Identifiability of dynamical networks with partial node measurements*

1st Julien M. Hendrickx
ICTEAM
Université catholique de Louvain
B1348 Louvain la Neuve, Belgium
Julien.Hendrickx@uclouvain.be

2nd Michel Gevers
ICTEAM
Université catholique de Louvain
B1348 Louvain la Neuve, Belgium
Michel.Gevers@uclouvain.be

3rd Alexandre S. Bazanella
Department of Automation and Energy
Universidade Federal do Rio Grande do Sul
Porto Alegre, Brazil
bazanella@ufrgs.br

Abstract—Much recent research has dealt with the identifiability of a dynamical network in which the node signals are connected by causal linear transfer functions and are excited by known external excitation signals and/or unknown noise signals. A major research question concerns the identifiability of the whole network - topology and all transfer functions - from the measured node signals and external excitation signals. So far all results on this topic have assumed that all node signals are measured. This paper presents the first results for the situation where not all node signals are measurable, under the assumptions that (1) the topology of the network is known, and (2) each node is excited by a known external excitation. Using graph theoretical properties, we show that the transfer functions that can be identified depend essentially on the topology of the paths linking the corresponding vertices to the measured nodes. A practical outcome is that, under those assumptions, a network can often be identified using only a small subset of node measurements.

Index Terms—Network Analysis and Control; System identification.

I. Introduction

This paper examines the identifiability of dynamical networks in which the node signals are connected by causal linear time-invariant transfer functions and are excited by known external excitation signals. Such networks can be looked upon as connected directed graphs in which the edges between the nodes (or vertices) are scalar transfer functions, and in which known external excitation signals enter into the nodes.

The identification of networks of linear time-invariant dynamical systems based on the measurement of all its node signals and of all known external excitation signals acting on the nodes has been the subject of much recent research [1]–[7]. It has been shown in [1], [5], [7] that identifiability can only be obtained provided prior knowledge is available about the structure of the network, and in particular the structure of the excitation. It is often the case that the excitation structure is known, i.e. one often knows at which nodes external excitation signals are applied. A number of conditions for the identifiability of the whole network have been derived under prior assumptions on the structure of the network, involving either its external excitation structure, or possibly also its internal structure [1], [5]–[7].

In all the results accumulated so far on the identifiability of a network of dynamical systems, it is assumed that all node signals are measured. In this paper we examine the situation where not all node signals are measured, but where the topology of the network is known; this means that the user knows a priori which nodes are connected by nonzero transfer functions. We also make the simplifying assumption that at each node a known external excitation is applied. In this context, a number of questions can be raised, such as

1) Can one identify the whole network with a restricted number of node measurements?
2) If so, are there a minimal number of nodes that need to be measured?
3) Are some nodes indispensable, in the sense that it is impossible to identify the network without measuring these nodes?
4) If one wants to identify a specific transfer function, can the topology tell us which node or nodes need to be measured?
5) Which transfer functions can be identified from the measure of a specific subset of nodes?

To answer these questions we shall heavily rely on properties from graph theory, using the connected directed graph corresponding to our network as our major tool.

To the best of our knowledge, the only other contributions that consider identification in networks using only a subset of measured nodes are [8], [9], [10]. However, the problem treated in these papers consists of the identification of a subset of the network's transfer functions - typically a single one - and hence is only one of the subproblems presented in this paper. In [8] networks driven only by a vector of white noises are considered, i.e. no known external excitation is available. Using the notion of d-separation of graphs, the authors derive sufficient conditions on which node signals need to be observed in order to guarantee the identifiability of a desired transfer function link. In [9] the objective is also to identify a specific transfer function link (or module), in a network that does have both known external excitation signals and/or noise

signals on the nodes. The authors present sufficient conditions for the selection of a set of measured node signals that will lead to the consistent identification of the desired module. The approach taken in [10] is quite different. It consists of estimating the desired but unobservable nodes from nodes that are measurable and contain information about them. Sparse measurements have also been considered in a different context in [11]; the goal there was to recover the network structure under the assumptions that the local dynamics are known, as opposed to re-identifying the dynamics and/or the whole network structure.

Our main contribution is to provide necessary and sufficient conditions under which all transfer functions of the network, or a subset of transfer functions, or a single transfer function can be identified from a given set of measured nodes, under the standing assumptions that the topology of the network is known and that there is a known external excitation on each node. Our results are existence results about identifiability; they are not algorithms for the estimation of the transfer functions. They all take the form of conditions on the topology of the graph associated to the network. We also present the computational complexity that is required to check these necessary and sufficient conditions.

In Section II we first describe the standard network matrix identifiability problem where all nodes are measured but where the topology is unknown and needs to be identified from data. We explain that without any knowledge of the topology the identification of the network's transfer functions from partial node measurements has no solution. We then show that, in order to relate the identifiability of a set of transfer functions to the selection of a set of measured nodes on the basis of the network topolgy, one needs to introduce the notion of generic identifiability. This notion is described intuitively in Section II together with a motivating example.

We then motivate the reason for addressing the problem of network identifiability with partial node measurements in Section III by analyzing three different 3-node networks. We show that the nodes that need to be measured to identify all transfer functions depend on the topology of the network and that, in some cases, a unique measurement suffices to identify the whole network. This already yields a positive answer to question 1 above. Our brief analysis of 3-node networks then leads us, in Section IV, to formulate a number of basic results pertaining to questions 2, 3 and 4 above. We also provide identifiability results for networks that have a special structure, such as a tree or a loop.

In Section V we focus on the identifiability of the transfer functions leaving a specific node i, i.e. the transfer functions G_{ji} that connect node i to its outgoing nodes. Our main result in that Section is a necessary and sufficient condition for the identifiability of a set of transfer functions leaving node i. This set of transfer functions is shown to be identifiable from a given set of measured nodes if and only if there are disjoint paths going from these outgoing nodes of i to the set of measured nodes.

In Section VI we address question 5 above. Instead of looking at a specific node within the network and examining its paths to a measured node or a set of measured nodes, as was done in Section V, we consider the converse approach. We consider a specific set of measured nodes and we ask which transfer functions can be identified from it. Our main result is a necessary and sufficient condition under which the whole network can be identified from a given set of node measurements.

In Section VII we examine the computational complexity of the algorithm to check the identifiability of the whole network or parts of it from a given set of measurements. We show for example that checking the identifiability of the whole network can be achieved at a computational cost of the order of $L^2 \times n$ where L is the number of nodes and n the number of unknown transfer functions in the network.

In Section VIII we will conclude and describe some challenging open problems that remain to be solved.

II. STATEMENT OF THE PROBLEM

The problem studied in this paper is part of the recent research on the question of identifiability of networks of dynamical systems. We first present the network structure and explain the network identifiability problem as it has so far been posed, i.e. with all nodes measured. We then pose a new network identifiability problem for the case when not all nodes are measured.

We adopt the standard network structure of [5], [7] for networks whose edges are labeled with scalar proper transfer functions. Thus, we consider that the network is made up of L nodes, with node signals denoted $\{w_1(t), \ldots, w_L(t)\}$, and that these node signals are related to each other and to external excitation signals $r_j(t), j = 1, \ldots, L$ by the following network equations, which we call the **network model** and in which the matrix $G^0(q)$ is called the **network matrix**:

$$w(t) = G^0(q)w(t) + r(t) + v(t). \tag{1}$$

In (1) q^{-1} is the delay operator, $w(t) = [w_1(t), \ldots, w_L(t)]^T$ is the vector of node signals, $r(t) = [r_1(t), \ldots, r_L(t)]^T$ is a vector of known external excitation signals, $v(t) = [v_1(t), \ldots, v_L(t)]^T$ is a vector of stochastic processes, and the dynamic network matrix is of the form

$$G^0(q) = \begin{bmatrix} 0 & G^0_{12}(q) & \cdots & G^0_{1L}(q) \\ G^0_{21}(q) & 0 & \ddots & G^0_{2L}(q) \\ \vdots & \ddots & \ddots & \vdots \\ G^0_{L1}(q) & G^0_{L2}(q) & \cdots & 0 \end{bmatrix}$$

The network (1) is assumed to have the following properties.

- $G^0_{ij}(q)$ are proper rational transfer functions
- the network is well-posed, that is $(I - G^0(q))^{-1}$ is proper and stable [12]
- there is a known external excitation signal $r_i(t)$ on each node; these are available to the user in order to produce informative experiments for identification

- the network is weakly connected[1]

In most papers on identifiability of networks based on measurements of all the nodes, the vector $r(t)$ of external excitation signals traditionally enters the nodes via a transfer function matrix $K^0(q)$, i.e. the driving term is $r(t) = K^0(q)\tilde{r}(t)$ where $\tilde{r}(t)$ is the vector of external excitations. In this paper on network identifiability using partial node measurements, we adopt the simplified network model (1) where $K^0(q) = I$. Observe that, by a simple change of variables, this is equivalent to assuming that in the traditional model the excitation matrix $K^0(q)$ is known and of full rank. The reason for making this simplifying assumption is that, as we shall see, the problem treated in this paper, even with this assumption, is complex enough and reveals significant new insights. We expect to be able to relax this assumption in future work.

The network model (1) can be rewritten in a more traditional input-output (I/O) form as follows:

$$w(t) = T^0(q)r(t) + \bar{v}(t) \tag{2}$$

where

$$T^0(q) \triangleq (I - G^0(q))^{-1} \tag{3}$$
$$\bar{v}(t) = (I - G^0(q))^{-1}v(t). \tag{4}$$

In this paper we address the question of the identifiability of the network matrix $G^0(q)$ for the case where not all nodes are measured, but where the topology of the network is known. The reason for the assumption on known topology is that, as we shall show in Theorem 5.3, when not all nodes are measured, some knowledge of the topology is required in order to identify the whole network (in the absence thus of any prior knowledge on the specific transfer functions $G_{ij}(q)$).

Thus we assume that we know that certain transfer functions $G_{ij}(q)$ are zero, and we say that a network matrix $G(q)$ is *consistent* with the topology if it satisfies these constraints. Moreover, we consider that, together with the network (1), there is a measurement equation

$$y(t) = Cw(t) \tag{5}$$

where C is a $p \times L$ matrix that reflects the selection of measured nodes. That is, each row of C contains one element 1 and $L-1$ elements 0. We shall denote by $\mathcal{C}$ the corresponding subset of nodes selected by C.

In this setting, the network under study is given by

$$w(t) = G^0(q)w(t) + r(t) + v(t) \tag{6}$$
$$y(t) = Cw(t) \tag{7}$$

which, in the input-output form, becomes

$$y(t) = CT^0(q)r(t) + C\bar{v}(t) \tag{8}$$

with $T(q)$ and $\bar{v}(t)$ defined by (3) and (4).

We now describe the **network matrix identifiability** problem for such networks; we start by summarizing the assumptions that are made throughout this paper.

---

[1]A precise definition will be given in Section IV.

Standing assumptions.

- The networks we examine are described by (6)-(7).
- The network matrix $G^0(q)$ has the properties defined above and its topology is known, i.e. one knows *a priori* that some of the $G_{ij}^0(q)$ are zero.
- The excitation vector $r(t)$ is sufficiently rich such that $CT^0(q)$ can be consistently estimated by standard identification of the open loop MIMO I/O model (8).

Since, for a given C, the matrix $CT^0(q)$ can be consistently identified from $\{y(t), r(t)\}$ data, it will be assumed to be known exactly. The **network matrix identifiability** problem is whether or not, under the standing assumptions, one can uniquely recover $G^0(q)$ from $CT^0(q)$. Specific questions related to network identifiability that are addressed in this paper are then:

- for a given C, which transfer functions $G_{ij}^0(q)$ can be uniquely recovered from $CT^0(q)$?
- under what conditions can we identify the whole network matrix $G^0(q)$ from $CT^0(q)$?

The identification of the transfer functions $G_{ij}(q)$ from $CT^0(q)$ rests on the following relationship

$$CT^0(q) = CT(q) = C(I - G(q))^{-1} \tag{9}$$

or, equivalently,

$$CT^0(q)(I - G(q)) = C \tag{10}$$

Since $CT^0(q)$ is assumed known, the question is whether the desired $G_{ij}(q)$ can be uniquely obtained by solving (10) for these unknowns, using the knowledge of the network topology. More precisely, we say that the network matrix G^0 is identifiable from the measurements C if it is the unique solution of (10) consistent with the topology. Similarly, a specific transfer function in G^0 is identifiable from the measurements C if $G_{ij} = G_{ij}^0$ for any solution G of (10) consistent with the topology.

Deciding whether $G^0(q)$ is uniquely recoverable from the identified and exact $CT^0(q)$ can thus be done by checking whether the solution $G(q)$ of (10) is unique. However, this is of limited interest because it does not take account of the information we have about the known topology of the network. Our ambition in this paper is to make statements about the identifiability of $G^0(q)$ for a given selection $\mathcal{C}$ of node measures before we actually compute $CT^0(q)$ from data or even collect the data, i.e. statements that are based not on the actual numbers that appear in the transfer functions of $CT^0(q)$, but on the topology of the network that is assumed to be known, and which can be represented by a graph associated to $G^0(q)$.

As a consequence, we will introduce the notion of *generic identifiability* of $G^0(q)$ because the topology tells us which of the $G_{ij}(q)$ can be nonzero, which impacts on the generic rank of $CT^0(q)$ and of its submatrices[2], but one cannot exclude the

---

[2]By the generic rank of a submatrix of $CT^0(q)$ we mean its rank for almost all $G^0(q)$ that are consistent with its associated graph.

possible situation where a given $G^0(q)$, that is consistent with the topology, happens to cause a drop in rank of $CT^0(q)$ or of its submatrices. Thus, a statement like: "The network matrix $G^0(q)$ that is consistent with a given topology is generically identifiable from a given choice $\mathcal{C}$ of measurements" will mean that $G^0(q)$ is identifiable for almost all choices of the elements $G_{ij}^0(q)$ of $G^0(q)$ that are not known to be zero.

We shall define this new notion of generic identifiability of the network in precise terms in Section V. In order to give the reader an intuitive feeling for this notion, we illustrate it with the following example.[3]

Example 1: Consider a network whose topology is defined by the following network matrix:

$$G^0 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ G_{21}^0 & 0 & 0 & 0 & 0 \\ G_{31}^0 & 0 & 0 & 0 & 0 \\ 0 & G_{42}^0 & G_{43}^0 & 0 & 0 \\ 0 & G_{52}^0 & G_{53}^0 & 0 & 0 \end{pmatrix}$$

and suppose we measure nodes 4 and 5 only. Simple calculations show that

$$CT^0 = \begin{pmatrix} G_{42}^0 G_{21}^0 + G_{43}^0 G_{31}^0 & G_{42}^0 & G_{43}^0 & 1 & 0 \\ G_{52}^0 G_{21}^0 + G_{53}^0 G_{31}^0 & G_{52}^0 & G_{53}^0 & 0 & 1 \end{pmatrix}$$

Clearly, from CT^0 we can uniquely identify $G_{42}^0, G_{43}^0, G_{52}^0, G_{53}^0$. The remaining elements, G_{21} and G_{31} are then recovered from CT^0 by solving

$$\begin{pmatrix} G_{42}^0 & G_{43}^0 \\ G_{52}^0 & G_{53}^0 \end{pmatrix} \begin{pmatrix} G_{21} \\ G_{31} \end{pmatrix} = \begin{pmatrix} T_{41}^0 \\ T_{51}^0 \end{pmatrix}$$

We conclude that G^0 is generically identifiable from measurements of nodes 4 and 5 only, because it is identifiable for almost all network matrices G consistent with the topology, namely all except those for which $G_{42}^0 G_{53}^0 = G_{52}^0 G_{43}^0$, which is a subset of measure zero.

III. Motivating examples

In order to motivate the reader, we now analyze a few 3-node networks and show that the nodes that allow identification of the whole network depend entirely on the topology of the network, and that the whole network can often be identified from the measurements of a small subset of nodes.

Consider first a network with 3 unknown transfer functions represented in Figure 1 and its corresponding true G^0 and true T^0. Calculations based on (10) show that identification of all 3 transfer functions requires the measurement of nodes 2 AND 3, and that measuring node 1 yields no information.

By contrast, the identification of the 3 unknown transfer functions in the network represented in Figure 2 is possible by measuring just one node: node 1 OR node 3.

Finally, in the network of Figure 3, all 5 transfer functions can be identified by measuring just two nodes: either nodes 1 AND 2, OR nodes 1 AND 3.

[3]Starting in this example and for the remaining of the paper, we omit the dependence on q whenever it creates no confusion.

$$G^0 = \begin{bmatrix} 0 & 0 & 0 \\ G_{21}^0 & 0 & G_{23}^0 \\ G_{31}^0 & 0 & 0 \end{bmatrix},$$

$$T^0 = \begin{bmatrix} 1 & 0 & 0 \\ G_{21}^0 + G_{23}^0 G_{31}^0 & 1 & G_{23}^0 \\ G_{31}^0 & 0 & 1 \end{bmatrix}$$

Fig. 1. Example of network with three transfer functions where two nodes (2 and 3) need to be measured.

$$G^0 = \begin{bmatrix} 0 & G_{12}^0 & G_{13}^0 \\ 0 & 0 & 0 \\ G_{31}^0 & 0 & 0 \end{bmatrix},$$

$$T^0 = \frac{1}{\Delta} \begin{bmatrix} 1 & G_{12}^0 & G_{13}^0 \\ 0 & 1 - G_{13}^0 G_{31}^0 & 0 \\ G_{31}^0 & G_{31}^0 G_{12}^0 & 1 \end{bmatrix}$$

Fig. 2. Example of network with three transfer functions where measuring one node (1 or 3) is sufficient. We use $\Delta \triangleq det(I - G^0) = 1 - G_{13}^0 G_{31}^0$

$$G^0 = \begin{bmatrix} 0 & G_{12}^0 & G_{13}^0 \\ G_{21}^0 & 0 & G_{23}^0 \\ 0 & G_{32}^0 & 0 \end{bmatrix},$$

$$T^0 = \frac{1}{\Delta} \begin{bmatrix} 1 - G_{23}^0 G_{32}^0 & G_{12}^0 + G_{13}^0 G_{32}^0 & G_{13}^0 + G_{12}^0 G_{23}^0 \\ G_{21}^0 & 1 & G_{23}^0 + G_{21}^0 G_{13}^0 \\ G_{32}^0 G_{21}^0 & G_{32}^0 & 1 - G_{12}^0 G_{21}^0 \end{bmatrix}$$

Fig. 3. Example of network with five transfer functions where measuring two nodes (1 and 2 or 1 and 3) is sufficient. We use $\Delta \triangleq 1 - G_{12}^0 G_{21}^0 - G_{23}^0 G_{32}^0 - G_{13}^0 G_{21}^0 G_{32}^0$.

These examples show that the number and the choice of measurements that are necessary to identify the network depends not only on the number of unknown transfer functions to be determined (the number of nonzero G_{ij}) but also on the topology of the network.

IV. Basic results

Inspired by our analysis of 3-node networks, we now establish a number of basic results regarding the identifiability of general L-node networks from a reduced set of node measurements. In particular we show that measurements of some nodes is indispensable, and we establish the minimum number of nodes that need to be measured for the identifiability of G^0.

We first introduce some notations and we define some concepts from graph theory (see e.g. [13], [14]). Observe that the graph associated with the network model is a graph with directed edges, i.e. if G_{ij}^0 is nonzero, it means that there is a directed edge from node j to node i. Conversely, we say that a transfer matrix G (or any matrix) is *consistent* with a directed graph if $G_{ij} \neq 0$ only if there is an edge (j, i). Observe that

the presence of an edge does not require the corresponding entry G_{ij} to be different from zero.

Notations and definitions:

- L = number of nodes;
- p = number of measured nodes;
- s = number of sinks, i.e. number of nodes with only incoming edges;
- n = number of unknown transfer functions;
- C = the $p \times L$ matrix that reflects the selection of nodes via $y(t) = Cw(t)$: thus each row of C contains one element 1 and $L-1$ elements 0;
- $\mathcal{C}$ = the subset of nodes selected by C;
- G_{TS}^0 = the restriction of the network matrix G^0 to the rows contained in a set T and the columns contained in a set S;
- $|\mathcal{A}|$ = cardinality of a set $\mathcal{A}$;
- N_i^+ = set of out-neighbors of node i, i.e. the set of nodes j for which $G_{ji}^0 \neq 0$;
- $d_i^+ = |N_i^+|$ = number of outgoing edges of node i;
- a *walk* denotes a series of adjacent directed edges (including trivial walks consisting of one node with no edge);
- a *loop* is a walk whose terminal node coincides with the initial node;[4]
- a *path* is a walk that never passes twice through the same node, i.e. a walk without loops.
- a directed graph is *weakly connected* if, for any partition of its vertices in two sets, there is at least one edge starting in one of the sets and ending in the other one.
- a *tree* is a graph that is weakly connected and has no loops even if one were to change the edges directions.

We can now establish the following basic results.

Theorem 4.1: 1) If w_i is a source, then $G_{ij}^0 = 0 \; \forall j$, $T_{ii}^0 = 1$ and $T_{ij}^0 = 0 \; \forall j \neq i$. The measurement of a source does not add any linearly independent equation to the system of equations (10). The identification of a transfer function on an outgoing edge from a source i requires that an external signal r_i is applied at the source.

2) If w_i is a sink, then $G_{ji}^0 = 0 \; \forall j$, $T_{ii}^0 = 1$ and $T_{ji}^0 = 0 \; \forall j \neq i$. Identifiability of the network requires that all sinks be measured. The application of an external signal r_i at a sink i yields no information.

Proof: 1) The first part follows from the definition of a source and from the calculation of T^0 from such G^0 using (3). It then follows that if C selects a source, say w_i, then the corresponding equation of (3) yields $T_{ii}^0 = 1$, which does not contribute any information for the identifiability of G^0. Finally, let i be a source with an outgoing edge G_{ki}^0. It follows that $w_i = r_i + v_i$ and thus $w_k = G_{ki}^0(r_i + v_i) + x$, where x contains only terms that do not involve G_{ki}^0. Hence the identification of G_{ki}^0 requires that $r_i \neq 0$.

2) The first part follows from the definition of a sink and from the calculation of T^0 using (3). Let node i be a sink and let node k be connected to i by a nonzero transfer function G_{ik}^0. Since node i is a terminal node of the path from k to i, no

---

[4]Note that a loop is typically called *cycle* in graph theory.

node signal other than w_i can give any information about G_{ik}^0. On the other hand, applying an excitation signal r_i to sink i yields no information, since no path leaves node i. ∎

We now make some observations concerning the number of useful equations that result from (10) for the computation of the G_{ij}. Each measured node contributes L equations, but some of these may not yield any information, because they result in 1=1 or 0=0.

First we note that $L-1 \leq n \leq L(L-1)$, the first inequality being a consequence of the connectedness of the graph. The number of equations is $p \times L$, so it is obvious that we need $p \geq \frac{n}{L}$. It now follows from (10) and Theorem 4.1 that each sink causes the appearance of one trivial equation $1 = 1$ in the sink's measurement, and also of one trivial equation $0 = 0$ at every other measurement. Hence the number of trivial equations caused by each sink equals p, and thus the total number of trivial equations due to the existence of sinks is ps. Therefore the number of useful equations is at most $n_e = pL - ps = p(L - s)$. We then have the following result.

Theorem 4.2: Identifiability of the whole network requires measurement of all sinks plus at least m more nodes such that

$$m + s \geq \frac{n}{L - s} \tag{11}$$

Proof: Given that the number of useful equations resulting from p measurements is at most $p(L - s)$, identifiability of a network with n unknowns and s sinks requires that $p(L-s) \geq n$, where $p = m + s$. This implies (11). ∎

The next theorem yields a simple result for networks that have the structure of a tree.

Theorem 4.3: For a tree it is necessary and sufficient to measure all the sinks, assuming that none of the G_{ij} that make up the tree are zero.

Proof: By Theorem 4.1 it is necessary to measure all the sinks for any graph, so it remains to prove sufficiency. In a tree every sink will be the terminal node of a path. Given that all transfer functions T_{ji}^0 from any input r_i to any sink is identifiable, in order to determine all the G_{kl}^0 in that path one can proceed backwards from the sink up to the root, since the transfer function from any given r_i to the sink is just the product of the G_{kl}^0 of each edge in the path from r_i to the sink, none of which is zero by our assumption. ∎

After a result for networks having a tree structure, the next result covers the case of loops.

Theorem 4.4: Let the nodes w_i, $i \in \mathcal{I}$ form one loop and assume that no other loop in the graph contains any of these nodes. Suppose moreover that all the transfer functions involved in the loop are nonzero. Then measuring any one of these nodes is sufficient to identify all transfer functions in the loop.

Proof: Let η be the cardinality of $\mathcal{I}$ and consider, without loss of generality, that the nodes in the loop are labeled $i = 1, \ldots, \eta$ sequentially, that is there is a link from each node i to node $i + 1$, so that the η transfer functions to be identified in

the loop are $G^0_{i+1,i}$, $i = 1, \dots, \eta - 1$ and $G^0_{1,\eta}$. Since an external excitation signal is assumed to enter each node, input-output identification provides all closed-loop transfer functions $T^0_{i,j}, i, j \in \mathcal{I}$, none of which are zero. Indeed,

$$T^0_{i,j} = \frac{1}{\Delta} G_{i,i-1} G_{i-1,i-2} \dots G_{1,\eta} G_{\eta,\eta-1} \dots G_{j+1,j}, \; i < j$$
$$T^0_{i,i} = \frac{1}{\Delta}$$
$$T^0_{i,j} = \frac{1}{\Delta} G_{i,i-1} G_{i-1,i-2} \dots G_{j+1,j}, \; i > j$$

where

$$\Delta = 1 - G_{1,\eta} \Pi_{i=1,\dots,\eta-1} G_{i+1,i}. \tag{12}$$

Now, suppose we measure only the "last" node $i = \eta$. Then we have identified all the transfer functions $T^0_{\eta,j}$:

$$T^0_{\eta,j} = \frac{1}{\Delta} G_{\eta,\eta-1} G_{\eta-1,\eta-2} \dots G_{j+1,j}, \; j = 1, \dots, \eta-1$$
$$T^0_{\eta,\eta} = \frac{1}{\Delta} \tag{13}$$

Now, notice that

$$G_{j+1,j} = \frac{T^0_{\eta,j}}{T^0_{\eta,j+1}}, \; j = 1, \dots, \eta-1$$

which gives each one of the transfer functions in the path from node 1 to node η, that is all transfer functions in the loop except $G_{1,\eta}$. Then this last transfer function can be obtained, from (12) and (13), as

$$G_{1,\eta} = \frac{1}{\Pi_{i=1,\dots,\eta-1} G_{i+1,i}} \left(1 - \frac{1}{T^0_{\eta,\eta}} \right).$$

The same reasoning holds if we measure any other node, since it is just a question of relabeling the nodes. ∎

V. PATH-BASED RESULTS

In this section, we consider a specific node i within the network and its out-going edges, i.e. the edges corresponding to the nonzero elements G^0_{ji} within the network matrix. Recall that we denote by N_i^+ the corresponding set of out-neighbors of node i. We show that the generic identifiability of an edge[5] or a group of edges leaving this node i can be related to the structure of the paths from the corresponding out-neigbors to the measured nodes.

Section V-A presents a linear algebraic reformulation of the identifiability problem, which involves submatrices of T^0. In Section V-B we formally define the notion of generic identifiability, needed because of the risk of exceptional rank drops in the submatrices of T^0. Section V-C establishes the link between the structure of paths in the network and the generic rank of certain submatrices of T^0. These relations are then used in Section V-D to obtain necessary and sufficient conditions for identifiability of out-going edges of a specific node, and some corollaries are derived in Section V-E.

[5]For reasons of brevity, we shall in future often refer to the *identifiability of an edge*, where this in fact means the identifiability of the transfer function corresponding to this edge.

A. A linear algebraic reformulation

Remember that CT^0 can be perfectly identified from $\{y, r\}$ data, and that therefore the transfer function G^0_{ji} of an edge (i, j) is identifiable if (10) implies $G_{ji} = G^0_{ji}$ for any G *consistent with the graph, i.e. with the topology*. Define $\Delta \triangleq G - G^0$, which is consistent with the graph if and only if G is. The next Lemma shows how the identifiability of G^0_{ji} depends on the kernel of a submatrix of the known CT^0, and hence on the rank of certain submatrices of CT^0.

Lemma 5.1: Let N_i^+ be the set of out-neighbors of node i. Let T^0_{C,N_i^+} denote the restriction of T^0 to the rows selected by C and to the columns corresponding to N_i^+, and let $\Delta_{N_i^+,i}$ denote the restriction of the i-th column of Δ to the rows corresponding to N_i^+. Then G^0_{ji} is identifiable from CT^0 if and only if

$$T^0_{C,N_i^+} \Delta_{N_i^+,i} = 0 \Rightarrow \Delta_{ji} = 0. \tag{14}$$

Proof: Substituting $G = G^0 + \Delta$ in (10) and remembering $T^0(I - G^0) = I$ shows that G_{ji} is identifiable if and only if

$$CT^0 \Delta = 0 \Rightarrow \Delta_{ji} = 0 \tag{15}$$

for any Δ consistent with the graph. The left hand side of (15) actually consists of L independent linear systems of the form

$$CT^0 \Delta_{:\ell} = 0, \quad \ell = 1, \dots L.$$

The function Δ_{ji} only appears in one system, with $\ell = i$, and none of the functions appearing in that system appear in any other one. Hence G_{ji} is identifiable if and only if

$$CT^0 \Delta_{:i} = 0 \Rightarrow \Delta_{ji} = 0 \tag{16}$$

for any $\Delta_{:i}$ consistent with the graph i.e. $\Delta_{ki} = 0$ if there is no edge $(i, k) \in G^0$. Remember that G_{ki}, and hence Δ_{ki}, may be nonzero only if $k \in N_i^+$. We use the notation $l \in \mathcal{C}$ to say that l is a measured node. Condition (16) can be rewritten as

$$\sum_{k \in N_i^+} T^0_{lk} \Delta_{ki} = 0, \forall l \in \mathcal{C} \Rightarrow \Delta_{ji} = 0, \tag{17}$$

which is equivalent to (14). ∎

The identifiability of G^0_{ji} is thus related to the rank of T^0_{C,N_i^+} and, as will be seen, that of certain of its submatrices. We will see in Section V-C how these are related to the topology.

B. Generic properties

As seen in Example 1, identifiability essentially depends on the known graph associated to G^0, except for network matrices G^0 that lie in subsets of measure 0. We now formalize this notion, using an approach similar to that in [15]. A *rational transfer matrix parametrization consistent with a given graph* is defined in the following way. For every edge (j, i), set constants p_{ij}, $n_{ij} \in \mathbb{Z}_0^+$, and parametrize $G_{ij}(z)$ by

$$G_{ij}(z) = \kappa_{ij} z^{-p_{ij}} \frac{z^{n_{ij}} + \beta^{(n_{ij}-1)}_{ij} z^{n_{ij}-1} + \dots + \beta^{(1)}_{ij} z + \beta^{(0)}_{ij}}{z^{n_{ij}} + \alpha^{(n_{ij}-1)}_{ij} z^{n_{ij}-1} + \dots + \alpha^{(1)}_{ij} z + \alpha^{(0)}_{ij}} \tag{18}$$

for real parameters κ_{ij}, α_{ij}^m and β_{ij}^m, $(1 \leq m < n_{ij})$. For pairs (j, i) not connected by an edge, let $G_{ij}(z) = 0$. We collect all parameters κ_{ij}, α_{ij}^m, β_{ij}^m in a vector P, and denote by $G(P, z)$ the transfer matrix obtained by a specific parameter.

We say that a property *generically*[6] *holds for a network matrix* G^0 if, for any rational transfer matrix parametrization $G(P, z)$ consistent with the graph associated to G^0, the property holds for $G(P, z)$ for all parameters P except possibly those lying on a zero measure set in $\Re^N$, where N is the total number of parameters.

notational remark.

In the remainder of this paper, and in order to simplify notations, we will say that a property generically holds for $T^0 = (I - G^0)^{-1}$ if for every parametrization $G(P, z)$ consistent with the graph associated to G^0, the property holds for $T(P, z) := (I - G(P, z))^{-1}$ for all P except possibly those lying on a zero measure set. We will use the same convention for properties holding for submatrices of T^0.

We have seen in Section V-A that identifiability is linked to the rank of certain matrices. Hence generic identifiability will be linked to the *generic rank* of certain submatrices of T^0, i.e. the size of their largest generically nonsingular submatrix. This implies checking if the determinant of a matrix related to G is generically nonzero. The following Lemma, when applied to Q being the determinant of a matrix related to G, provides a convenient way of establishing this. See proof in Appendix A.

Lemma 5.2: Let $Q(.) : \mathbb{C}^{L \times L} \to \mathbb{C}$ be an analytic function and consider a network matrix $G^0(z)$. If there exists a matrix $A \in \mathbb{C}^{L \times L}$ consistent with the graph associated to $G^0(z)$ such that $Q(A) \neq 0$, then $Q(G^0(z))$ is generically not identically zero as a function of z (for polynomial or rational $Q(.)$, it then has finitely many roots). Otherwise, $Q(G(z)) \equiv 0$ for every $G(z)$ consistent with the graph. ∎

This leads to the following definition of a *generically identifiable network matrix*.

Definition 1: A network matrix $G^0(z)$ is generically identifiable from a set of measured nodes defined by C in (5) if, for any rational transfer matrix parametrization $G(P, z)$ consistent with the directed graph associated to $G^0(z)$, there holds

$$C(I - G(P, z))^{-1} = C(I - \tilde{G}(z))^{-1} \Rightarrow G(P, z) = \tilde{G}(z) \quad (19)$$

for all parameters P except possibly those lying on a zero measure set in $\Re^N$, where $\tilde{G}(z)$ is any network matrix consistent with the graph. ∎

This definition naturally extends to the generic identifiability of a specific transfer function in G (or edge), or of a group of these. We note that (19) is exactly parallel to the definition

---

Fig. 4. Illustration of the notions of vertex-disjoint paths, disconnecting sets, and Lemma 5.3. The highlighted paths $(2, 5, 8)$ and $(3, 6, 9)$ are vertex-disjoint paths from $\mathcal{A} = \{1, 2, 3\}$ to $\mathcal{C} = \{7, 8, 9\}$, and their existence implies $b_{\mathcal{A} \to \mathcal{C}} \geq 2$. An alternative set of such paths is $(1, 4, 5, 7)$ and $(3, 6, 9)$. On the other hand, every path starting from $\mathcal{A}$ and arriving in $\mathcal{C}$ goes through $\mathcal{B} = \{5, 6\}$, which is thus an $\mathcal{A} - \mathcal{C}$ disconnecting set. It follows then from Lemma 5.3 that $b_{\mathcal{A} \to \mathcal{C}} \leq 2$ and thus that (i) $\mathcal{B}$ is a minimal $\mathcal{A} - \mathcal{C}$ disconnecting set, and (ii) there is no set of more than two vertex-disjoint paths from $\mathcal{A}$ to $\mathcal{C}$.

of identifiability of a given transfer matrix, since (9) can be rewritten $C(I - G^0)^{-1} = C(I - G)^{-1}$.

C. Disconnecting sets, vertex-disjoint paths and matrix rank

We say that a group of paths are mutually vertex disjoint if no two paths of this group contain the same vertex. Consider two subsets of nodes $\mathcal{A}$ and $\mathcal{C}$. We let $b_{\mathcal{A} \to \mathcal{C}}$ be the maximum number of mutually vertex disjoint paths starting in $\mathcal{A}$ and ending in $\mathcal{C}$. We say that a set of nodes $\mathcal{B}$ is an $\mathcal{A} - \mathcal{C}$ *disconnecting set* if every path starting in $\mathcal{A}$ and ending in $\mathcal{C}$ contains at least one node in $\mathcal{B}$, which implies that there would be no path from $\mathcal{A}$ to $\mathcal{C}$ if $\mathcal{B}$ were removed. These notions are illustrated in Figure 4. Note that $\mathcal{B}$ can intersect $\mathcal{A}$ and/or $\mathcal{C}$. In particular, $\mathcal{A}$ and $\mathcal{C}$ are always $\mathcal{A} - \mathcal{C}$ disconnecting sets.

The following Lemma, also illustrated in Figure 4, links the notions of disconnecting sets and vertex disjoint paths.

Lemma 5.3: Consider two subsets of nodes $\mathcal{A}$ and $\mathcal{C}$. The maximum number $b_{\mathcal{A} \to \mathcal{C}}$ of mutually vertex disjoint paths from $\mathcal{A}$ to $\mathcal{C}$, is also the size of the smallest $\mathcal{A} - \mathcal{C}$ disconnecting set. Moreover, under the standing assumption that the network is weakly connected, it can be computed in $O(Ln)$ operations.

Proof: The equality between $b_{\mathcal{A} \to \mathcal{C}}$ and the size of the smallest $\mathcal{A} - \mathcal{C}$ disconnecting set is the directed vertex disjoint version of Menger's theorem, see e.g. [17]. Computing $b_{\mathcal{A} \to \mathcal{C}}$ can be recast as solving a max-flow problem, see for example Section 24.2 of [18]. There exist many efficient ways of solving max-flow problems. Since the maximum flow is bounded by L, the classical Ford-Fulkerson Algorithm (see e.g. [14, Section 10.5.1]), for example, terminates in $O(Ln)$ operations provided there are at least $L - 1$ edges, which is the case if the network is weakly connected. ∎

---

[6]The word "structurally" is also sometimes used, see e.g. [16].

The next lemma will be useful to use the bounds derived in terms of $b_{\mathcal{A} \to \mathcal{C}}$.

Lemma 5.4: For any sets of nodes $\mathcal{A}, \mathcal{A}'$ and $\mathcal{C}$ there holds

$$b_{\mathcal{A} \cup \mathcal{A}' \to \mathcal{C}} \leq b_{\mathcal{A} \to \mathcal{C}} + b_{\mathcal{A}' \to \mathcal{C}}$$

Proof: Consider a set of $b_{\mathcal{A} \cup \mathcal{A}' \to \mathcal{C}}$ vertex disjoint paths from $\mathcal{A} \cup \mathcal{A}'$ to $\mathcal{C}$. Let b be the number among those starting from a node in $\mathcal{A}$. There must thus be at least $b_{\mathcal{A} \cup \mathcal{A}' \to \mathcal{C}} - b$ starting from $\mathcal{A}'$ (there can be more). By definition, $b_{\mathcal{A} \to \mathcal{C}} \geq b$ because we have already found b vertex disjoint paths from $\mathcal{A}$ to $\mathcal{C}$. Similarly, $b_{\mathcal{A}' \to \mathcal{C}} \geq b_{\mathcal{A} \cup \mathcal{A}' \to \mathcal{C}} - b$. Hence there holds

$$b_{\mathcal{A} \to \mathcal{C}} + b_{\mathcal{A}' \to \mathcal{C}} \geq b + b_{\mathcal{A} \cup \mathcal{A}' \to \mathcal{C}} - b = b_{\mathcal{A} \cup \mathcal{A}' \to \mathcal{C}}.$$

∎

Our main result in this subsection is the establishment of the link between the generic rank of submatrices of T^0 and the number of vertex-disjoint paths between the sets corresponding to the selected columns and rows of T^0. A link between generic rank and vertex-disjoint paths was obtained in the pioneering paper [19], where the rank of the matrix $C(sI - A)^{-1}B$ of a system $\dot{x} = Ax + Bu$, $y = Cx$ was related with paths from inputs to outputs in a graph defined by A, B and C.

Our next Proposition differs from the main result of [19] in several ways. First the paths defined in the graph associated to the matrix $C(sI - A)^{-1}B$ in [19] are those of the whole network that connects the inputs of $C(sI - A)^{-1}B$ to its outputs, whereas we consider the paths connecting a subset of these inputs to a subset of the outputs. Secondly, the matrices A, B, C appearing in $C(sI - A)^{-1}B$ are real matrices, while we examine the generic rank of a submatrix of $(I - G(P,z))^{-1}$ where $G(P,z)$ is a matrix of transfer functions. Finally, the definition of the nodes in the graph associated to $C(sI - A)^{-1}B$ differs from that used in this paper, because our nodes are linked by transfer functions; this means that if we were to represent $T^0(z) = (I - G^0(z)))^{-1}$ as a state space representation as is done in [19], then the nodes of the graph associated to our $G(z)$ would be a small subset of those associated to this state space representation. For all these reasons, the next Proposition is not just an application of the main theorem of [19] and requires a specific proof.

Proposition 5.1: [7] Let $\mathcal{A}, \mathcal{C}$ be two sets of nodes of a directed graph associated to a network matrix $G^0(z)$. Let $T^0_{\mathcal{C},\mathcal{A}}(z)$ be the restriction of $T^0(z) = (I - G^0(z))^{-1}$ to the rows corresponding to $\mathcal{C}$ and columns corresponding to $\mathcal{A}$. Then the generic rank of $T^0_{\mathcal{C},\mathcal{A}}(z)$ is $b_{\mathcal{A} \to \mathcal{C}}$.

Proof: The proof will consist of two parts. The first one establishes that the rank is generically at least $b_{\mathcal{A} \to \mathcal{C}}$, and uses the interpretation of $b_{\mathcal{A} \to \mathcal{C}}$ in terms of the number of vertex-disjoint paths.

Part 1: Generically $\mathrm{Rank}(T^0_{C,A}) \geq b_{\mathcal{A} \to \mathcal{C}}$

---

[7]Remember the important notational convention adopted for T^0 and $T^0_{\mathcal{C},\mathcal{A}}(z)$.

Select $b_{\mathcal{A} \to \mathcal{C}}$ vertex-disjoint (directed) paths from $\mathcal{A}$ to $\mathcal{C}$, and let A be the adjacency matrix of the directed graph consisting only of these paths, i.e. $A_{ij} = 1$ if the edge (j,i) is on one of the paths and $A_{ij} = 0$ otherwise. It is then a standard result in graph-theory (see e.g. [14, Section 6.10]) that $[A^k]_{ji}$ is the number of walks of length exactly k from j to i in that graph. Since the graph consists of disjoint directed paths, this implies that (i) $A^k = 0$ if k is larger than the longest of the vertex-disjoint paths, and hence (ii) $(I - A)^{-1} = \sum_{k=0}^{\infty} A^k$. As a result (iii) $[(1 - A)^{-1}]_{ji} = \sum_{k=0}^{\infty}[A^k]_{ji}$ is the total number of walks of any length from i to j in the graph containing only the vertex disjoint paths. In particular, let now $\tilde{\mathcal{A}} \subseteq \mathcal{A}$ be the set of starting points of the paths, and $\tilde{\mathcal{C}} \subseteq \mathcal{C}$ the set of their arrival points, with obviously $|\tilde{\mathcal{A}}| = |\tilde{\mathcal{C}}| = b_{\mathcal{A} \to \mathcal{C}}$. Therefore if $i \in \tilde{\mathcal{A}}$ and $j \in \tilde{\mathcal{C}}$ and if they are on the same path, then $[(1 - A)^{-1}]_{ji} = 1$. Otherwise $[(1 - A)^{-1}]_{ji} = 0$ (as there is no walk from the origin of one path to the end of another one). The restriction $[(1 - A)^{-1}]_{\tilde{\mathcal{C}}, \tilde{\mathcal{A}}}$ of $[(1 - A)^{-1}]$ is thus a permutation matrix of size $b_{\mathcal{A} \to \mathcal{C}}$, whose determinant is nonzero. By Lemma 5.2 this implies that $det(T^0_{\tilde{\mathcal{C}}, \tilde{\mathcal{A}}})$ is generically nonzero, implying that the rank of $T^0_{\tilde{\mathcal{C}}, \tilde{\mathcal{A}}}$ is generically $b_{\mathcal{A} \to \mathcal{C}}$, and hence the generic rank of $T^0_{\mathcal{C}, \mathcal{A}}$ is at least $b_{\mathcal{A} \to \mathcal{C}}$, since $T^0_{\tilde{\mathcal{C}}, \tilde{\mathcal{A}}}$ is a submatrix of $T^0_{\mathcal{C}, \mathcal{A}}$.

The proof of the second part relies on the equivalent interpretation of $b_{\mathcal{A} \to \mathcal{C}}$ in terms of the size of the minimal $\mathcal{A} - \mathcal{C}$ disconnecting set.

Part 2: Generically $\mathrm{Rank}(T^0_{C,A}) \leq b_{\mathcal{A} \to \mathcal{C}}$

Let $\mathcal{B}$ be an $\mathcal{A} - \mathcal{C}$ disconnecting set of minimal size $b_{\mathcal{A} \to \mathcal{C}}$, the existence of which is guaranteed by Lemma 5.3. Let $\mathcal{S} \subset \{1, \ldots, L\}$ be the set of nodes that can be reached by a path from a node in $\mathcal{A}$ without intersecting any node of the disconnecting set $\mathcal{B}$, and let $\mathcal{P} = \{1, \ldots, L\} \setminus (\mathcal{S} \cup \mathcal{B})$. We have thus partitioned the L nodes into 3 disjoint sets: $\mathcal{P}, \mathcal{S}$ and $\mathcal{B}$. There holds $\mathcal{A} \subseteq \mathcal{S} \cup \mathcal{B}$ (nodes in $\mathcal{A}$ are all in $\mathcal{S}$ except if they belong to $\mathcal{B}$). There also holds $\mathcal{C} \subseteq \mathcal{P} \cup \mathcal{B}$. Indeed, there would otherwise be a node of $\mathcal{C}$ in $\mathcal{S}$, meaning that it could be reached from a node in $\mathcal{A}$ without going through $\mathcal{B}$, in contradiction with $\mathcal{B}$ being an $\mathcal{A} - \mathcal{C}$ disconnecting set.

After re-ordering of the indices, the matrices G^0 and T^0 can be rewritten as

$$G^0 = \begin{pmatrix} G_{PP} & G_{PB} & 0 \\ G_{BP} & G_{BB} & G_{BS} \\ G_{SP} & G_{SB} & G_{SS} \end{pmatrix} \text{ and} \qquad (20)$$

$$T^0 \triangleq (I - G^0)^{-1} = \begin{pmatrix} T_{PP} & T_{PB} & T_{PS} \\ T_{BP} & T_{BB} & T_{BS} \\ T_{SP} & T_{SB} & T_{SS} \end{pmatrix} \qquad (21)$$

We focus on the rows P and columns S and B, keeping in mind that $T^0 = I + G^0 T^0$. There holds

$$T_{PS} = I_{PS} + [GT]_{PS} = 0 + G_{PP}T_{PS} + G_{PB}T_{BS} + 0 T_{SS}$$
$$T_{PB} = I_{PB} + [GT]_{PB} = 0 + G_{PP}T_{PB} + G_{PB}T_{BB} + 0 T_{SB},$$

from which follows

$$\begin{pmatrix} I - G_{PP} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} T_{PB} & T_{PS} \\ T_{BB} & T_{BS} \end{pmatrix} = \begin{pmatrix} G_{PB} \\ I \end{pmatrix} \begin{pmatrix} T_{BB} & T_{BS} \end{pmatrix}.$$
$$(22)$$

The right hand side of the equality has a rank at most $b_{\mathcal{A}\to\mathcal{C}}$ because $(T_{BB} \quad T_{BS})$ has $b_{\mathcal{A}\to\mathcal{C}}$ rows. The same holds thus true for the left-hand side. Observe now that the left-hand side is square and generically invertible; it is indeed invertible if we replace G_{PP} by 0, and the generic invertibility then follows from Lemma 5.2. As a consequence, the rank of $T_{P\cup B, B\cup S}$, the second matrix of the left hand side, is also at most $b_{\mathcal{A}\to\mathcal{C}}$. The claim of part 2 follows then from the fact that $T^0_{C,A}$ is a submatrix of $T_{P\cup B, B\cup S}$, because we have seen that $\mathcal{A} \subseteq \mathcal{B} \cup \mathcal{S}$ and $\mathcal{C} \subseteq \mathcal{P} \cup \mathcal{B}$. ∎

The result of Proposition 5.1 can intuitively be understood as follows. In the system represented by T^0, an edge can carry a one-dimensional information about the effect of a given external excitation, and a vertex can only let a one-dimensional information about a given external excitation transit through it. Suppose first that the graph only consists of two paths starting in $\mathcal{A}$ and ending in $\mathcal{C}$. If the paths are vertex-disjoint, then nodes in $\mathcal{A}$ can transmit a two-dimensional information about a given external signal to those in $\mathcal{C}$, one dimension per path. On the other hand, if the two paths intersect in one vertex, only a one-dimensional information can transit through this vertex and reach $\mathcal{C}$. Proposition 5.1 extends this intuitive idea to graphs with more edges than just those on the paths and to larger number of paths. Since we know by Lemma 5.3 that the largest number of vertex-disjoint paths between two sets is the size of the smallest disconnecting set, this allows characterizing exactly the dimension of the information transmitted, i.e. the rank of $T^0_{C,A}$.

D. Necessary and sufficient conditions for generic identifiability

With the help of Proposition 5.1 we can now derive one of the main results of this paper, namely necessary and sufficient conditions for the generic identifiability of transfer functions leaving a given node i. The reformulation of the identifiability of a transfer function leaving node i by (14) naturally leads one to consider conditions for the generic identifiability of a group of edges leaving the same node i, as these are all related to the same matrix T^0_{C,N_i^+}.

Theorem 5.1: Let N_i^* be a subset of N_i^+ and denote $\bar{N}_i^* \triangleq N_i^+ \backslash N_i^*$. The transfer functions corresponding to edges from i to N_i^* can generically all be identified when measuring nodes $\mathcal{C}$ using the identified CT^0 if and only if the following two conditions hold:

$$b_{N_i^*\to\mathcal{C}} = |N_i^*| \tag{23}$$
$$b_{N_i^+\to\mathcal{C}} = b_{N_i^*\to\mathcal{C}} + b_{\bar{N}_i^*\to\mathcal{C}} = |N_i^*| + b_{\bar{N}_i^*\to\mathcal{C}} \tag{24}$$

Proof: Let us fix a G consistent with the graph defined by G^0 and the corresponding $T = (I - G)^{-1}$. It follows from (16) that we can recover the transfer functions of all edges from i to N_i^* if and only if the equality $CT\Delta_{:,i} = 0$ implies $\Delta_{N_i^*,i} = 0$ for every $\Delta_{:,i}$ for which $\Delta_{ki} = 0$ for every $k \notin N_i^+$. This can be rewritten as

$$T_{C,N_i^*}\Delta_{N_i^*,i} + T_{C,\bar{N}_i^*}\Delta_{\bar{N}_i^*,i} = 0 \to \Delta_{N_i^*,i} = 0 \tag{25}$$

Observe first that T_{C,N_i^*} must have rank $|N_i^*|$ for this condition to hold; otherwise one could find a $\Delta_{N_i^*,i} \neq 0$ for which $T_{C,N_i^*}\Delta_{N_i^*,i} = 0$, which, with $\Delta_{\bar{N}_i^*,i} = 0$, would contradict the condition. We can rewrite (25) as

$$T_{C,N_i^*}\Delta_{N_i^*,i} = -T_{C,\bar{N}_i^*}\Delta_{\bar{N}_i^*,i} \to \Delta_{N_i^*,i} = 0$$

If the image sets of T_{C,N_i^*} and $T_{C,\bar{N}_i^*}$ have a nontrivial intersection, then we could find $\tilde{\Delta}_{N_i^*,i} \neq 0$ and $\tilde{\Delta}_{\bar{N}_i^*,i} \neq 0$ such that $T_{C,N_i^*}\tilde{\Delta}_{N_i^*,i} = -T_{C,\bar{N}_i^*}\tilde{\Delta}_{\bar{N}_i^*,i} \neq 0$, and the condition is not satisfied. On the other hand, if the image sets of T_{C,N_i^*} and $T_{C,\bar{N}_i^*}$ have no nontrivial intersection, then $T_{C,N_i^*}\Delta_{N_i^*,i} = -T_{C,\bar{N}_i^*}\Delta_{\bar{N}_i^*,i}$ implies both $T_{C,N_i^*}\Delta_{N_i^*,i} = 0$ and $T_{C,\bar{N}_i^*}\Delta_{\bar{N}_i^*,i} = 0$. When T_{C,N_i^*} has rank $|N_i^*|$, the former equality implies $\Delta_{N_i^*,i} = 0$.

We have thus shown that the transfer functions of edges from i to N_i^* can all be identified from CT if and only if (i) $\text{Rank}(T_{C,N_i^*}) = |N_i^*|$ and (ii) the image sets of T_{C,N_i^*} and $T_{C,\bar{N}_i^*}$ have no nontrivial intersection, i.e. are linearly independent. The latter condition is equivalent to $\text{Rank}(T_{C,(N_i^*\cup\bar{N}_i^*)}) = \text{Rank}(T_{C,N_i^*}) + \text{Rank}(T_{C,\bar{N}_i^*})$. For any G (and corresponding T) consistent with the graph associated to G^0, this equality, together with $\text{Rank}(T_{C,N_i^*}) = |N_i^*|$, are thus necessary and sufficient for the identifiability of the transfer functions corresponding to the edges from i to $\bar{N}_i^*$. This is in particular the case for any matrix $G(P,z)$ in any parametrization of the transfer matrices consistent with that graph. Generic identifiability of the edges from i to $\bar{N}_i^*$ is thus equivalent to $\text{Rank}(T^0_{C,N_i^*}) = |N_i^*|$ and $\text{Rank}(T^0_{C,(N_i^*\cup\bar{N}_i^*)}) = \text{Rank}(T^0_{C,N_i^*}) + \text{Rank}(T^0_{C,\bar{N}_i^*})$ holding generically, and the equivalence with (23) and (24) then follows from Proposition 5.1. ∎

Comment: Condition (24) can also be formulated as

$$b_{N_i^+\to\mathcal{C}} \geq b_{N_i^*\to\mathcal{C}} + b_{\bar{N}_i^*\to\mathcal{C}} = |N_i^*| + b_{\bar{N}_i^*\to\mathcal{C}}, \tag{26}$$

because it follows from the sub-additivity Lemma 5.4 and $N_i^+ = N_i^* \cup \bar{N}_i^*$ that

$$b_{N_i^+\to\mathcal{C}} \leq b_{N_i^*\to\mathcal{C}} + b_{\bar{N}_i^*\to\mathcal{C}}.$$

This formulation will be used in the proof of Corollary 5.3.

An immediate corollary of Theorem 5.1 is obtained when one considers all out-neighbors of node i.

Corollary 5.1: The transfer functions from node i to its out-neighbors N_i^+ can generically all be identified from CT^0 if and only if $b_{N_i^+\to\mathcal{C}} = d_i^+ \triangleq |N_i^+|$.

Proof: The result follows from Theorem 5.1 applied to $N_i^* = N_i^+$, in which case $\bar{N}_i^*$ is an empty set. ∎

Corollary 5.1 can be intuitively understood in the following way. We want to recover d_i^+ transfer functions of edges leaving i, so we need a d_i^+-dimensional information about the effect of r_i. Moreover, the information we have comes from the out-neighbors of i and arrives at our measured nodes $\mathcal{C}$. Hence the recovery will be possible if and only if

a d_i^+-dimensional information is transmitted from these out-neighbors to $\mathcal{C}$, which requires $b_{N_i^+ \to \mathcal{C}} = d_i^+$ by Proposition 5.1. In case we only want to recover the transfer function of the edges arriving at a subset N_i^* of the out-neighbors of i as in Theorem 5.1, then the situation is more complex because the information received at $\mathcal{C}$ from N_i^* is mixed with information about other edges leaving i. One then has to check if the specific information about N_i^* can be isolated in all the information arriving at $\mathcal{C}$ from the out-neighbors of i, which is what condition (24) is about. It can indeed be interpreted as requiring all the loss of information-dimension from N_i^+ to $\mathcal{C}$ to concern exclusively information about $N_i^+ \setminus N_i^*$, leaving that about N_i^* intact.

Given the definition of $b_{\mathcal{A} \to \mathcal{C}}$, an alternative formulation of the previous result is as follows.

Corollary 5.2: The transfer functions from node i to its out-neighbors N_i^+ can generically all be identified from CT^0 if and only if there exist vertex-disjoint directed paths[8] leaving all out-neighbors of i and arriving at the measured nodes defined by C. ∎

Corollary 5.2 is illustrated by the example in Figure 5. Remember that known external signals r_i are applied to each node, which we have not added on the figure for visibility reasons. Node i has three outgoing nodes, each of which has a vertex-disjoint directed path to the measured nodes 7, 8 and 9, namely the paths $(1, 5, 7), (2, 4, 8)$ and $(3, 6, 9)$; they are represented by dashed green arrows. As a result, the dotted red transfer functions G_{1i}^0, G_{2i}^0 and G_{3i}^0 can all be identified from these three measured nodes.

Fig. 5. Example illustrating Corollary 5.2: 3 vertex-disjoint dashed green paths to the 3 measured nodes; the 3 dotted red edges are identifiable.

We stress that the sufficient condition in Corollary 5.2 does not require all paths from the nodes N_i^+ to the measured nodes to be disjoint, but only the existence of a set of mutually disjoint paths. In other words, there may very well exist many other paths than those used in the condition, and there is no requirement on those, nor on their intersections with those used in the condition. For example, Figure 5 illustrates that the conditions of Corollary 5.2 apply even though node 2 has another path to node 8, namely $(2, 6, 8)$ which has a common node with the path $(3, 6, 9)$.

[8]The vertex-disjoint condition applies also for the departure and arrival nodes.

Particularizing Theorem 5.1 to a set N_i^* consisting of a single node immediately leads to a necessary and sufficient condition for identifying a single transfer function.

Theorem 5.2: Consider an edge (i, j) and its corresponding transfer function G_{ji}^0, and let N_i^+ be the set of out-neighbors of i. The transfer function G_{ji}^0 can be generically uniquely identified by measuring the nodes $\mathcal{C}$ if and only if

$$b_{N_i^+ \to \mathcal{C}} = b_{N_i^+ \setminus \{j\} \to \mathcal{C}} + 1 \qquad (27)$$

Proof: The result follows directly from Theorem 5.1 applied to $N_i^* = \{j\}$, taking into account the fact that $b_{j \to \mathcal{C}}$ is 1 if there is a path from j to $\mathcal{C}$ and 0 otherwise. ∎

E. Additional results

In this Subsection we present several results that apply to specific cases and that can be directly derived from the previous results. We start by giving conditions for identifying a group of edges that are only sufficient (not necessary) but that are simpler than the ones given in Theorem 5.1.

Corollary 5.3: Consider a node i, and let $N_i^* \subseteq N_i^+$ be a subset of its out-neighbors with $|N_i^*| = d_i^*$. Suppose in addition that the two following conditions hold

(i) There exist d_i^* vertex disjoint directed paths joining the nodes of N_i^* to the measured nodes $\mathcal{C}$,

(ii) There is no path from any node of $N_i^+ \setminus N_i^*$ to any node of $\mathcal{C}$.

Then all transfer functions from node i to nodes in N_i^* can be generically identified from the measured nodes.

Proof: Since $|N_i^*| = d_i^*$, there holds $b_{N_i^* \to \mathcal{C}} \leq d_i^*$. Hence it follows from condition (i) that $b_{N_i^* \to \mathcal{C}} = d_i^*$. Condition (ii) implies that $b_{N_i^+ \setminus N_i^* \to \mathcal{C}} = 0$. Now $b_{N_i^+ \to \mathcal{C}}$ is by definition always larger than or equal to $b_{N_i^* \to \mathcal{C}}$ because $N_i^* \subseteq N_i^+$. It also follows from Lemma 5.4 that

$$b_{N_i^+ \to \mathcal{C}} \leq b_{N_i^* \to \mathcal{C}} + b_{N_i^+ \setminus N_i^* \to \mathcal{C}} = b_{N_i^* \to \mathcal{C}} \qquad (28)$$

where the last equality follows from $b_{N_i^+ \setminus N_i^* \to \mathcal{C}} = 0$. Combining this with $b_{N_i^+ \to \mathcal{C}} \geq b_{N_i^* \to \mathcal{C}}$ yields the desired result by Theorem 5.1. This implies that

$$b_{N_i^+ \to \mathcal{C}} \geq b_{N_i^* \to \mathcal{C}} = b_{N_i^* \to \mathcal{C}} + b_{N_i^+ \setminus N_i^+ \to \mathcal{C}}. \qquad (29)$$

The result then follows from Theorem 5.1 and Lemma 5.4. ∎

The next two results concern the generic identification of the whole network, starting with a rather simple but very telling necessary condition. It is related again to the need of obtaining information of sufficiently high dimension for the out-neighbors of every node.

Corollary 5.4: The network can be generically identified from CT^0 only if $\mathcal{C}$ contains at least as many nodes as the highest out-degree present in the network.

Proof: This follows from a direct application of the condition of Corollary 5.2 to a node i with the highest out-degree. ∎

Our next (and final) result in this subsection deals with the number of nodes that are necessary and sufficient for identification of the whole network: it shows that we never need to measure all L nodes to secure network identifiability. It also confirms that, without any knowledge of the topology, we need to measure at least all but one of the nodes, as we cannot exclude the possibility of the graph being fully connected. To prove it, we first need the following Lemma.

Lemma 5.5: Suppose we measure the set $\mathcal{C}_k$ defined as containing all nodes except k. If the network cannot be generically fully identified from $C_k T^0$, then there exists a node k' with $N_{k'}^+ \supseteq N_k^+ \cup \{k\}$, and thus $d_{k'}^+ > d_k^+$.

Proof: If a node i does not have k as out-neighbor, then all its outgoing edges are generically identifiable. Indeed, all out-neighbors belong to $\mathcal{C}_k$ and are thus all connected to $\mathcal{C}_k$ by trivial zero-length vertex disjoint paths. Suppose now that node i has k as out-neighbor. If k has an out-neighbor j that is not an out-neighbor of i, then there exists d_i^+ vertex disjoint paths from N_i^+ to $\mathcal{C}_k$: the path (k, j), and $d_i^+ - 1$ zero-length trivial paths from the other out-neighbors of i to themselves (since they belong to $\mathcal{C}_k$). Hence the network is generically fully identifiable. So if the network is not generically fully identifiable, then k and all its out-neighbors must be out-neighbors of i, which proves the claim with $k' = i$. ∎

Theorem 5.3: $p = L - 1$ is sufficient for identifiability, in the sense that there always exists a set of $L - 1$ measured nodes allowing to generically fully identify the network. In particular, measuring all nodes except one of those with the highest out-degree is always sufficient. In a fully connected network (that is, $n = L(L - 1)$), $p = L - 1$ is also necessary.

Proof: Necessity for the fully connected case is obvious, since to identify $n = L(L - 1)$ unknowns we need at least $p = \frac{n}{L} = L - 1$ measurements. To prove sufficiency, consider the set $\mathcal{C}_k$ defined in Lemma 5.5, where k is a node with the maximal out-degree. It then follows from Lemma 5.5 that this $\mathcal{C}_k$ allows generically full identification of the network, for otherwise there would be a node k' with a higher out-degree, which is a contradiction. ∎

In this Section, we have started our identifiability analysis by looking at a given node and its outgoing edges. We have given necessary and sufficient conditions for the identifiability of one specific outgoing edge, or a subset of outgoing edges, or all of them. These conditions are based on the existence of disjoint paths from these outgoing edges to the measured nodes. In particular, our results are useful to decide which nodes need to be measured if one wants to identify a particular transfer function: see Theorem 5.2. In addition, we have shown that it is never necessary to measure all L nodes of a network, but that $L - 1$ measures are sufficient. Several of our results and examples have actually shown that special structures within the network often allow one to identify the network using a

much smaller number of measurements than $L - 1$: see e.g. Theorem 4.4 and Example 5.

VI. MEASUREMENT-BASED RESULTS

In this section, instead of starting from a given node and its outgoing edges, we look at the converse approach. We consider a measured node, or a set of measured nodes, and we examine which transfer functions are identifiable from that measured node or from this set of measured nodes. In the first result we consider a single measured node.

Theorem 6.1: Let j be a measured node, and consider a node i that has a path to node j. Then all transfer functions along that path can generically be identified if there is no other walk that connects i to j.

Proof: Let N_i^* of Theorem 5.1 contain only the out-neighbor of node i that is on the path to j mentioned in the theorem, and let $\mathcal{C}$ contain only j. By the assumption in the statement, there is no path from any node in $N_i^+ \setminus N_i^*$ to j since this would constitute another path from i to j. The result then follows by applying Theorem 5.1 to the successive nodes along the path from i to j. ∎

Theorem 6.1 is illustrated by the example in Figure 6; remember again that known signals r_i are added to each node, which are not shown on the figure. It follows from this theorem that the 7 transfer functions on the dashed green-colored paths can all be generically identified from the measurement of node 9. If in addition node 7 is also measured, then the 10 transfer functions of the network can all be generically identified from the two measured nodes 7 and 9. The intuition behind Theorem 6.1 is that for each edge of the path we need to recover a specific one-dimensional information about the effect of the input at its starting node. The presence of the path from i to j guarantees that this one-dimensional information reaches j, and the absence of another walk to j guarantees that it is not mixed with other information about the same input, i.e. about other edges leaving the node.

Fig. 6. Example illustrating Theorem 6.1: all transfer functions on the dashed green edges can be generically identified from the measurement of node 9.

The following result extends Theorem 6.1 by providing a necessary and sufficient condition for the generic identifiability of all transfer functions on a path to a single measured node.

Theorem 6.2: Let j be the only measured node and consider a node i that has a path to node j - let's call it path $\mathcal{P}$. All

transfer functions along $\mathcal{P}$ can be (generically) identified if and only if any other walk from i to j contains $\mathcal{P}$ as a prefix.[9]

Proof: *Necessity:* Suppose there is a walk $\mathcal{W}$ from i to j that does not contain $\mathcal{P}$ as prefix. Since they both start from i, $\mathcal{P}$ and $\mathcal{W}$ begin by a common part, possibly reduced to node i without any edge. Let k be the last node of this initial common part, that is, $\mathcal{P}$ and $\mathcal{W}$ are identical until k and different afterwards. This last common node k cannot be j for otherwise $\mathcal{P}$ would be a prefix of $\mathcal{W}$. Hence there is a node after k along $\mathcal{P}$, which we call $\ell_{\mathcal{P}}$, and a node after k along $\mathcal{W}$, which we call we call $\ell_{\mathcal{W}}$. We apply Theorem 5.2 to the edge $(k, \ell_{\mathcal{P}})$. Clearly $b_{N_k^+ \to \mathcal{C}} = 1$ because $\mathcal{C} = \{j\}$ contains only one node. Moreover, there is by definition a walk from $\ell_{\mathcal{W}}$ to j and, since $\ell_{\mathcal{W}} \in N_k^+$, there is a path from $\ell_{\mathcal{W}} \in N_k^+ \setminus \{\ell_{\mathcal{P}}\}$ to j, so that $b_{N_k^+ \setminus \{\ell_{\mathcal{P}}\} \to \mathcal{C}} = 1$. It follows then from Theorem 5.2 that $(k, \ell_{\mathcal{P}})$ cannot be generically identified.

Sufficiency: Suppose now there exists a node $k \neq j$ on $\mathcal{P}$ and its successor $\ell_{\mathcal{P}}$ is such that the edge $(k, \ell_{\mathcal{P}})$ is not generically identifiable. Clearly, $b_{\ell_{\mathcal{P}} \to \mathcal{C}} = 1$. Hence it follows from Theorem 5.2 that $b_{N_k^+ \setminus \{\ell_{\mathcal{P}}\} \to \mathcal{C}} = 1$, which means there exists another neighbor, that we call $\ell_{\mathcal{W}}$, from which there is a path $\mathcal{P}'$ to $\mathcal{C} = \{j\}$. We can then build a walk $\mathcal{W}$ from i to j by aggregating (i) the restriction of $\mathcal{P}$ to its first nodes until it arrives at k, (ii) the edge $(k, \ell_{\mathcal{W}})$ and (iii) the path $\mathcal{P}'$ from $\ell_{\mathcal{W}}$ to j, and this walk does not contain $\mathcal{P}$ as a prefix. ∎

Finally, the results of Section V allow us to produce a necessary and sufficient condition for the generic identifiability of all edges of the network from a given set of measured nodes, i.e. a given choice of C. This is another main result of this paper.

Theorem 6.3: All edges of a network can generically be identified if and only if $b_{N_i^+ \to \mathcal{C}} = d_i^+$ for every i.

Proof: The result follows immediately from Corollary 5.1, or the equivalent Corollary 5.2, applied to all nodes. ∎

Theorem 6.3 can be put in other (more intuitive) words as follows: all edges can be identified if and only if for every node i there exist d_i^+ vertex-disjoint paths from the set of neighbors of i to the nodes of $\mathcal{C}$.

VII. Algorithmic Complexity

Our results allow determining whether a given set $\mathcal{C}$ of measured node allows recovering a specific edge (i, j), a specific set of edges, or all edges in the network. Let us now analyze the algorithmic complexity of these issues. We have seen in Lemma 5.3 that $b_{\mathcal{A} \to \mathcal{C}}$ can be computed in $O(Ln)$ for any sets $\mathcal{A}, \mathcal{C}$ using, for example, the Ford-Fulkerson algorithm.

It follows from Theorem 6.3 that checking if all edges can be identified can be achieved by computing $b_{N_i^+ \to \mathcal{C}}$ for the L nodes i, at a cost $L.O(Ln) = O(L^2 n)$. If we only want to determine if a specific edge (i, j) can be identified, then by Theorem 5.2 we can achieve this by comparing $b_{N_i^+ \to \mathcal{C}}$

with $b_{N_i^+ \setminus \{j\} \to \mathcal{C}}$, the computation of which has a cost $O(Ln)$. Finally, suppose we are given a $\mathcal{C}$ and we want to determine the exact set of edges that can be identified. We then need to compute $b_{N_i^+ \to \mathcal{C}}$ for each of the L nodes i and $b_{N_i^+ \setminus \{j\} \to \mathcal{C}}$ for each of the n edges (i, j), at a total cost of $(L + n)O(nL) = O((L + n)nL = O(n^2 L)$ if we assume that the network is weakly connected, so that $L \leq n + 1$.

VIII. Conclusions

The results so far on the global identifiability of a network of dynamical systems have been built on the assumption that all nodes are measured. In this paper, we have addressed the network identifiability problem in the situation where not all nodes are measured, but where they are all excited by a known external excitation signal. We have first shown that network identifiability with partial node measurements is impossible without knowledge about the topology. We have then developed an identifiability theory for a network matrix that is based on the topology of its associated graph, and not on the particular numbers that appear in the unknown network matrix. This has led us to define and exploit the notion of generic identifiability of a network matrix.

We have first shown that the node measurements needed for network identifiability depend entirely on the topology of the network. In doing so, we have observed that the measurement of all sinks are indispensable.

We have then provided a series of results on identifiability. Some of these are based on looking at a particular node and its out-neighbours, and their paths to measured nodes; others have addressed the question of which transfer functions can be identified from the measurement of a particular node or a subset of nodes.

Our first main result, based on the first approach, is a necessary and sufficient condition for identifiability of one edge, a set of edges, or all edges leaving a particular node. Our second main result is a necessary and sufficient condition for identifiability of all transfer functions of the network from a selected set of measured nodes. We have also shown that these necessary and sufficient conditions can be checked by algorithms that run in polynomial time, an important feature for large networks.

An interesting outcome of our work is that networks can often be identified by measuring only a small subset of nodes.

Future research questions will include the search for a reduced set of measured nodes that allow identification of the whole network, as well as the search for informative experiment designs.

References

[1] J. Gonçalves and S. Warnick, "Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks," *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1670–1674, 2008.

[2] D. Materassi and G. Innocenti, "Topological identification in networks of dynamical systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1860–1870, 2010.

[3] A. Dankers, P. Van den Hof, P. Heuberger, and X. Bombois, "Dynamic network structure identification with prediction error methods - basic examples," in *USB Proc. 16-th IFAC Symp on System Identification (SYSID 2012)*. IFAC, 2012, pp. 876–881.

---

[9]A path $\mathcal{P}_1$ is a prefix to another path $\mathcal{P}_2$ if the initial nodes of $\mathcal{P}_2$ are those of $\mathcal{P}_1$.

[4] A. Chiuso and G. Pillonetto, "A Bayesian approach to sparse dynamic network identification," *Automatica*, vol. 48, pp. 1553–1565, 2012.

[5] H. Weerts, A. Dankers, and P. Van den Hof, "Identifiability in dynamic network identification," in *USB Proc. 17th IFAC Symp. on System Identification*, Beijing, P.R. China, 2015, pp. 1409–1414.

[6] D. Hayden, Y. Chang, J. Goncalves, and C. Tomlin, "?Sparse network identifiability via compressed sensing," *Automatica*, vol. 68, pp. 9–17, 2016.

[7] M. Gevers, A. Bazanella, and A. Parraga, "On the identifiability of dynamical networks," in *USB Proc. IFAC World Congress 2017*. Toulouse, France: IFAC, July 2017, pp. 11 069–11 074.

[8] D. Materassi and M. Salapaka, "Identification of network components in presence of unobserved nodes," in *USB Proc. of 54th IEEE Conference on Decision and Control*, Osaka, Japan, December 2015, pp. 1563–1568.

[9] 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*, vol. 61, no. 4, pp. 937–952, April 2016.

[10] J. Linder and M. Enqvist, "Identification and prediction in dynamic networks with unobservable nodes," in *CD-ROM Proc. IFAC World Congress 2017*. Toulouse, France: IFAC, July 2017, pp. 11 063–11 068.

[11] A. Mauroy and J. Hendrickx, "Spectral identification of networks using sparse measurements," *SIAM Journal of Applied Dynamical Systems (SIADS)*, vol. 16, no. 1, pp. 479–513, 2016.

[12] M. Araki and M. Saeki, "A quantitative condition for the well-posedness of interconnected dynamical systems," *IEEE*, vol. 28, no. 5, pp. 625–637, May 1983.

[13] R. Diestel, *Graph theory*. New York, USA: Springer Verlag, 2000.

[14] M. Newman, *Networks: an introduction*. Oxford university press, 2010.

[15] J. van der Woude, "Disturbance decoupling by measurement feedback for structured transfer matrix systems," *Automatica*, vol. 32, no. 3, pp. 357–363, 1996.

[16] C.-T. Lin, "Structural controllability," *IEEE Transactions on Automatic Control*, vol. 19, no. 3, pp. 201–208, 1974.

[17] T. Böhme, F. Göring, and J. Harant, "Menger's theorem," *Journal of Graph Theory*, vol. 37, no. 1, pp. 35–36, 2001.

[18] J. Erickson, "Algorithms, Etc." 2015. [Online]. Available: http://jeffe.cs.illinois.edu/teaching/algorithms/

[19] J. van der Woude, "A graph-theoretic characterization for the rank of the transfer matrix of a structured system," *Mathematics of Control, Signals, and Systems (MCSS)*, vol. 4, no. 1, pp. 33–40, 1991.

[20] B. Mityagin, "The zero set of a real analytic function," *arXiv preprint arXiv:1512.07276*, 2015.

APPENDIX

Proof of Lemma 5.2

We first show that the absence of A such that $Q(A) \neq 0$ implies $Q(G(z)) \equiv 0$ for any G consistent with the graph associated to $G^0(z)$. Indeed, if there is a $G(z)$ such that $Q(G(z)) \not\equiv 0$, then there is a z^* such that $Q(G(z^*)) \neq 0$ and we obtain the desired A by taking $G(z^*)$.

We now show that the existence of A such that $Q(A) \neq 0$ implies $Q(G^0(z)) \neq 0$ generically. Consider a parametrization $G(P, z)$ of rational transfer functions consistent with the graph associated to G^0, and let $\tilde{Q} : (P, z) \to \tilde{Q}(P, z) = Q(G(P, z))$ as a function of both z and the parameters collected in P. Suppose, to obtain a contradiction, that the implication does not hold, that is there exists a nonzero-measure set $\mathbb{P}_0$ of parameters P such that $Q(G(P, z))$, as a function of z, is identically zero. This implies that $\tilde{Q}(P, z) = 0$ for every couple $(P, z) \in \mathbb{P}_0 \times \mathbb{C}$, a set whose measure is also nonzero. Now, it follows from the assumption on Q and the parametrization by rational functions that $\tilde{Q}$ is analytic. And it is a classical result that analytic functions that are not identically zero vanish only on a zero-measure set [20]. In particular, the fact that $\tilde{Q}$ vanishes on $\mathbb{P}_0 \times \mathbb{C}$ implies that it is identically 0.

We now show that this contradicts the existence of A consistent with the graph for which $Q(A) \neq 0$. Observe indeed that $A = G(P^*, 1)$ for a parametrization P^* defined by letting $k_{ij}^* = A_{ij}$, and $\beta_{ij}^{t*} = \alpha_{ij}^{t*} = 0$ for every i, j and t. Hence $\tilde{Q}(P^*, 1) = Q(G(P^*, 1)) = Q(A) \neq 0$, in contradiction with $\tilde{Q} \equiv 0$. Therefore, the existence of A implies that $Q(G(P, z))$ is identically zero (as a function of z) only on a zero-measure set of parameters P. The last part of the result follows from the fact that single-variable polynomials have a finite number of roots when they are not identically zero, and the same holds for rational functions. ∎

Julien M. Hendrickx Julien M. Hendrickx received an engineering degree in applied mathematics and a PhD in mathematical engineering from the Université catholique de Louvain, Belgium, in 2004 and 2008, respectively.

He has been a visiting researcher at the University of Illinois at Urbana Champaign in 2003-2004, at the National ICT Australia in 2005 and 2006, and at the Massachusetts Institute of Technology in 2006 and 2008. He was a postdoctoral fellow at the Laboratory for Information and Decision Systems of the Massachusetts Institute of Technology 2009 and 2010, holding post-doctoral fellowships of the F.R.S.-FNRS (Fund for Scientific Research) and of the Belgian American Education Foundation. Since September 2010, he is a faculty member of the Université catholique de Louvain, in the Ecole Polytechnique de Louvain.

Doctor Hendrickx is the recipient of the 2008 EECI award for the best PhD thesis in Europe in the field of Embedded and Networked Control Systems, and of the Alcatel-Lucent-Bell 2009 award for a PhD thesis on original new concepts or application in the domain of information or communication technologies.

Michel Gevers Michel Gevers obtained an Electrical Engineering degree from the University of Louvain, Belgium, in 1968, and a Ph.D. degree from Stanford University, California, in 1972, under the supervision of Tom Kailath. He is an IFAC Fellow, a Fellow of the IEEE, a Distinguished Member of the IEEE Control Systems Society. He holds Doctor Honoris Causa degrees from the University of Brussels and from Linköping University, Sweden. He has been President of the European Union Control Association (EUCA) from 1997 to 1999, and Vice President of the IEEE Control Systems Society in 2000 and 2001.

Michel Gevers is Professor Emeritus at the Department of Mathematical Engineering of the University of Louvain, Louvain la Neuve, and scientific collaborator at the Free University Brussels (VUB). He has been for 20 years the coordinator of the Belgian Interuniversity Network DYSCO (Dynamical Systems, Control, and Optimization) funded by the Federal Ministry of Science. He has spent long-term visits at the University of Newcastle, Australia, and the Technical University of Vienna, and was a Senior Research Fellow at the Australian National University from 1983 to 1986. His present research interests are in system identification, experiment design for identification of linear and nonlinear systems, and identifiability and informativity issues in networks of linear systems.

Michel Gevers has been Associate Editor of Automatica, of the IEEE Transactions on Automatic Control, of Mathematics of Control, Signals, and Systems (MCSS) and Associate Editor at Large of the European Journal of Control. He has published about 280 papers and conference papers, and two books: "Adaptive Optimal Control - The Thinking Man's GPC", by R.R. Bitmead, M. Gevers and V. Wertz (Prentice Hall, 1990), and "Parametrizations in Control, Estimation and Filtering Problems: Accuracy Aspects", by M. Gevers and G. Li (Springer-Verlag, 1993).

Alexandre Bazanella Alex Bazanella received his PhD degree in Electrical Engineering in 1997. He is currently a Full Professor with the Department of Automation and Energy of Universidade Federal do Rio Grande do Sul, in Porto Alegre, Brazil. His main research interests are presently in system identification and data-driven control design, but he has also authored a number of papers in nonlinear systems theory, particularly its application to electric power systems and electrical machines. He is the author of two books: "Control Systems - Fundamentals and Design Methods" (in portuguese), and "Data-driven Controller Design: the H2 Approach" (Springer, 2011). Dr. Bazanella has served as associate editor of the IEEE Transactions on Control Systems Technology from 2002 to 2008 and as an Editor of the Journal of Control, Automation and Electrical Systems from 2008 to 2012. He has held visiting professor appointments at Universidade Federal da Paraba in 2001 and at Université catholique de Louvain, where he spent a sabbatical year in 2006. Dr. Bazanella is a Senior Member of the IEEE, and is currently vice-president of the Brazilian Automation Society and vice-coordinator of the area of Electrical Engineering for CAPES, with the Brazilian Ministry of Education.