Preprints of the 17th IFAC Symposium on System Identification
Beijing International Convention Center
October 19-21, 2015. Beijing, China

Identifiability in dynamic network identification

Harm H.M. Weerts [*], **Arne G. Dankers** [**] and **Paul M.J. Van den Hof** [*]

[*] *Control Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands (email: h.h.m.weerts@tue.nl, p.m.j.vandenhof@tue.nl)*
[**] *Department of Electrical Engineering, University of Calgary, Canada (email: adankers@hifieng.com)*

Abstract: Dynamic networks are structured interconnections of dynamical systems driven by external excitation and disturbance signals. We develop the notion of *network identifiability*, a property of a parameterized model set that ensures that module dynamics are uniquely related to the filters that specify the one-step-ahead predictors of all node signals in the network. It can be used to specify which presence of excitation signals will result in a unique representation of the network dynamics in a particular network model parametrization. This uniqueness is necessary for detecting the topology of the network from measured data, and for consistently estimating the network dynamics. We combine aspects of the classical notion of system identifiability with a uniqueness-oriented parametrization concept, and extend this to the situation of highly structured model sets. All node signals in the network are treated in a symmetric way. The presented concept and theory allow for the incorporation of particular structural prior knowledge of the network structure.

Keywords: System identification, dynamic networks, identifiability.

1. INTRODUCTION

With the increasing complexity of dynamical systems that appear in engineering and other domains of science, the study of interconnected systems or dynamic networks has become an important area of attention in different fields, such as robotics, smart grids, social networks, transportation systems etcetera, and has led to interesting problems of distributed control and optimization. In system identification literature, where the majority of the work is focused on open-loop or feedback controlled systems, there is also an increasing interest in data-driven modeling problems related to dynamic networks. One particularly interesting area is that of topology detection, which has been addressed in e.g. [Materassi and Salapaka, 2012] where Wiener filters are used to reconstruct the network topology. In [Chiuso and Pillonetto, 2012] a Bayesian viewpoint is taken and regularization techniques are applied to obtain sparse estimates. Topology detection in a large scale network is done in [Sanandaji et al., 2011] using compressive sensing methods, and in a biological network in [Yuan et al., 2011] using also sparse estimation techniques. The problem of identifying a single module in a network has been addressed in [Van den Hof et al., 2013], where a framework has been introduced for prediction error identification in dynamic networks. Using this framework predictor input selection [Dankers et al., 2016] and errors-in-variables problems have been addressed [Dankers et al., 2014].

When identifying a full network model in order to extract the network topology, care has to be taken that there exists a unique mapping from the identified object to a dynamic network structure. In [Gonçalves and Warnick, 2008, Adebayo et al., 2012] specific local conditions for such a unique mapping are considered for a particular deterministic transfer function. Uniqueness properties of a model structure for purely stochastic networks are studied in [Materassi and Salapaka, 2012] where the assumption is made that each node is driven by an independent white noise source. In this paper we are going to address this problem in the context of prediction error identification and by utilizing the concept of identifiability. This allows to consider not only the deterministic mapping between measured signals, but to also include disturbance modelling, while allowing correlated disturbances on the different node signals.

The notion of identifiability is a classical notion in system identification, but the concept has been used in different settings. The classical definition as present in [Ljung, 1976] is a consistency-oriented concept concerned with estimates converging to the true underlying system (system identifiability) or to the true underlying parameters (parameter identifiability). These concepts were applied to the closed-loop case as well, see e.g. [Söderström et al., 1976]. A second and more dominant approach in the current literature is to consider identifiability as a property of a parametrized model set, referring to a unique one-to-one relationship between parameters and the predictor model, see e.g. [Ljung, 1999]. For an interesting account of these concepts see also the more recent work of [Bazanella et al., 2010].

In these classical concepts the structure/topology of the considered situation has been fixed and restricted to the common open-loop or closed-loop cases. When identifying models in dynamic networks we have to deal with additional structural properties in both the system setup and in our parametrized models. These properties concern e.g. the choices where external excitation and disturbance signals are present, and how they are modeled, whether or not disturbances can be correlated, and if some modules in the network are known and need not be parametrized.

Employing the dynamic network and prediction error identification framework as described in [Van den Hof et al., 2013], we introduce the notion of *network identifiability* which is concerned with the question whether there is a unique one-to-one relation between our network model dynamics and the one-step-ahead predictor model that is used to identify the network. This uniqueness property has been relatively easy to show in the classical open-loop and closed-loop settings, but becomes nontrivial in a dynamic network setting. Network identifiability is not consistency-driven, nor related to a unique parameter mapping, and as such it is not a simple generalization of those classical notions. In spirit it relates to the notion of discriminability of model sets as introduced in a deterministic setting in [Van den Hof, 1989, 1994].

The framework that we consider is very much flexible, and we do not make the (common) prior assumption that all disturbance signal are independent or white. All node variables are predicted based on other node variables and external variables, and in this sense are treated in a fully symmetric way. Our objective is to find conditions on the presence and location of external excitation signals and disturbance signals, as well as conditions on the parametrized model set, that allow a unique representation of the full network, including the topology. These conditions are necessary for consistency of the estimates.

This paper will proceed by defining the network setup (section 2), and subsequently formulating the network predictor model that is used as a basis for identification (section 3). In section 4 the concept of network identifiability will be introduced, motivated and illustrated with some examples. In section 5 results for the case where the interconnection structure in a network is (partially) known are given, after which conclusions are formulated.

2. DYNAMIC NETWORK SETTING

Following the basic setup of [Van den Hof et al., 2013], a dynamic network is built up out of L scalar *internal variables* or *nodes* w_j, $j = 1, \ldots, L$. Each internal variable is defined by:

$$w_j(t) = \sum_{l \in \mathcal{N}_j} G^0_{jl}(q) w_l(t) + \sum_{k \in \mathcal{N}_j^r} R^0_{jk}(q) r_k(t) + v_j(t) \quad (1)$$

where q^{-1} is the delay operator, i.e. $q^{-1}w_j(t) = w_j(t-1)$;

- G^0_{jl}, $l \in \mathcal{N}_j$ and R^0_{jk}, $k \in \mathcal{N}_j^r$ are proper rational transfer functions;
- v_j is *process noise*, modeled as a realization of a stationary stochastic process with rational spectral density;
- r_k is an *external variable* that can directly or indirectly be manipulated by the user;
- $\mathcal{N}_j$ is the set of indices of internal variables with direct causal connections to w_j, i.e. $l \in \mathcal{N}_j$ iff $G^0_{jl} \neq 0$;
- $\mathcal{N}_j^r$ is the set of indices of external variables with direct causal connections to w_j, i.e. $k \in \mathcal{N}_j^r$ iff $R^0_{jk} \neq 0$.

The situation that we would like to consider is the full network constructed by combining (1) for all node signals,

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} 0 & G^0_{12} & \cdots & G^0_{1L} \\ G^0_{21} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & G^0_{L-1\ L} \\ G^0_{L1} & \cdots & G^0_{L\ L-1} & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} + R^0(q) \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_K \end{bmatrix} + H^0(q) \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_L \end{bmatrix}$$

where $K \in \mathbb{N}_0$, and transfer function matrix H^0 is monic, stable and stably invertible, and e a multivariate white noise process with diagonal covariance matrix $\Gamma > 0$ describing the process noise via $v_j = \sum_{i=1}^{L} H^0_{ji}(q) e_i(t)$. Using obvious notation this results in the matrix equation:

$$w = G^0 w + R^0 r + H^0 e. \quad (2)$$

For a dynamic network as defined above, the resulting identification problem then becomes to identify the topology and/or the network dynamics (G^0, R^0, H^0) on the basis of measured node variables $\{w_j, j = 1, \cdots L\}$ and external variables $\{r_k, k = 1, \cdots K\}$ with potentially unknown sets $\mathcal{N}_j$ and $\mathcal{N}_j^r$.
The dynamic network formulation above relates that what has been called the *Dynamic Structure Function (DSF)* as considered for disturbance-free systems in [Adebayo et al., 2012].

3. NETWORK PREDICTOR

The one-step-ahead predictor for the network model (2) can be derived in a similar way as in the classical prediction error case (Ljung [1999]). To this end we rewrite the network model as

$$w = G^0 w + R^0 r + (H^0 - I)e + e$$

and substitute the expression for e from (2):

$$e = (H^0)^{-1}[(I - G^0)w - R^0 r]$$

into the expression $(H^0 - I)e$, leading to

$$w = [I - (H^0)^{-1}(I - G^0)]w + (H^0)^{-1} R^0 r + e.$$

We can now define the one-step-ahead predictor as:

$$\hat{w}_j(t|t-1) := \mathbb{E}\{w_j(t) | w_j(t-1)^-; w_k(t)^-, k \neq j; r(t)^-\}$$

where $w_j(t-1)^-$ reflects the past of $w_j(t)$. The predictor can then simply be shown to be given by

$$\hat{w}(t|t-1) = [I - (H^0)^{-1}(I - G^0)]w + (H^0)^{-1} R^0 r. \quad (3)$$

Note that in this derivation of the predictor, the different node variables (signals) are treated completely symmetric. All node variables are considered as output variables in the predictor, but at the same time they also serve as inputs for other node variables. Compared to the usual multivariate predictor [Ljung, 1999] in (3) the term $(H^0)^{-1}G^0 w$ is added such that relations between "outputs" are explicitly added instead of implicitly via $(H^0)^{-1}$. A network model structure is now defined as follows:

Definition 1. (network model structure). A network model structure for a network of L nodes and K external signals is defined as a set of parametrized transfer functions:

$$\mathcal{M}(\theta) := \{G(q,\theta), H(q,\theta), R(q,\theta), \theta \in \Theta\};$$

- $G(\theta) \in \mathbb{R}^{L\times L}(z)$, with zeros on the diagonal, all entries proper, and $(I - G(q,\theta))^{-1}$ proper and stable;
- $H(\theta) \in \mathbb{R}^{L\times L}(z)$, monic, stable and stably invertible;
- $R(\theta) \in \mathbb{R}^{L\times K}(z)$, all elements proper;
- the model structure defines a well-posed network [Dankers, 2014];
- define $U(q,\theta) := [H(q,\theta)\ R(q,\theta)]$ for notational convenience.

The requirement that $(I - G(q,\theta))^{-1}$ is proper and stable is induced by the requirement that all mappings from external variables to nodes should be stable and causal. A network model structure determines a one-step-ahead predictor for the internal variables $w(t)$ according to (3):

$$\hat{w}(t|t-1;\theta) = [I - H^{-1}(\theta)(I - G(\theta))]w + H^{-1}(\theta)R(\theta)r \quad (4)$$

which leads to a prediction error $\varepsilon(t,\theta) = w(t) - \hat{w}(t|t-1;\theta)$ that is given by

$$\varepsilon(t;\theta) = H(q,\theta)^{-1}[(I - G(q,\theta))w(t) - R(q,\theta)r(t)]. \quad (5)$$

The predictor model is determined by three transfer functions, namely $G(q,\theta)$, $H(q,\theta)$ and $R(q,\theta)$. However in the expression for the predictor itself only two transfer function matrices play a role, namely

$$[I - H^{-1}(\theta)(I - G(\theta))] \quad \text{and} \quad H^{-1}(\theta)R(\theta).$$

Since the analysis of these transfer function matrices is rather complicated, we transform them to an equivalent set of transfer function matrices that subsequently allow us to study aspects of identifiability.

Proposition 1. Consider a network model structure $\mathcal{M}$. Then two models represented by respectively θ_1, θ_2 induce the same predictor if and only if

$$T(q,\theta_1) = T(q,\theta_2), \quad (6)$$

$$\text{where } T(q,\theta) := [I - G(q,\theta)]^{-1}U(q,\theta) \quad (7)$$

Proof of the proposition is included in a report version of the paper [Weerts et al., 2015]. Due to the equivalence of the predictor filters and $T(q,\theta)$, the latter can be used to show properties of the former. In an identification context we can expect that the predictor filters (or equivalently $T(q,\theta)$) can be identified from data. The question that will be addressed in the next section is then under which conditions, knowledge of T leads to a unique representation of a model in the model class. Note that whereas in the predictor model the transfer functions from w, r to $\hat{w}$ are considered, in the above Proposition they are equivalently replaced by the transfer functions from ε, r to w.

4. NETWORK IDENTIFIABILITY

For introducing the notion of network identifiability we now move to the question whether there exists a unique relation between the predictor filters, equivalently represented by the transfer function $T(q,\theta)$ (7) and the transfer functions: $G(q,\theta), H(q,\theta)$ and $R(q,\theta)$. Different interconnection structures and corresponding G, H and R can result in the same mapping $r, e \to w$, i.e. T. This is different from classical identifiability as there the question is about a unique mapping from parameters to module dynamics. Note that in the open-loop situation the unique relation between predictor filters and module dynamics is trivial (Ljung [1999]). However in the dynamic network case this becomes dependent on several structural properties of the models.

Definition 2. (Global network identifiability at θ_0). A network model structure $\mathcal{M}$ is *globally network identifiable at* θ_0 if for all $\theta \in \Theta$ the following implication holds:

$$T(q,\theta) = T(q,\theta_0) \Rightarrow \begin{cases} G(q,\theta) = G(q,\theta_0) \\ U(q,\theta) = U(q,\theta_0). \end{cases}$$

Note that in contrast with the global identifiability concept as used in Ljung [1999], uniqueness of the module dynamics is addressed rather than uniqueness of the parameter values. A related concept has been exploited in [Gonçalves and Warnick, 2008] in the context of topology reconstruction on the basis of an identified network model, however outside the scope of prediction error identification.

In general we will be interested in uniqueness properties over the whole model set rather than only locally in a particular point θ_0, especially since the original network is supposed to be unknown. Therefore we have the following definition.

Definition 3. (Global network identifiability). A network model structure $\mathcal{M}$ is *globally network identifiable* if for all $\theta_1, \theta_2 \in \Theta$ the following implication holds:

$$T(q,\theta_1) = T(q,\theta_2) \Rightarrow \begin{cases} G(q,\theta_1) = G(q,\theta_2) \\ U(q,\theta_1) = U(q,\theta_2). \end{cases}$$

The following example uses a disturbance free network to illustrate global network identifiability at θ_0.

Example 1. Given the disturbance-free systems $\mathcal{S}_1$, $\mathcal{S}_2$ in Figure 1 with $A(q) \neq 0, -1$, and $B(q) \neq 0$, both rational transfer functions. Consider the model structure $\mathcal{M}(\theta)$ with (omitting arguments q, θ)

$$G = \begin{bmatrix} 0 & G_{12} & G_{13} \\ G_{21} & 0 & G_{23} \\ G_{31} & G_{32} & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (8)$$

Since we have a disturbance free system we discard a noise model here, without loss of generality.

The transfer function matrices $T_1^0(q)$ and $T_2^0(q)$ related to the networks $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively, are given by:

$$T_1^0(q) = \begin{bmatrix} 1 & 0 \\ A & 1 \\ AB+1 & B \end{bmatrix}, \quad T_2^0(q) = \begin{bmatrix} 1 & 0 \\ (A+1)B & 1 \\ A+1 & 0 \end{bmatrix}, \quad (9)$$

with

$$G_1^0 = \begin{bmatrix} 0 & 0 & 0 \\ A & 0 & 0 \\ 0 & B & 0 \end{bmatrix}, \quad G_2^0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & B \\ A & 0 & 0 \end{bmatrix}. \quad (10)$$

These transfer functions map the external signals r to the node signals w. In order to investigate whether each of the two systems can be represented uniquely within the model structure, we refer to (7), and analyze whether the equation

$$T_i^0(q) = [I - G(q,\theta)]^{-1}R(q,\theta). \quad (11)$$

for $i = 1, 2$ has a unique solution for $G(q,\theta)$. To this end we premultiply (11) with $[I - G(q,\theta)]$.

For network $\mathcal{S}_1$ we then obtain the relation (omitting argument q)

$$\begin{bmatrix} 1 & -G_{12}(\theta) & -G_{13}(\theta) \\ -G_{21}(\theta) & 1 & -G_{23}(\theta) \\ -G_{31}(\theta) & -G_{32}(\theta) & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ A & 1 \\ AB+1 & B \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (12)$$

Solving the corresponding six equations for the parameterized transfer functions $G_{ij}(\theta)$ shows that, from the first row in the matrix it follows that $G_{13}(\theta) = G_{12}(\theta) = 0$.

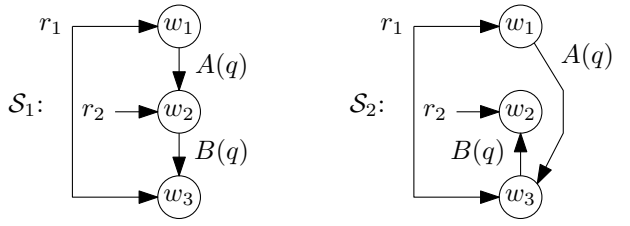

Fig. 1. Systems $\mathcal{S}_1$ and $\mathcal{S}_2$

Solving the second row leads to $G_{23}(\theta) = 0$ and $G_{21}(\theta) = A$, while solving the third row delivers $G_{13}(\theta) = 0$ and $G_{32}(\theta) = B$. As a result the original system $\mathcal{S}_1$ is uniquely recovered.
When applying the same reasoning to network $\mathcal{S}_2$ we obtain

$$\begin{bmatrix} 1 & -G_{12}(\theta) & -G_{13}(\theta) \\ -G_{21}(\theta) & 1 & -G_{23}(\theta) \\ -G_{31}(\theta) & -G_{32}(\theta) & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ (A+1)B & 1 \\ A+1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (13)$$

Solving this system of equations for the second column on the right hand side leads to $G_{12}(\theta) = G_{32}(\theta) = 0$, while the solution for the first column delivers $G_{13}(\theta) = 0$, $G_{31}(\theta) = A$ and

$$-G_{21}(\theta) + (A+1)B - G_{23}(\theta)(A+1) = 0 \quad (14)$$

or equivalently $G_{21}(\theta) = (A+1)(B - G_{23}(\theta))$. This shows that not only $G_{21}(\theta) = 0$, $G_{23}(\theta) = B$ is a valid solution, but actually an infinite number of solutions exists. As a result the chosen model structure is global network identifiable at the θ_0 that corresponds to $\mathcal{S}_1$, but it is not globally network identifiable at the θ_0 that corresponds to $\mathcal{S}_2$. An interpretation is that in $\mathcal{S}_2$ the contributions from w_1 and w_3 both solely depend on r_1 making them indistinguishable, which is reflected in the modeled transfer function matrix $R(q, \theta)$. □

Remark 1. The dynamic networks in Example 1 can be extended with disturbance signals. In the case that the vector of node signals in $\mathcal{S}_1$ and $\mathcal{S}_2$ are contaminated by a three-dimensional vector noise signal with dynamic correlation among the three disturbances, then this correlation will have to be parametrized in a square noise model $H(q, \theta)$. Without any further structural constraints on $H(q, \theta)$ the results on global network identifiability at θ_0 will remain the same as above. If the three node signals are contaminated by uncorrelated noise sources, the noise model $H(q, \theta)$ can be chosen diagonal, and the network model structure will become globally network identifiable, due to the independent noise excitation of each node. □

Next we consider the general case that there can be both disturbance signals and external variables present. For this situation the following result for global network identifiability can be shown.

Theorem 1. A network model structure $\mathcal{M}$ is globally network identifiable if there exists a nonsingular transfer function matrix $P \in \mathbb{R}^{K+L \times K+L}(z)$ such that

$$U(q, \theta)P(q) = [D(q, \theta) \ F(q, \theta)]$$

with $D(\theta) \in \mathbb{R}^{L \times L}(z)$, diagonal and full rank for all $\theta \in \Theta$, and $F(\theta) \in \mathbb{R}^{L \times K}(z)$.
In the case that the off-diagonal terms of $G(q, \theta)$ are fully parametrized and all transfer functions in $\mathcal{M}(\theta)$ are parametrized independently [1] the condition is also necessary. □

The proof of the theorem is collected in the appendix. Note that the condition can be interpreted as the possibility to give $U(q, \theta)$ a leading diagonal matrix by column operations. Essentially the theorem states that in case G is fully and independently parametrized each node should have some excitation source which has a component that is independent from excitation sources at the other node signals. The excitation sources can be either disturbances or external variables. There is an implicit requirement in the theorem that U has at most $K+1$ parametrized transfer functions on each row due to the transformation to diagonal form by matrix $P(q)$. Transforming a fully and independently parametrized full rank $U(q, \theta)$ to diagonal form by matrix P would require P to depend on θ.
In [Gonçalves and Warnick, 2008] results are presented in a deterministic setting that are equivalent to Theorem 1.

Example 2. Suppose we model correlated noise by having off-diagonal terms in H, in the model structure $\mathcal{M}(\theta)$ with

$$G = \begin{bmatrix} 0 & g_{12} & g_{13} \\ g_{21} & 0 & g_{23} \\ g_{31} & g_{32} & 0 \end{bmatrix}, \ H = \begin{bmatrix} h_{11} & h_{12} & 0 \\ h_{21} & h_{22} & 0 \\ 0 & 0 & h_{33} \end{bmatrix}, \ R = \begin{bmatrix} r_{11} & 0 \\ 0 & r_{22} \\ 0 & 0 \end{bmatrix},$$

where all nonzero elements are parameterized transfer functions, and $r_{11} \neq 0$, $r_{22} \neq 0$, and h_{ii} monic. Then a simple permutation matrix P can be found to create $U(q, \theta)P = [D(q, \theta) \ F(q, \theta)]$ with $D(q, \theta) = \text{diag}([r_{11} \ r_{22} \ h_{33}])$ and by Theorem 1 the model structure is globally network identifiable. □

In the next section structure restrictions in G are considered that allow more than $K+1$ parametrized transfer functions on the rows of U.

5. NETWORK IDENTIFIABILITY WITH STRUCTURE RESTRICTIONS IN G

If there is prior knowledge on particular parts of the network, e.g. knowledge that some interconnections are 0, or some modules (e.g. controllers) have known dynamics, the parametrization of $G(q, \theta)$ can be restricted which has its impact on the conditions for global identifiability.

Example 3. Consider the situation of network $\mathcal{S}_2$ in Example 1 with the additional prior knowledge that $G_{21} = 0$. Then we can fix $G_{21}(\theta) = 0$ and as a result $G_{23}(\theta) = B$. For the considered (restricted) model structure global network identifiability now holds true in the parameter value that corresponds to $\mathcal{S}_2$. □

As illustrated in Example 3 when information about the 'true' network is used then one can obtain results that allow us to distinguish between certain networks, however we are mainly interested in results that allow us to distinguish between all networks in a model structure. Below we present a theorem that shows when a model structure is network identifiable when we allow structural restrictions on $G(q, \theta)$. These structural restrictions are represented in parts of $G(\theta)$ that not not be parametrized.

Theorem 2. The network model structure $\mathcal{M}(\theta)$ is globally network identifiable if there exists a permutation matrix

[1] i.e. parameters used in one transfer function entry are different from parameters used in any other transfer function entry.

$P_r \in \mathbb{R}^{L\times L}$ and a nonsingular transfer function matrix $P \in \mathbb{R}^{(K+L)\times(K+L)}(z)$ such that for all $\theta \in \Theta$:

a) $\tilde{G}(\theta) := P_r G(\theta) P_r^{-1}$ is structured according to $\tilde{G}(\theta) = \begin{bmatrix}\tilde{G}_l(\theta) & \tilde{G}_r\end{bmatrix}$, with $\tilde{G}_l(\theta) \in \mathbb{R}^{L\times(L-M)}(z)$, and $\tilde{G}_r \in \mathbb{R}^{L\times M}(z)$, with M is the number of nonparameterized columns, and

b) $\tilde{U}(\theta) := P_r U(\theta) P$ is structured according to

$$\tilde{U}(\theta) = \begin{bmatrix}\tilde{U}_l(\theta) & \tilde{U}_r(\theta)\end{bmatrix} \text{ with } \tilde{U}_l(\theta) = \begin{bmatrix}\tilde{D}(\theta) & 0 \\ 0 & \tilde{F}(\theta)\end{bmatrix}$$

where $\tilde{D} \in \mathbb{R}^{(L-M)\times(L-M)}(z)$, diagonal and full rank for all $\theta \in \Theta$, and $\tilde{F}(\theta) \in \mathbb{R}^{M\times M}(z)$, and $\tilde{U}_r(\theta) \in \mathbb{R}^{L\times K}(z)$ can be fully parametrized. □

The proof of the theorem is collected in the appendix. Permutation matrix P_r is interpreted as a matrix that relabels the nodes such that the first $L-M$ nodes all have an independent component in their (external) excitation. $P(q)$ then creates the diagonal block $\tilde{D}(q,\theta)$ that explicitly shows independent excitation like in Theorem 1 to make the parameterized part $\tilde{G}_l(q,\theta)$ unique. The remaining part $\tilde{G}_r(q)$ then is unique because it is not parametrized. Theorem 2 states that when the structure of $G(q,\theta)$ is restricted then requirements on the excitation of the network can be relaxed. Note that the structure restrictions can either be formulated in terms of absence of links (zero transfers), or in known dynamical links. In both basis the corresponding entries of $G(\theta)$ do not need to be parametrized. When all transfer functions that share the same input node signal are known (the corresponding column in G is fixed) then this node does not need to have independent excitation. The result of the theorem is illustrated in an example.

Example 4. Consider network $\mathcal{S}_3$ in Figure 2. Based on the a-priori knowledge that there are no transfer functions leaving node 3, we choose the model structure $\mathcal{M}(\theta)$ according to (omitting arguments):

$$G = \begin{bmatrix} 0 & g_{12} & 0 \\ g_{21} & 0 & 0 \\ g_{31} & g_{32} & 0 \end{bmatrix},\ H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ 0 & h_{22} & h_{23} \\ 0 & 0 & h_{33} \end{bmatrix},\ R = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Global network identifiability of the model structure will be assessed by Theorem 2. Using a $P(q)$ that swaps the 2nd and 4th columns of $[H \quad R]$ delivers $\tilde{U}_l$ with $\tilde{D} = \mathrm{diag}([h_{11}\ 1])$. The third column of G contains no parameters and can take the role of $\tilde{G}_r$ with $M = 1$. Then by Theorem 2 the model is globally network identifiable. Note that in the chosen network nodes $1, 2$ are independently excited, while node 3 is not. Network identifiability is ensured because the transfer functions that have node 3 as input are not parameterized.

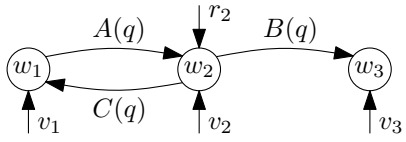

Fig. 2. Network $\mathcal{S}_3$

It is possible that even when a model structure is not globally network identifiable, some elements in the model structure can be identified uniquely when these have independent excitation.

Definition 4. (Global network identifiability of a G-row). Within $\mathcal{M}(\theta)$ a row j of $G(q,\theta)$, indicated by $G_{j\star}(q,\theta)$, is *globally network identifiable* if for all $\theta_1, \theta_2 \in \Theta$ the following implication holds:

$$T(q,\theta_1) = T(q,\theta_2) \Rightarrow G_{j\star}(q,\theta_1) = G_{j\star}(q,\theta_2)$$

These rows of G are globally network identifiable when each transfer function on that row is either independently excited or not parameterized, this is formulated in the next corollary.

Corollary 1. Within $\mathcal{M}(\theta)$ row j of $G(q,\theta)$ is *globally network identifiable* if there exists a permutation matrix $P_r \in \mathbb{R}^{L\times L}$ and nonsingular transfer function matrix $P \in \mathbb{R}^{K+L\times K+L}(z)$ such that for all $\theta \in \Theta$ the matrices

a) $\tilde{G}(\theta) := P_r G(\theta) P_r^{-1}$ has structure $\begin{bmatrix} G_{tl}(q,\theta) & G_{tr}(q,\theta) \\ G_{bl}(q,\theta) & G_{br}(q) \end{bmatrix}$ with $\tilde{G}_{tl}(\theta) \in \mathbb{R}^{(L-M)\times(L-M)}(z)$, $\tilde{G}_{bl}(\theta) \in \mathbb{R}^{M\times(L-M)}(z)$, $\tilde{G}_{tr}(\theta) \in \mathbb{R}^{(L-M)\times M}(z)$, $\tilde{G}_{br} \in \mathbb{R}^{M\times M}(z)$, such that $P_r G_{j\star}(q,\theta)$ permutes row j to row $i > L-M$, and

b) $\tilde{U}(\theta) := P_r U(\theta) P$ is structured as in Theorem 2.

The proof of the corollary follows from the proof of Theorem 2.

6. CONCLUSIONS

The objective of this paper has been to obtain conditions on the presence and location of excitation and disturbance signals and conditions on the parameterized model set such that a unique representation of the full network can be obtained. A property called global network identifiability has been defined to ensure this unique representation, and results have been derived to analyze this property. The three key ingredients for a network identifiable model set are: presence and location of external excitation signals, modeled correlations between disturbances, and prior (structural) knowledge on the network that is incorporated in the model.

Appendix A. PROOF OF THEOREM 1

First sufficient will be proven. By definition of $T(\theta)$,

$$(I - G(\theta_1))^{-1} U(\theta_1) = (I - G(\theta_2))^{-1} U(\theta_2) \qquad \text{(A.1)}$$

and postmultiplication with P leads to

$$(I - G(\theta_1))^{-1}[D(\theta_1)\ F(\theta_1)] = (I - G(\theta_2))^{-1}[D(\theta_2)\ F(\theta_2)].$$

The left square $L \times L$ blocks in both sides of the equation can now be inverted to deliver $D(\theta_1)^{-1}(I - G(\theta_1)) = D(\theta_2)^{-1}(I - G(\theta_2))$. Due to zeros on the diagonal of $G(\theta)$ and the diagonal structure of $D(\theta)$ it follows that $D(\theta_1) = D(\theta_2)$ and consequently $G(\theta_1) = G(\theta_2)$. Then by (A.1) it follows that $U(\theta_1) = U(\theta_2)$.

For the necessary part we first observe that for global network identifiability, $U\theta)$ has to have full row rank. This can be verified by noting that if $U(\theta)$ is rank deficient and $U(\theta_1) = U(\theta_2)$ then (A.1) can be satisfied even when $G(\theta_1) \neq G(\theta_2)$.

With $U(\theta)$ being full row rank, there exists a P matrix such that $U(\theta)P = [X(\theta)\ F(\theta)]$ with $X(\theta)$ square and

invertible.
Equality of T-matrices then implies $(I-G(\theta_1))^{-1}X(\theta_1) = (I-G(\theta_2))^{-1}X(\theta_2)$, and by inverting these equations:

$$X(\theta_1)^{-1}(I-G(\theta_1)) = X(\theta_2)^{-1}(I-G(\theta_2)). \qquad \text{(A.2)}$$

If G is fully and independently parametrized, then each column of G has $L-1$ independently parametrized transfer functions. Since $\mathcal{M}(\theta)$ is network identifiable each row of $X^{-1}(\theta)$ can only have one independently parameterized transfer function, otherwise there would be more parametrized transfer functions than constraints that need to be satisfied. Since X must be full rank and can only have one independently parametrized transfer function per row, a column permutation (included in P) can be found that makes X diagonal, which concludes the proof for necessity. □

Appendix B. PROOF OF THEOREM 2

Global network identifiability will be shown for $\tilde{G}$ and $\tilde{U}$, which, with the expressions of the theorem, is shown to be equivalent to a similar property for G and U.
With the expressions for $\tilde{G}$ and $\tilde{U}$, it follows that $\tilde{T}(\theta) = (I-\tilde{G}(\theta))^{-1}\tilde{U}(\theta)$ can be written as

$$\tilde{T}(\theta) = (I - P_r G(\theta) P_r^{-1})^{-1} P_r U(\theta) P = P_r T(\theta) P,$$

which implies that global network identifiability can be equivalently analysed through $\tilde{G}$ and $\tilde{U}$.
We denote

$$I-\tilde{G}(\theta) = \begin{bmatrix} \tilde{G}_{tl}(\theta) & \tilde{G}_{tr} \\ \tilde{G}_{bl}(\theta) & \tilde{G}_{br} \end{bmatrix}$$

with $\tilde{G}_{tl}(\theta) \in \mathbb{R}^{(L-M)\times(L-M)}(z)$, $\tilde{G}_{bl}(\theta) \in \mathbb{R}^{(M)\times(L-M)}(z)$, $\tilde{G}_{tr} \in \mathbb{R}^{(L-M)\times(M)}(z)$, $\tilde{G}_{br} \in \mathbb{R}^{(M)\times(M)}(z)$. Note that $\tilde{G}_{tl}(\theta)$ and $\tilde{G}_{br}$ are invertible due to well-posedness. With the matrix inversion lemma it follows that

$$(I-\tilde{G}(\theta))^{-1} = \begin{bmatrix} \tilde{Q}(\theta) & -\tilde{Q}(\theta)\tilde{G}_{tr}\tilde{G}_{br}^{-1} \\ -\tilde{G}_{br}^{-1}\tilde{G}_{bl}(\theta)\tilde{Q}(\theta) & (\tilde{G}_{br}-\tilde{G}_{bl}(\theta)\tilde{G}_{tl}^{-1}(\theta)\tilde{G}_{tr})^{-1} \end{bmatrix},$$

where $\tilde{Q} = (\tilde{G}_{tl}(\theta) - \tilde{G}_{tr}\tilde{G}_{br}^{-1}\tilde{G}_{bl}(\theta))^{-1}$.
Evaluating the equivalence relation $(I-\tilde{G}(\theta_1))^{-1}\tilde{U}(\theta_1) = (I-\tilde{G}(\theta_2))^{-1}\tilde{U}(\theta_2)$ with the expression for $\tilde{U}(\theta)$ now implies $\tilde{Q}(\theta_1)\tilde{D}(\theta_1) = \tilde{Q}(\theta_2)\tilde{D}(\theta_2)$ and

$$-\tilde{G}_{br}^{-1}\tilde{G}_{bl}(\theta_1)\tilde{Q}(\theta_1)\tilde{D}(\theta_1) = -\tilde{G}_{br}^{-1}\tilde{G}_{bl}(\theta_2)\tilde{Q}(\theta_2)\tilde{D}(\theta_2).$$

The previous two equations result in $-\tilde{G}_{br}^{-1}\tilde{G}_{bl}(\theta_1) = -\tilde{G}_{br}^{-1}\tilde{G}_{bl}(\theta_2)$ such that $\tilde{G}_{bl}(\theta_1) = \tilde{G}_{bl}(\theta_2)$. Taking the inverse of $\tilde{Q}(\theta_1)\tilde{D}(\theta_1) = \tilde{Q}(\theta_2)\tilde{D}(\theta_2)$ and using $\tilde{G}_{bl}(\theta_1) = \tilde{G}_{bl}(\theta_2)$ results in

$$(\tilde{G}_{tl}(\theta_1) - \tilde{C}(\theta_1))\tilde{D}(\theta_1)^{-1} = (\tilde{G}_{tl}(\theta_2) - \tilde{C}(\theta_2))\tilde{D}(\theta_2)^{-1},$$

where $\tilde{C}(\theta_1) = \tilde{C}(\theta_2) = \tilde{G}_{tr}\tilde{G}_{br}^{-1}\tilde{G}_{bl}(\theta_2)$. The diagonal of $\tilde{G}$ is 1 because $(I-\tilde{G})_{ii} = (P_r)_{ij}(I-G)_{jj}(P_r)_{ji} = (G)_{jj}$, where $(\cdot)_{ij}$ indicates the ij-th element of the matrix. Then $(\tilde{D}^{-1}(\theta_1))_{ii} = (\tilde{D}^{-1}(\theta_2))_{ii}$ follows, and hence $(\tilde{G}_{tl}(\theta_1) = (\tilde{G}_{tl}(\theta_2)$, resulting into $\tilde{U}(\theta_1) = \tilde{U}(\theta_2)$. □

REFERENCES

J. Adebayo, T. Southwick, V. Chetty, E. Yeung, Y. Yuan, J. Gonçalves, J. Grose, J. Prince, G.-B. Stan, and S. Warnick. Dynamical structure function identifiability conditions enabling signal structure reconstruction. In *Proc. 51st IEEE Conf. Decision and Control (CDC)*, pages 4635–4641, 2012.

A.S. Bazanella, M. Gevers, and L. Miskovic. Closed-loop identification of mimo systems: a new look at identifiability and experiment design. *European Journal of Control*, 16(3):228–239, May–June 2010.

A. Chiuso and G. Pillonetto. A bayesian approach to sparse dynamic network identification. *Automatica*, 48 (8):1553–1565, 2012.

A.G. Dankers. *System identification in dynamic networks*. PhD dissertation, Delft University of Technology, 2014.

A.G. Dankers, P.M.J. Van den Hof, X. Bombois, and P.S.C. Heuberger. Errors-in-variables identification in dynamic networks. In *Proc. 19th IFAC World Congress*, pages 2335–2340, 2014.

A.G. Dankers, P.M.J. Van den Hof, P.S.C. Heuberger, and X. Bombois. Identification of dynamic models in complex networks with predictior error methods - predictor input selection. *IEEE Trans. Automatic Control*, 61(4), 2016. To appear.

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

L. Ljung. On the consistency of prediction error identification methods. In R.K. Mehra and D.G. Lainiotis, editors, *System Identification: Advances and Case Studies*, pages 121–164. New York, 1976.

L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ, 1999.

D. Materassi and M. Salapaka. On the problem of reconstructing an unknown topology via locality properties of the Wiener filter. *IEEE Trans. Automatic Control*, 57 (7):1765–1777, 2012.

B.M. Sanandaji, T.L. Vincent, and M.B. Wakin. Exact topology identification of large-scale interconnected dynamical systems from compressive observations. In *Proc. American Control Conference (ACC)*, pages 649–656, San Francisco, CA, USA, 2011.

T. Söderström, L. Ljung, and I. Gustavsson. Identifiability conditions for linear multivariable systems operating under feedback. *IEEE Trans. Automatic Control*, 21 (6):837–840, 1976.

P.M.J. Van den Hof. A deterministic approach to approximate modelling of input-output data. In *Proc. 28th IEEE Conference on Decision and Control*, pages 659–664, Tampa, Florida, 1989.

P.M.J. Van den Hof. Model sets and parametrizations for identification of multivariable equation error models. *Automatica*, 30(3):433–446, 1994.

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

H.H.M. Weerts, A.G. Dankers, and P.M.J. Van den Hof. Identifiability in dynamic network identification. Technical report, Eindhoven University of Technology, `http://pvandenhof.nl/Reportfiles/Weerts&etal_SYSID2015_report.pdf`, 2015.

Y. Yuan, G-B. Stan, S. Warnick, and J. Gonçalves. Robust dynamical network structure reconstruction. *Automatica*, 47:1230–1235, 2011.