conic model uncertainty in hierarchical systems. *IEEE Transactions on Circuits and Systems*, pages 388–396, June 1983.
- [24] G. Scorletti and G. Duc. An LMI approach to decentralized H_∞ control. *International Journal of Control*, 74(3):211–224, 2001.
- [25] T. Vincent, C. Novara, K. Hsu, and K. Poolla. Input design for structured nonlinear system identification. *Automatica*, 46:990–998, 2010.
- [26] H. Weerts, A. Dankers, and P. Van den Hof. Identifiability in dynamic network identification. In *Proc. IFAC Symposium on System Identification*, pages 1409–1414, Beijing, China, 2015.
- [27] H. Weerts, P. Van den Hof, and A. Dankers. Identifiability of dynamic networks with part of the nodes noise-free. In *Proc. IFAC International Workshop on Adaptation and Learning in Control and Signal Processing*, Eindhoven, The Netherlands, 2016.
- [28] K. Zhou and J. Doyle. *Essentials of Robust Control*. Prentice Hall, Upper Saddle River, New Jersey, 1998.