



Generic identifiability of subnetworks in a linear dynamic network: The full measurement case[☆]

Shengling Shi^a, Xiaodong Cheng^{a,b}, Paul M.J. Van den Hof^{a,*}

^a Control Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

^b Department of Engineering, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom



ARTICLE INFO

Article history:

Received 4 August 2020

Received in revised form 20 October 2021

Accepted 26 October 2021

Available online 18 December 2021

Keywords:

System identification

Identifiability

Dynamic networks

Graph theory

ABSTRACT

Identifiability conditions for single or multiple modules in a dynamic network specify under which conditions the considered modules can be uniquely recovered from the second-order statistical properties of the measured signals. Conditions for generic identifiability of multiple modules, i.e. a subnetwork, are developed for the situation that all node signals are measured and excitation of the network is provided by both measured excitation signals and unmeasured disturbance inputs. Additionally, the network model set is allowed to contain non-parametrized modules that are fixed, and e.g. reflect modules of which the dynamics are known to the user. The conditions take the form of path-based conditions on the graph of the network model set. Based on these conditions, synthesis results are formulated for allocating external excitation signals to achieve generic identifiability of particular subnetworks. If there are a sufficient number of measured external excitation signals, the formulated results give rise to a generalized indirect type of identification algorithm that requires only the measurement of a subset of the node signals in the network.

© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Due to the increasing complexity of current technological systems, the study of large-scale interconnected dynamic systems receives considerable attention (Cheng & Scherpen, 2021; Ishizaki, Chakraborty, & Imura, 2018; Mesbahi & Egerstedt, 2010). The resulting dynamic networks can adequately describe a wide class of complex engineering systems appearing in various applications, including multi-robot coordination (Mesbahi & Egerstedt, 2010), power grids (Ishizaki et al., 2018) and gene networks (Adebayo et al., 2012). For data-driven modeling problems in structured dynamic networks, different types of network models have been used. Connecting to prediction-error identification methods, the most popular modeling framework is based on a network of transfer functions, initially introduced in Gonçalves and Warnick (2008) and Van den Hof, Dankers, Heuberger, and Bombois (2013), where vertices represent the internal signals, that are

available for measurement, and directed edges denote transfer functions which are called modules. While there are alternatives, e.g. in the form of state-space models (Haber & Verhaegen, 2014; Yu & Verhaegen, 2018), in this paper we will adhere to the former so-called *module* framework.

Estimating network properties from data can be addressed in terms of different problem settings. One can e.g. consider the estimation of the network topology, i.e. the interconnection structure of the network (Chiuso & Pilonetto, 2012; Hayden, Chang, Goncalves, & Tomlin, 2016; Materassi & Salapaka, 2012; Sanandaji, Vincent, & Wakin, 2011; Shi, Bottegal, & Van den Hof, 2019; Zorzi & Chiuso, 2017). Another problem is to identify a single module in a network with given topology. This includes the selection of internal signals that need to be measured and excited for achieving consistent module estimates (Dankers, Van den Hof, Bombois, & Heuberger, 2016; Everitt, Bottegal, & Hjalmarsson, 2018; Gevers & Bazanella, 2015; Gevers, Bazanella, & da Silva, 2018; Linder & Enqvist, 2017; Materassi & Salapaka, 2020; Ramaswamy & Van den Hof, 2021; Van den Hof et al., 2013). Identification of the full network dynamics, for given network topology, is addressed in e.g., Chiuso and Pilonetto (2012), Fonken, Ferizbegovic, and Hjalmarsson (2020) and Weerts, Van den Hof, and Dankers (2018b).

In this paper, we focus on *network identifiability*, which is a concept that is independent of the particular identification method chosen. Based on the results for deterministic network reconstruction problems in Adebayo et al. (2012) and Gonçalves

[☆] This project has received funding from the European Research Council (ERC), Advanced Research Grant SYSDYNET, under the European Union's Horizon 2020 research and innovation programme (grant agreement No 694504). The material in this paper was partially presented at the 21st IFAC World Congress (IFAC 2020), July 12–17, 2020, Berlin, Germany. This paper was recommended for publication in revised form by Associate Editor Gianluigi Pilonetto under the direction of Editor Torsten Söderström.

* Corresponding author.

E-mail addresses: s.shi@tue.nl (S. Shi), xc336@cam.ac.uk (X. Cheng), p.m.j.vandenhof@tue.nl (P.M.J. Van den Hof).

and Warnick (2008), the concept of global network identifiability was introduced for an identification setting in Weerts, Dankers, and Van den Hof (2015) and Weerts, Van den Hof, and Dankers (2018a), as a property that reflects the ability to distinguish between network models in a parametrized model set based on measured data. In the literature, there are two notions of network identifiability, namely, *global identifiability* (van Waarde, Tesi, & Camlibel, 2020; Weerts et al., 2015, 2018a) that requires *all* model to be distinguishable from the other models in the model set,¹ and *generic identifiability* (Bazanella, Gevers, Hendrickx, & Parraga, 2017; Hendrickx, Gevers, & Bazanella, 2019; Weerts, Van den Hof, & Dankers, 2018c), which means that *almost all* models can be distinguished from all other models in the model set. It has been shown in Bazanella et al. (2017) and Hendrickx et al. (2019) that by considering generic identifiability, the algebraic conditions for identifiability can be recast into path-based conditions on the topology/graph of the network models, which largely simplifies the analysis.

Network identifiability is typically dependent on several structural properties of the model set, such as the network topology, the modeled correlation structure of process noises, the presence and location of external excitation signals and the availability of measured vertex signals. Conditions for network identifiability have been analyzed for different problem settings. In the setting with full excitation (Bazanella et al., 2017; Hendrickx et al., 2019; van Waarde et al., 2020), all vertices are excited by external excitation signals while only a subset of vertices is measured. In contrast, the full measurement setting in Weerts et al. (2018a, 2018c) assumes that all vertex signals are measured, while only a subset of them is excited. Recent contributions in Bazanella, Gevers, and Hendrickx (2019) and Shi, Cheng, and Van den Hof (2021b) also address the combined situation.

In this paper our objective is to derive path-based conditions for generic identifiability of only a subset of modules (subnetwork) in the network, while we assume all internal signals in the network to be available for measurement, the so-called full measurement case. Additionally, we will require the conditions to be suitable for solving the synthesis problem too, i.e. the allocation of a minimum set of external excitation signals, so as to achieve generic identifiability of the subnetwork.

For analyzing this problem, we start from the path-based results for generic identifiability of a subnetwork as presented in Bazanella et al. (2017) and Hendrickx et al. (2019). In these works identifiability is defined as a property of a single network, while disturbance signals are not exploited. In our approach we follow up on the more general model-set type of definition of identifiability as introduced in Weerts et al. (2018a, 2018c), that is more aligned to the use of this concept in an identification setting. Moreover, this step allows to incorporate the following attractive features that are all addressed in the current paper: (a) it allows to include the effect of unmeasured disturbance signals in the network as an information source that can contribute to satisfying identifiability conditions; (b) it allows to include modules in the network that are a priori known to the user and thus do not need to be identified. In this setting we will develop novel analysis tools for generic identifiability of subnetworks that are formulated in terms of disconnecting sets in the graph of the network models, and we will show that this leads to a new and effective synthesis procedure for allocating external excitation signals for achieving identifiability of a subnetwork. This synthesis problem was not addressed in Bazanella et al. (2017), Hendrickx et al. (2019) and Weerts et al. (2018a, 2018c). As a side result, a

¹ There are actually two versions of global identifiability, reflecting whether either one particular model in the set can be distinguished or all models in the set (Weerts et al., 2018a).

new generalized indirect identification method is described that follows immediately from the identifiability conditions, in the situation that a sufficient number of measured excitation signals is present. This method points to a subset of internal signals that in the considered situation would be sufficient for uniquely identifying the subnetwork.

The synthesis problem as formulated above is important for actually designing experimental setups for subnetwork identification. A related synthesis problem for full network identifiability has been addressed in Cheng, Shi, and Van den Hof (2022), but requires completely different tools for analysis.

The paper proceeds as follows. After introducing preliminaries and the problem in Section 2, algebraic and path-based conditions for generic identifiability are formulated in Sections 3 and 4. Disconnecting-set-based conditions are then derived in Section 5, leading to synthesis approaches in Section 6. Then a generalized indirect identification method that directly follows from the identifiability conditions is presented in Section 7.

Preliminary results of this paper were presented in Shi, Cheng, and Van den Hof (2020). In the current paper comprehensive algebraic and path-based conditions are formalized for a generalized situation, including the step from single modules to subnetworks. Additionally a novel indirect identification method is presented.

2. Preliminaries and problem formulation

2.1. Dynamic networks

The dynamic network models the relation among a set of measured scalar *internal signals* $\mathcal{W} \triangleq \{w_1(t), \dots, w_L(t)\}$ with $L = |\mathcal{W}|$ denoting the cardinality of \mathcal{W} , a set of measured excitation signals $\mathcal{R} \triangleq \{r_1(t), \dots, r_K(t)\}$ with $K = |\mathcal{R}|$, and a set of unmeasured disturbance signals $\{v_1(t), \dots, v_L(t)\}$. The model is written as

$$w(t) = G(q)w(t) + R(q)r(t) + v(t), \quad (1)$$

where $G(q)$ and $R(q)$ are matrices of rational transfer operators with delay operator q^{-1} , i.e. $q^{-1}w_i(t) = w_i(t-1)$; $w(t)$, $r(t)$ and $v(t)$ are the column vectors that collect all the internal signals, excitation signals and disturbances, respectively. In addition, $v(t)$ is a vector of zero-mean stationary stochastic processes with power spectrum $\Phi_v(\omega)$, which is modeled as a filtered white noise vector $e(t)$ according to:

$$v(t) = H(q)e(t), \quad (2)$$

where $e(t)$ has a covariance matrix Λ . Depending on whether $\Phi_v(\omega)$ is of full rank or not, $H(q)$ can either be square or have more rows than columns (Weerts et al., 2018a). Combining (1) and (2) leads to a dynamic network model that is supposed to satisfy the following **Assumption 1**.

Assumption 1.

- $G(q)$ has zero diagonal elements, and $G(q)$ is proper and stable;
- The network is well-posed in the sense that all principal minors of $\lim_{z \rightarrow \infty} (I - G(z))$ are non-zero;
- $[I - G(q)]^{-1}$ is stable;
- $R(q)$ is a proper and stable rational transfer matrix;
- $H(q)$ is minimum phase and monic if square; for the non-square case, i.e. when $\Phi_v(\omega)$ is singular, $H(q)$ is structured as $H(q) = \begin{bmatrix} H_a(q) \\ H_b(q) \end{bmatrix}$, with $H_a(q)$ square, monic and minimum phase (Weerts et al., 2018a);
- The covariance matrix Λ of $e(t)$ is positive definite.

In the above assumptions, (b) and (c) ensure that every transfer function from external signals to internal signals is proper and stable (Dankers, 2014); the stability of $G(q)$ in (a) guarantees that the noise filter defined later in (3) is also inversely stable, which is typical for the modeling of stationary stochastic processes.

Both the excitation signals in \mathcal{R} and the entries (white noises) in $e(t)$ are called *external signals*, and the set of all external signals is denoted by \mathcal{X} . The entries in $G(q)$ are referred to as *modules*. In addition, the dynamic network model leads to mappings from the excitation signals to internal signals:

$$w(t) = T_{\mathcal{W}\mathcal{X}}(q) \begin{bmatrix} r(t) \\ e(t) \end{bmatrix} = T_{\mathcal{W}\mathcal{R}}(q)r(t) + \bar{v}(t), \quad (3)$$

where $T_{\mathcal{W}\mathcal{X}}(q) \triangleq [I - G(q)]^{-1}X(q)$, $X(q) \triangleq [R(q) \ H(q)]$, $T_{\mathcal{W}\mathcal{R}}(q) \triangleq [I - G(q)]^{-1}R(q)$, $\bar{v}(t) \triangleq [I - G(q)]^{-1}H(q)e(t)$.

The power spectrum $\Phi_{\bar{v}}(\omega)$ of $\bar{v}(t)$ satisfies $\Phi_{\bar{v}}(\omega) = [I - G(e^{i\omega})]^{-1}H(e^{i\omega})\Lambda H^T(e^{-i\omega})[I - G(e^{-i\omega})]^{-T}$. When applying common statistical identification methods to (3) that are based on first and second moment information of the measured signals, typically the objects $T_{\mathcal{W}\mathcal{R}}(q)$ and $\Phi_{\bar{v}}(\omega)$ can be consistently estimated from measured signals $w(t)$ and $r(t)$, if $r(t)$ is persistently exciting (Ljung, 1987). This motivates to use $T_{\mathcal{W}\mathcal{R}}(q)$ and $\Phi_{\bar{v}}(\omega)$ as bases for network identifiability (Weerts et al., 2018a).

Given subsets $\bar{\mathcal{W}} \subseteq \mathcal{W}$ and $\bar{\mathcal{X}} \subseteq \mathcal{X}$, $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}(q)$ denotes a submatrix of $T_{\mathcal{W}\mathcal{X}}(q)$ with the rows and columns corresponding to the signals in $\bar{\mathcal{W}}$ and $\bar{\mathcal{X}}$. If $\bar{\mathcal{W}}$ contains only one signal w_k , $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}(q)$ is often simply written as $T_{k\bar{\mathcal{X}}}(q)$. The above notation applies similarly to submatrices of other matrices and vectors.

2.2. Model sets

Network identifiability will be defined on the basis of a network model set, which is introduced as follows. The network model (1), (2) is completely specified by a quadruple $M \triangleq (G(q), R(q), H(q), \Lambda)$. By parametrizing the entries of the network matrices in a rational form and then collecting the parameters into a parameter vector θ , a parametrized model set can be defined.

Definition 1. Consider a rational parametrization of a network model according to

$$M(\theta) = (G(q, \theta), R(q, \theta), H(q, \theta), \Lambda(\theta)).$$

Then a network model set \mathcal{M} is defined as $\mathcal{M} = \{M(\theta) | \theta \in \Theta \subseteq \mathbb{R}^n\}$, where $M(\theta)$ satisfies Assumption 1 for all $\theta \in \Theta$.

In addition, there can be certain entries in $G(q, \theta)$, $R(q, \theta)$ and $H(q, \theta)$ that are fixed/known and thus do not depend on the parameters. These entries are said to be *known* or *fixed*, and all models in \mathcal{M} contain the same known entries. The entries in the network matrices that depend on the parameters are said to be *unknown* or *parametrized*. For example, the absence of an interconnection between internal signals is represented by a fixed 0 in $G(q, \theta)$; entries in $G(q, \theta)$ may be particularly designed controllers that are fixed and known. Similarly, entries in $H(q, \theta)$, $R(q, \theta)$ and $\Lambda(\theta)$ can be fixed, e.g., equal to 1 or 0. In the situation of $R(q)$, this implies that it can be specified upfront on which nodes the r -signals enter the network. Note that the dependency of transfer matrices on q, θ and the dependency of power spectra on ω are sometimes omitted for simplicity of notation.

The structural information of a model set is reflected by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} \triangleq \mathcal{W} \cup \mathcal{X}$ is a set of vertices representing both the internal signals and the external signals, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes a set of directed edges representing those entries in $G(q, \theta)$ and $X(q, \theta)$ that are not fixed to zero: a directed edge from w_i to w_j exists, i.e. $(w_i, w_j) \in \mathcal{E}$, if G_{ji} is not fixed to

zero; Similarly, $(e_k, w_j) \in \mathcal{E}$ and $(r_p, w_j) \in \mathcal{E}$ if H_{jk} and R_{jp} are not fixed to zero, respectively. In this way, any parametrized model set or network model induces a directed graph \mathcal{G} . Note that w_i now represents both a signal and a vertex, and its dependency on t is sometimes omitted for simplicity of notation.

For deriving the results in the sequel of this paper, we will need an additional concept.

Definition 2. Given a model set \mathcal{M} with its graph \mathcal{G} , the set \mathbb{G}^* is defined as the set of all $G(q)$ matrices that satisfy the following conditions: (i) they meet Assumption 1; (ii) they have the same fixed entries as $G(q, \theta)$ in \mathcal{M} ; (iii) entries are strictly proper if the corresponding entries in $G(q, \theta)$ are (parametrized to be) strictly proper.

In Definition 2, \mathbb{G}^* is the largest possible set of $G(q)$ matrices that have the same fixed entries and feedthrough structure as $G(q, \theta)$ in \mathcal{M} . So while the modules G in \mathcal{M} can be restricted in order, the order of entries in \mathbb{G}^* is not bounded. This implies $\{G(q, \theta) | \theta \in \Theta\} \subseteq \mathbb{G}^*$. In Assumption 4 we will introduce a technical condition on $\{G(q, \theta) | \theta \in \Theta\}$ with respect to \mathbb{G}^* for proving necessary conditions for identifiability.

2.3. Network identifiability

We focus on the identifiability of a subset of unknown modules in one row of the G matrix, i.e. a subnetwork of modules that share the same output node w_j . For this output node w_j , we define an important set of signals:

- \mathcal{W}_j : all the internal signals that have unknown directed edges (modules) to w_j .

For any subset $\bar{\mathcal{W}}_j \subseteq \mathcal{W}_j$, $G_{j\bar{\mathcal{W}}_j}$ denotes a row vector containing a subset of unknown modules in the j th row of G . For the identifiability of $G_{j\bar{\mathcal{W}}_j}$, we follow the concept of global network identifiability in Weerts et al. (2018a), and extend it with a generic version introduced in Bazanella et al. (2017) and Hendrickx et al. (2019). In this respect we follow an approach that was suggested in Weerts et al. (2018c).

Definition 3. Given a network model set \mathcal{M} , consider any subset $\bar{\mathcal{W}}_j \subseteq \mathcal{W}_j$, a parameter vector $\theta_0 \in \Theta$ and the following implication:

$$\left. \begin{aligned} T_{\mathcal{W}\mathcal{R}}(q, \theta_0) &= T_{\mathcal{W}\mathcal{R}}(q, \theta_1) \\ \Phi_{\bar{v}}(\omega, \theta_0) &= \Phi_{\bar{v}}(\omega, \theta_1) \end{aligned} \right\} \Rightarrow G_{j\bar{\mathcal{W}}_j}(q, \theta_0) = G_{j\bar{\mathcal{W}}_j}(q, \theta_1) \quad (4)$$

for all $\theta_1 \in \Theta$. Then $G_{j\bar{\mathcal{W}}_j}(q, \theta)$ is said to be

- globally identifiable in \mathcal{M} from (w, r) if the implication (4) holds for all $\theta_0 \in \Theta$;
- generically identifiable in \mathcal{M} from (w, r) if the implication (4) holds for almost all $\theta_0 \in \Theta$.

In Definition 3, the notion ‘‘almost all’’ excludes a subset of Lebesgue measure zero from Θ . The identifiability concept in Definition 3 concerns whether $G_{j\bar{\mathcal{W}}_j}$ is unique given the objects that can typically be identified from the first and second moment information of the measured signals, as motivated following (3). If a model set is not identifiable, any identification method that relies on the first and second moments for estimating the network cannot be guaranteed to (asymptotically) identify a unique network model. In addition, recall that the network matrices in \mathcal{M} may contain known entries, and this restriction of \mathcal{M} may simplify the conditions under which the implication (4) holds.

To further simplify (4), we use the following assumption on the feedthrough terms in G .

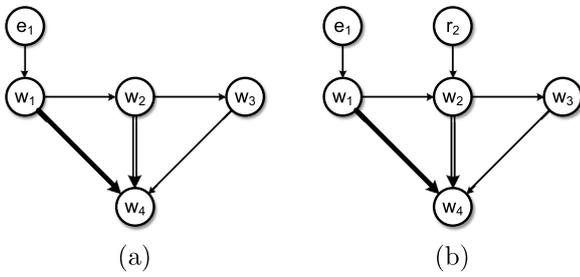


Fig. 1. Generic identifiability of G_{41} is considered (thick line), and G_{42}^0 is known (double-line edge). G_{41} is not generically identifiable in (a) but becomes generically identifiable in (b) if an extra signal r_2 is allocated at w_2 .

Assumption 2 (Weerts et al., 2018a).

- (a) either all modules $G(q, \theta)$ are parametrized to be strictly proper, or
- (b) the parametrized network model does not contain any algebraic loops,² and $H^\infty(\theta)\Lambda(\theta)[H^\infty(\theta)]^T$ is diagonal for all $\theta \in \Theta$, with $H^\infty(\theta) \triangleq \lim_{z \rightarrow \infty} H(z, \theta)$.

This assumption ensures that the spectral factorization of $\Phi_{\bar{v}}$ admits a unique spectral factor $(I - G)^{-1}H$ (Weerts et al., 2018a), i.e., the mapping from the white noises to the internal signals as in (3). Then we can reformulate Definition 3 by considering the uniqueness of network modules given $T_{\mathcal{W}\mathcal{X}} = [T_{\mathcal{W}\mathcal{R}} \ (I - G)^{-1}H]$.

Proposition 1. Given a network model set \mathcal{M} that satisfies Assumption 2, consider any subset $\bar{\mathcal{W}}_j \subseteq \mathcal{W}_j$, $\theta_0 \in \Theta$ and the following implication:

$$T_{\mathcal{W}\mathcal{X}}(q, \theta_0) = T_{\mathcal{W}\mathcal{X}}(q, \theta_1) \Rightarrow G_{j\bar{\mathcal{W}}_j}(q, \theta_0) = G_{j\bar{\mathcal{W}}_j}(q, \theta_1),$$

for all $\theta_1 \in \Theta$. Then $G_{j\bar{\mathcal{W}}_j}(q, \theta)$ is globally (generically) identifiable in \mathcal{M} from (w, r) if and only if the above implication holds for all (almost all) $\theta_0 \in \Theta$.

Proof. This result is a direct consequence of Propositions 1 and 2 in Weerts et al. (2018a). ■

In contrast to Definition 3, Proposition 1 concerns the mapping $T_{\mathcal{W}\mathcal{X}}$ from both r and e signals to the internal signals. Therefore, the r and e signals play the same role in the identifiability analysis. This is also a key difference from the identifiability concept in Bazanella et al. (2017) and Hendrickx et al. (2019), where the noise information is not exploited. Proposition 1 also extends trivially to the single module case where $G_{j\bar{\mathcal{W}}_j}$ only contains one entry, and to the full network case where $\bar{\mathcal{W}}_j = \mathcal{W}_j$ and j varies over all $j \in [1, L]$.

2.4. Problem formulation

The identifiability issue is first illustrated in Example 1.

Example 1. Consider a network model set whose graph is shown in Fig. 1(a), where there is one white noise e_1 and four measured internal signals. All edges are parametrized except for G_{42} which is fixed to G_{42}^0 . The identifiability of module G_{41} is considered, and recall that the dependence of transfer functions on q and θ is omitted for simplicity of notation. Then based on Proposition 1, the identifiability of G_{41} concerns the uniqueness

² There exists an algebraic loop around node w_{n_1} if there exists a sequence of integers n_1, \dots, n_k such that $G_{n_1 n_2}^\infty G_{n_2 n_3}^\infty \dots G_{n_k n_1}^\infty \neq 0$, with $G_{n_1 n_2}^\infty := \lim_{z \rightarrow \infty} G_{n_1 n_2}(z)$.

of G_{41} given the mapping $T_{\{w_1, w_2, w_3, w_4\}\{e_1\}}$ from e_1 to all the internal signals. According to Fig. 1(a), the mapping satisfies that $T_{\{w_1\}\{e_1\}} = H_{11}$, $T_{\{w_2\}\{e_1\}} = G_{21}H_{11}$, $T_{\{w_3\}\{e_1\}} = G_{32}G_{21}H_{11}$ and $T_{\{w_4\}\{e_1\}} = H_{11}(G_{41} + G_{42}^0 G_{21} + G_{43}G_{32}G_{21})$, where G_{42}^0 and the submatrices of $T_{\{w_1, w_2, w_3, w_4\}\{e_1\}}$ are given, but H_{11} and the other entries in G are unknown. The above equations further lead to

$$T_{\{w_4\}\{e_1\}} = T_{\{w_1\}\{e_1\}}G_{41} + G_{42}^0 T_{\{w_2\}\{e_1\}} + G_{43}T_{\{w_3\}\{e_1\}},$$

which is not sufficient for a unique solution of G_{41} due to the two unknowns G_{41} and G_{43} . Therefore, G_{41} is neither globally nor generically identifiable in Fig. 1(a).

However, if an additional signal r_2 is allocated at w_2 as in Fig. 1(b), additional mappings from r_2 to the measured internal signals are given. Then it can be found that

$$G_{41} = (T_{\{w_4\}\{e_1\}} - \frac{T_{\{w_2\}\{e_1\}}T_{\{w_4\}\{r_2\}}}{T_{\{w_2\}\{r_2\}}})/T_{\{w_1\}\{e_1\}},$$

which shows that a unique G_{41} can be found if the above inverses exist. Recall that the mappings depend on parameter θ , and thus G_{41} is globally (generically) identifiable if the inverses exist for all (almost all) $\theta \in \Theta$.

As shown in the above example, identifiability is related to the uniqueness of modules given the mapping $T_{\mathcal{W}\mathcal{X}}$. In addition, it is possible to achieve identifiability by allocating extra excitation signals in the network.

Given a model set and its associated graph, we first develop sufficient and necessary conditions for the generic identifiability of a subnetwork $G_{j\bar{\mathcal{W}}_j}$ when all the internal signals are measured. More importantly, to avoid the complex algebraic verification as in Example 1, we aim to develop graphical conditions that can be verified efficiently by only inspecting the network topology. Furthermore, when a subnetwork is not generically identifiable, we aim to develop graphical synthesis approaches that allocate additional excitation signals to automatically achieve the generic identifiability of a subnetwork.

2.5. Notations and definitions

The following notations are used in the paper. Matrix $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}(q, \theta)$ is called having generically full rank if it has full rank for almost all θ . More generally, a property that depends on θ is said to hold generically if it holds for almost all θ . A matrix C is called a selection matrix if it consists of a subset of rows of an identity matrix.

In a graph \mathcal{G} , for a directed edge (w_i, w_j) , w_i is called an *in-neighbor* of w_j , and w_j is an *out-neighbor* of w_i . The set $\mathcal{N}_{\bar{\mathcal{V}}}^+$ contains all out-neighbors of the vertices in $\bar{\mathcal{V}}$, and the set $\mathcal{N}_{\bar{\mathcal{V}}}^-$ all in-neighbors. A (directed) *path* from w_i to w_j is a sequence of vertices and out-going edges starting from w_i to w_j without repeating any vertex, and w_j is said *reachable* by w_i . A single vertex is also regarded as a directed path to itself. *Internal vertices* of a path are the vertices excluding the starting and the ending vertices.

Two directed paths are called *vertex disjoint* if they do not share any vertex, including the starting and ending vertices, otherwise they *intersect*. Given two subsets of vertices \mathcal{V}_1 and \mathcal{V}_2 , $b_{\mathcal{V}_1 \rightarrow \mathcal{V}_2}$ denotes the maximum number of vertex disjoint paths from \mathcal{V}_1 to \mathcal{V}_2 . A vertex set \mathcal{D} is a $\mathcal{V}_1 - \mathcal{V}_2$ *disconnecting set* if it intersects with all paths from \mathcal{V}_1 to \mathcal{V}_2 , and it is a minimum disconnecting set if it has the minimum cardinality among all $\mathcal{V}_1 - \mathcal{V}_2$ disconnecting sets (Schrijver, 2003). Note that \mathcal{D} may also include vertices in $\mathcal{V}_1 \cup \mathcal{V}_2$.

Lemma 1 (Hendrickx et al., 2019). For a directed graph and given a $\mathcal{V}_1 - \mathcal{V}_2$ disconnecting set \mathcal{D} , consider the division of all the vertices \mathcal{V}

into three disjoint sets $S \cup \mathcal{D} \cup \mathcal{P}$ as follows: set S contains all vertices reachable by \mathcal{V}_1 without intersecting \mathcal{D} , and $\mathcal{P} = \mathcal{V} \setminus (\mathcal{D} \cup S)$. Then it holds that no directed edge exists from S to \mathcal{P} .

The duality between vertex disjoint paths and disconnecting sets is explained in Menger's theorem.

Theorem 1 (Schrijver, 2003). *Let $\mathcal{V}_1, \mathcal{V}_2$ be two subsets of the vertices in a directed graph. The maximum number of vertex disjoint paths from \mathcal{V}_1 to \mathcal{V}_2 equals the cardinality of a minimum disconnecting set from \mathcal{V}_1 to \mathcal{V}_2 .*

3. Algebraic conditions for identifiability

We first investigate under which conditions a subnetwork $G_{j\bar{\mathcal{V}}_j}$ is globally and generically identifiable. An algebraic rank condition is first presented in this section which will serve as a basis for deriving a more attractive graphical condition in the next section. Recall (3) which leads to $(I - G)T_{\mathcal{W}\mathcal{X}} = X$. It then follows from Proposition 1 that identifiability essentially reflects the unique solutions of modules in G from a given $T_{\mathcal{W}\mathcal{X}}$ and thus is related to the rank of $T_{\mathcal{W}\mathcal{X}}$. We first define the following set related to the output w_j :

- \mathcal{X}_j : the set of all the external signals that do not have any unknown edge to w_j .

Then the following sufficient conditions for identifiability can be obtained from Theorem V.1 in Hendrickx et al. (2019) analogously.

Lemma 2. *Given a model set \mathcal{M} that satisfies Assumption 2. Then $G_{j\bar{\mathcal{V}}_j}(q, \theta)$ is globally (generically) identifiable in \mathcal{M} from (w, r) if the following conditions hold:*

$$\text{rank}[T_{\bar{\mathcal{V}}_j\mathcal{X}_j}(q, \theta)] = |\bar{\mathcal{V}}_j|, \text{ and} \quad (5a)$$

$$\text{rank}[T_{\mathcal{W}_j\mathcal{X}_j}(q, \theta)] = \text{rank}[T_{\bar{\mathcal{V}}_j\mathcal{X}_j}(q, \theta)] + \text{rank}[T_{(\mathcal{W}_j \setminus \bar{\mathcal{V}}_j)\mathcal{X}_j}(q, \theta)] \quad (5b)$$

for all (almost all) $\theta \in \Theta$.

The above result shows that identifiability of $G_{j\bar{\mathcal{V}}_j}$ holds if (i) the mapping from \mathcal{X}_j to the inputs of the target modules has full row rank, and (ii) the rows in the above mapping and the ones in the mapping from \mathcal{X}_j to the other inputs of w_j are linearly independent. Here, we consider the signals in \mathcal{X}_j instead of \mathcal{X} , because the ones in $\mathcal{X} \setminus \mathcal{X}_j$, which have unknown edges to w_j , contribute to extra unknowns in the analysis in the form of a parametrized disturbance model for v_j , and thus are not helpful for the identifiability analysis. If the corresponding noise model is fixed/known, the corresponding white noise source becomes an element of \mathcal{X}_j and does act as an excitation source in the identifiability analysis.

The sufficient conditions in Lemma 2 are implied in the proof of Theorem V.1 of Hendrickx et al. (2019). While in that proof they were shown to be also necessary, that necessity proof cannot be applied directly to our current setting, owing to a different definition of generic identifiability: Definition 3 is based on a particular model set, while the identifiability concept in Hendrickx et al. (2019) concerns a single network model and its associated graph. To establish the necessity of the rank conditions in this work, we first introduce two technical conditions on the model set \mathcal{M} .

Assumption 3. All the parametrized entries in $M(\theta)$ are parametrized independently.

Assumption 4. Given the network model set \mathcal{M} and the corresponding \mathbb{G}^* defined in Definition 2, $\{G(q, \theta) | \theta \in \Theta\}$ is a non-empty open subset³ of \mathbb{G}^* .

Then the necessity of the rank conditions for identifiability can be obtained as follows.

Theorem 2. *Given a model set \mathcal{M} that satisfies Assumption 2, if \mathcal{M} also satisfies Assumptions 3 and 4, conditions (5a)–(5b) in Lemma 2 are also necessary conditions for the results of the lemma.*

The proof of Theorem 2 is presented in Shi, Cheng, and Van den Hof (2021c) and a more comprehensive analysis can be found in Shi (2021). Theorem 2 extends the identifiability conditions in Weerts et al. (2018a, 2018c) from a full network and a single module to any subnetwork, such that the full network and the single module identifiability become special cases. In addition, with extra technical efforts in the analysis of Theorem 2, Assumption 4 becomes less conservative than the assumption in Weerts et al. (2018a, 2018c) for the necessity of identifiability.

The rank tests in Theorem 2 may require to invert $(I - G)$ for computing the external-to-internal mappings, which can be computationally expensive when the network is large-scale. The following result shows that the rank of any submatrix of $T_{\mathcal{W}\mathcal{X}}$ can be also evaluated without inverting the matrix $I - G$ for computing $T_{\mathcal{W}\mathcal{X}}$.

Lemma 3. *Consider a dynamic network and any subsets $\bar{\mathcal{W}} \subseteq \mathcal{W}$ and $\bar{\mathcal{X}} \subseteq \mathcal{X}$. Define the matrix $F(\bar{\mathcal{W}}, \bar{\mathcal{X}})$ as*

$$F(\bar{\mathcal{W}}, \bar{\mathcal{X}}) \triangleq \begin{bmatrix} [G(q) - I]_{\mathcal{W}(\mathcal{W} \setminus \bar{\mathcal{W}})} & X_{\mathcal{W}\bar{\mathcal{X}}}(q) \end{bmatrix}. \quad (6)$$

Then it holds that $\text{rank}(T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}(q)) = \text{rank}[F(\bar{\mathcal{W}}, \bar{\mathcal{X}})] + |\bar{\mathcal{W}}| - L$, where L is the total number of internal signals.

Proof. Recall that $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}} = C(I - G)^{-1}\bar{X}$, where $\bar{X} = X_{\mathcal{W}\bar{\mathcal{X}}}$ and C is a selection matrix that extracts the rows corresponding to $\bar{\mathcal{W}}$. Define

$$\bar{F} = \begin{bmatrix} G - I & \bar{X} \\ C & 0 \end{bmatrix}, \quad (7)$$

then according to van der Woude (1991), \bar{F} can be written as

$$\begin{bmatrix} I & 0 \\ C(G - I)^{-1} & I \end{bmatrix} \begin{bmatrix} G - I & 0 \\ 0 & C(I - G)^{-1}\bar{X} \end{bmatrix} \begin{bmatrix} I & (G - I)^{-1}\bar{X} \\ 0 & I \end{bmatrix}.$$

Thus, $\text{rank}(\bar{F}) = \text{rank}(T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}) + L$. Since F is obtained from \bar{F} by removing the columns in \bar{F} that correspond to the nonzero entries of C , which through the structure of \bar{F} are linearly independent in F , it follows that $\text{rank}(F) = \text{rank}(\bar{F}) - |\bar{\mathcal{W}}|$, which proves the result. ■

With the above result, the rank condition in Theorem 2 can be reformulated as follows.

Proposition 2. *Consider a network model set and the matrix F defined in (6). The conditions (5a)–(5b) in Theorem 2 are equivalently reformulated as $F(\bar{\mathcal{V}}_j, \mathcal{X}_j)$ having full row rank and*

$$\text{rank}[F(\mathcal{W}_j, \mathcal{X}_j)] = \text{rank}[F(\bar{\mathcal{V}}_j, \mathcal{X}_j)] + \text{rank}[F(\mathcal{W}_j \setminus \bar{\mathcal{V}}_j, \mathcal{X}_j)] - L,$$

for all (almost all) $\theta \in \Theta$.

³ Openness is considered in the metric space equipped with the H_∞ norm.

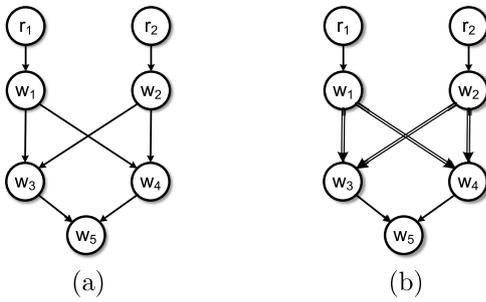


Fig. 2. Two network model sets where (b) contains known modules (double-line edges).

4. Path-based conditions for generic identifiability

When analyzing generic identifiability, the results of Lemma 2 and Proposition 2 show that the rank of the parametrized matrices have to be evaluated over almost all $\theta \in \Theta$. The resulting generic rank of transfer matrices in dynamic networks can be evaluated by determining the number of vertex disjoint paths between particular vertices in the graph induced by the model set, i.e.

$$b_{\bar{x} \rightarrow \bar{y}} = \text{rank } T_{\bar{y}\bar{x}}(q, \theta) \text{ for almost all } \theta \in \Theta. \quad (8)$$

This result from Hendrickx et al. (2019) is very attractive as it allows us to avoid numerical rank evaluations, and therefore significantly facilitates the analysis of generic identifiability. However, the result in Hendrickx et al. (2019) has been derived for the particular situation that $X = I$, and that all edges in the graph \mathcal{G} are parametrized, i.e. without having fixed/known modules in the model set. In this section we will extend the results beyond these limitations.

The issue caused by known transfer functions is explained in the following example.

Example 2. Consider two network model sets in Fig. 2, where the model set in (b) contains known modules. In addition, assume $R_{11} = R_{22} = 1$ for simplicity. Then in Fig. 2(a), $b_{\{r_1, r_2\} \rightarrow \{w_3, w_4\}} = 2$, i.e. there exist maximally two paths $r_1 \rightarrow w_1 \rightarrow w_3$ and $r_2 \rightarrow w_2 \rightarrow w_4$ that are vertex disjoint. This implies that $\text{rank}(T_{\{w_3, w_4\}\{r_1, r_2\}}) = 2$ generically according to (8), which can also be seen from

$$\det(T_{\{w_3, w_4\}\{r_1, r_2\}}) = G_{31}G_{42} - G_{41}G_{32}. \quad (9)$$

As each module is an analytic function of independent parameters, the determinant in (9) is a non-constant analytic function of θ and thus is non-zero for almost all θ . This matches with (8) that $\text{rank}(T_{\{w_3, w_4\}\{r_1, r_2\}}) = b_{\{r_1, r_2\} \rightarrow \{w_3, w_4\}} = 2$ generically.

However, $T_{\{w_3, w_4\}\{r_1, r_2\}}$ is not of full rank in Fig. 2(b) if the known modules $G_{31}^0, G_{42}^0, G_{41}^0, G_{32}^0$ take values such that $G_{31}^0 G_{42}^0 - G_{41}^0 G_{32}^0 = 0$. Then $\det(T_{\{w_3, w_4\}\{r_1, r_2\}}) = 0$ for any θ and thus (8) does not hold. ■

As shown in Example 2, the presence of known modules may cause the equality in (8) to fail. In order to circumvent this, we introduce an additional condition based on the concept of structural rank.

Definition 4 (Steffen, 2005). The structural rank of a matrix is the maximum rank of all matrices with the same nonzero pattern. A matrix has full structural rank if it can be permuted so that the diagonal has no zero entries.

The property of structural full rank depends solely on the sparsity pattern of the matrix and does not rely on the numerical values of the entries. In order to characterize and exclude the situations as presented in Example 2, we introduce the following assumption.

Assumption 5. In model set \mathcal{M} , the rank of any submatrix of $[(G(q, \theta) - I)X(q, \theta)]$, which does not depend on θ , is equal to its structural rank. ■

The above assumption prevents known entries to be chosen in such a specific manner that they cause a rank deficiency of $T_{\bar{y}\bar{x}}$, which is not captured by the number of vertex disjoint paths. Using this assumption, equality (8) can be extended from Hendrickx et al. (2019) to incorporate known transfer operators in the model set, as shown in the following result.

Proposition 3. Consider a model set \mathcal{M} with graph \mathcal{G} that satisfies Assumptions 3 and 5. Then for any subsets $\bar{\mathcal{W}} \subseteq \mathcal{W}$ and $\bar{\mathcal{X}} \subseteq \mathcal{X}$, it holds that

$$b_{\bar{x} \rightarrow \bar{y}} = \text{rank}[T_{\bar{y}\bar{x}}(q, \theta)] \text{ for almost all } \theta \in \Theta.$$

The proof of the above result is presented in Appendix. By combining Proposition 3 with Theorem 2, the graphical condition for generic identifiability can be obtained, that follows immediately from the previous results:

Theorem 3. Consider a model set \mathcal{M} that satisfies Assumptions 2, 3 and 5. Then $G_{j\bar{\mathcal{W}}_j}(q, \theta)$ is generically identifiable in \mathcal{M} from (w, r) if the following conditions hold:

$$b_{x_j \rightarrow \bar{\mathcal{W}}_j} = |\bar{\mathcal{W}}_j|, \text{ and} \quad (10a)$$

$$b_{x_j \rightarrow \mathcal{W}_j} = b_{x_j \rightarrow \bar{\mathcal{W}}_j} + b_{x_j \rightarrow \mathcal{W}_j \setminus \bar{\mathcal{W}}_j}. \quad (10b)$$

When \mathcal{M} additionally satisfies Assumption 4, the above conditions are also necessary.

The above result shows that generic identifiability of the target modules in $G_{j\bar{\mathcal{W}}_j}$ is ensured if (i) the inputs $\bar{\mathcal{W}}_j$ of the target modules are excited by x_j in the sense that x_j has a sufficient number of vertex disjoint paths to these inputs, and (ii) the inputs $\bar{\mathcal{W}}_j$ receive excitation that is independent of the excitation for the remaining inputs $\mathcal{W}_j \setminus \bar{\mathcal{W}}_j$. The result of Theorem 3 is also illustrated in Example 3. An important special case of Theorem 3 is when $\bar{\mathcal{W}}_j = \mathcal{W}_j$, i.e. all parametrized modules that map into w_j are in the subnetwork under study. In this case, conditions (10a) and (10b) can be simplified into a single condition as $b_{x_j \rightarrow w_j} = |\mathcal{W}_j|$.

Example 3. The generic identifiability of G_{41} in the network model set in Fig. 1(a) can be verified using Theorem 3 as follows. In this model set, $\mathcal{X}_4 = \{e_1\}$, $\mathcal{W}_4 = \{w_1, w_3\}$ and $\bar{\mathcal{W}}_4 = \{w_1\}$. As there exists a path from e_1 to w_1 , we have $b_{x_4 \rightarrow \bar{\mathcal{W}}_4} = 1 = |\bar{\mathcal{W}}_4|$ and thus (10a) is satisfied. In addition, since all paths from e_1 to $\{w_1, w_3\}$ intersect with w_1 , we have $b_{x_4 \rightarrow \mathcal{W}_4} = b_{x_4 \rightarrow \mathcal{W}_4 \setminus \bar{\mathcal{W}}_4} = 1$, which shows that $b_{x_4 \rightarrow \mathcal{W}_4} < b_{x_4 \rightarrow \bar{\mathcal{W}}_4} + b_{x_4 \rightarrow \mathcal{W}_4 \setminus \bar{\mathcal{W}}_4}$, and thus (10b) is not satisfied, which shows that G_{41} is not generically identifiable in Fig. 1(a).

Compared to Theorem 2 and the analysis in Example 1, the conditions in Theorem 3 are fully graphical and thus can be evaluated more efficiently by only inspecting the network topology. Standard graphical algorithms are available for computing the maximum number of vertex disjoint paths (Schrijver, 2003). Note that the conditions (10a)–(10b) are similar to the ones in Hendrickx et al. (2019), but now extended to a different situation. Besides the incorporation of known modules and the

model-set-based identifiability in Definition 3, the set \mathcal{X}_j contains both measured excitation signals and unmeasured white noises, instead of only r signals as considered in Hendrickx et al. (2019). Thus, to satisfy the graphical conditions in the theorem, the white noises can compensate for a possible lack of excitation signals.

5. Disconnecting-set-based conditions

5.1. Generic identifiability based on disconnecting sets

The graphic condition in Theorem 3 can be applied to analyzing generic identifiability in a given model set. However, it does not explicitly indicate where to allocate external signals such that a particular set of modules becomes generically identifiable. To solve this synthesis question, a new analytic result is developed in this section by exploring the duality between vertex disjoint paths and disconnecting sets, as also indicated in Hendrickx et al. (2019). In this section, we will exploit this relationship in a novel way, by suitably exploiting it for the situation of subnetwork generic identifiability, while providing a solution to the synthesis question as well. This synthesis question will be addressed in more detail in Section 6.

A set of vertices is called a disconnecting set from vertex set \mathcal{V}_1 to \mathcal{V}_2 if all paths from \mathcal{V}_1 to \mathcal{V}_2 pass through the set, see Section 2.5. Its relevance in the analysis of generic identifiability is illustrated in Example 4.

Example 4. Starting from Fig. 1(a), if r_2 is allocated at w_2 as in Fig. 1(b) such that $\mathcal{X}_4 = \{e_1, r_2\}$ and $b_{\mathcal{X}_4 \rightarrow \mathcal{W}_4} = 2$, we have $b_{\mathcal{X}_4 \rightarrow \mathcal{W}_4} = b_{\mathcal{X}_4 \rightarrow \tilde{\mathcal{W}}_4} + b_{\mathcal{X}_4 \rightarrow \mathcal{W}_4 \setminus \tilde{\mathcal{W}}_4} = 2$, and thus G_{41} becomes generically identifiable in Fig. 1(b) based on Theorem 3. It is important to note that the vertex w_2 , where the new r_2 is allocated, is a disconnecting set from the target input w_1 to the other inputs of w_4 . ■

It is observed from the above example that when G_{ji} is not generically identifiable, its identifiability can be achieved when the vertices in a disconnecting set from $\{w_i\}$ to the other in-neighbors of w_j are excited. This observation is generalized to obtain a new identifiability condition based on disconnecting sets.

Theorem 4. Consider the situation of Theorem 3. The conditions (10a)–(10b) can equivalently be written as either one of the following two equivalent conditions:

(a) There exists a $\mathcal{X}_j - \mathcal{W}_j \setminus \tilde{\mathcal{W}}_j$ disconnecting set \mathcal{D} such that $b_{\mathcal{X}_j \rightarrow \tilde{\mathcal{W}}_j \cup \mathcal{D}} = |\mathcal{D}| + |\tilde{\mathcal{W}}_j|$.

(b) There exists a set of external signals $\tilde{\mathcal{X}}_j \subseteq \mathcal{X}_j$ and a $\tilde{\mathcal{X}}_j - \mathcal{W}_j \setminus \tilde{\mathcal{W}}_j$ disconnecting set \mathcal{D} such that

$$b_{\tilde{\mathcal{X}}_j \rightarrow \tilde{\mathcal{W}}_j \cup \mathcal{D}} = |\mathcal{D}| + |\tilde{\mathcal{W}}_j|. \quad (11)$$

The proof of the above result is presented in Appendix. As visualized in Fig. 3, this result states that the target modules in $G_{j\tilde{\mathcal{W}}_j}$ are generically identifiable, if the inputs of the target modules in $\tilde{\mathcal{W}}_j$ and the signals in \mathcal{D} are excited by external signals through vertex disjoint paths, where \mathcal{D} contains the vertices that block the paths from the target inputs to the other inputs that have unknown edges to w_j . In contrast to Theorem 3, the conditions in Theorem 4 explicitly state that the signals in $\tilde{\mathcal{W}}_j \cup \mathcal{D}$ should be excited. When the target modules are not identifiable, the above result will be exploited in Section 6 to allocate additional excitation signals for achieving identifiability.

In Theorem 4, condition (a) more closely resembles Theorem 3 as it involves all external signals in \mathcal{X}_j ; on the other hand, the condition (b) considers any subset of signals in \mathcal{X}_j , and this ability of considering subsets of \mathcal{X}_j will be exploited for the design of synthesis approaches.

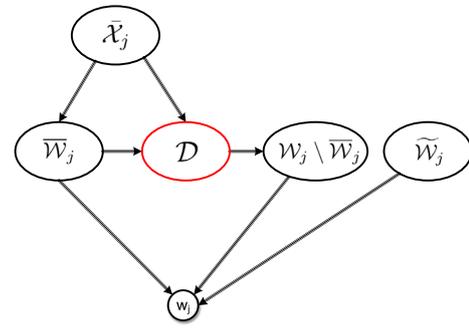


Fig. 3. Visualization of condition (b) in Theorem 4, where $\tilde{\mathcal{W}}_j$ collects the internal signals that have known edges to w_j .

5.2. Algebraic interpretation of disconnecting sets

A disconnecting set that disconnects external signals from internal signals in the graph of a network model set, gives rise to an attractive factorization of the external-to-internal mapping for the associated dynamic network.

Theorem 5. Consider a network model M with a disconnecting set $\mathcal{D} \subseteq \mathcal{V}$ from $\tilde{\mathcal{X}} \subseteq \mathcal{X}$ to $\tilde{\mathcal{W}} \subseteq \mathcal{W}$. Then there exists a proper transfer matrix $K(q)$ such that

$$T_{\tilde{\mathcal{W}}\tilde{\mathcal{X}}}(q) = K(q)T_{\mathcal{D}\tilde{\mathcal{X}}}(q). \quad (12)$$

The proof of the above result is presented in Appendix. The decomposition in (12) means that if all paths from $\tilde{\mathcal{X}}$ to $\tilde{\mathcal{W}}$ intersect with \mathcal{D} , then the signals in \mathcal{D} act as auxiliary signals that contain all information from the signals in $\tilde{\mathcal{X}}$ that is relevant for the signals in $\tilde{\mathcal{W}}$. The explicit form of the K matrix is specified in (26). The factorization in Theorem 5 will serve as a basis for formulating an indirect identification method for identifying $G_{j\tilde{\mathcal{W}}_j}$. This will be addressed in Section 7.

5.3. Relation with the parallel path and loop condition

In dynamic network identification, there is an important condition used to select measured internal signals for identifying a single module G_{ji} . The condition was first formulated in Dankers et al. (2016) as the parallel path and loop condition and selects a set of internal signals as follows. Determine a set of internal signals \mathcal{D} according to:

- Every parallel path from w_i to w_j , i.e. a path that is not the edge G_{ji} , passes through an internal vertex in \mathcal{D} , and
- Every loop through w_j passes through a vertex in \mathcal{D} .

If in a network the signals w_i , w_j and \mathcal{D} are retained, and all other internal signals are eliminated (immersed), e.g., because they cannot be measured, then in the immersed network that leaves the measured internal signals invariant, the target module G_{ji} remains invariant too. As a result, this condition has become an important tool for selecting internal signals to be measured for identification of a single module, see e.g. also Ramaswamy and Van den Hof (2021) and Weerts, Linder, Enqvist, and Van den Hof (2020). It can be shown that the above condition has a very strong link to disconnecting sets and Theorem 4.

Proposition 4. Consider a model set where all non-zero modules in $G(q, \theta)$ are parametrized. Consider the module G_{ji} and a set of internal signals \mathcal{D} with $\{w_i\} \notin \mathcal{D}$. Then \mathcal{D} is an $\{w_i\} - \mathcal{W}_j \setminus \{w_i\}$ disconnecting set if and only if \mathcal{D} contains an internal vertex of every parallel path from w_i to w_j and a vertex of every loop around w_j .

The proof of the above result is presented in [Appendix](#). While the parallel path and loop condition is used to select internal signals for identifying G_{ji} , that same condition now has the role of selecting internal signals to be externally excited for single module identifiability, according to the results of [Theorem 4](#). Moreover, a set of signals, which satisfies the parallel path and loop condition, can now effectively be found by constructing a (minimum) disconnecting set.

6. Signal allocation for generic identifiability

If $G_{j\bar{w}_j}$ is not generically identifiable in a given network model set, extra excitation signals can be allocated to achieve generic identifiability of $G_{j\bar{w}_j}$. In this section an algorithm will be developed to address this synthesis problem by exploiting condition (b) in [Theorem 4](#).

Consider a network model set with a set of initial external signals x_j^0 that do not have unknown directed edges to w_j . The algorithms aim to allocate a minimum number of additional excitation signals x_j^a such that generic identifiability of $G_{j\bar{w}_j}$ is guaranteed. In this synthesis problem, it is assumed that if r_k is allocated directly at w_j , its corresponding transfer function R_{jk} is known.

Condition (b) of [Theorem 4](#) is based on the existence of a $\bar{x}_j - w_j \setminus \bar{w}_j$ disconnecting set given \bar{x}_j ; however, the set \bar{x}_j depends on the allocated excitation signals and thus is not available before signal allocation. Therefore, in the synthesis approach we first find a candidate disconnecting set \mathcal{D} without relying on the excitation signals, and then construct \bar{x}_j by allocating signals such that (i) \mathcal{D} becomes a $\bar{x}_j - w_j \setminus \bar{w}_j$ disconnecting set and (ii) $\mathcal{D} \cup \bar{w}_j$ are excited by \bar{x}_j to satisfy (11). A candidate set \mathcal{D} can be motivated by the following corollary.

Corollary 1. *Assume that one of the conditions in [Theorem 4](#) is satisfied, then \mathcal{D} is a $\bar{x}_j - w_j \setminus \bar{w}_j$ disconnecting set if and only if it is a $\bar{w}_j \cup \bar{x}_j - w_j \setminus \bar{w}_j$ disconnecting set subject to $\mathcal{D} \cap \bar{w}_j = \emptyset$.*

Proof. The “if” part is trivial. The “only if” part is proved by contradiction. If \mathcal{D} is not a $\bar{w}_j - w_j \setminus \bar{w}_j$ disconnecting set, i.e. there exists a path from \bar{w}_j to $w_j \setminus \bar{w}_j$ which does not intersect with \mathcal{D} , then there is also a path from x_j (or \bar{x}_j) via \bar{w}_j to $w_j \setminus \bar{w}_j$ and the path does not intersect with \mathcal{D} , which contradicts that \mathcal{D} is a disconnecting set. In addition, following from (11), $\mathcal{D} \cap \bar{w}_j = \emptyset$ must hold. ■

The above result indicates that a $\bar{w}_j - w_j \setminus \bar{w}_j$ disconnecting set \mathcal{D} subject to $\mathcal{D} \cap \bar{w}_j = \emptyset$ can be computed first, which is independent of the allocated external signals. Then extra excitation signals can be allocated such that signals in $\mathcal{D} \cup \bar{w}_j$ are excited, and \mathcal{D} becomes a disconnecting set from the allocated signals to $w_j \setminus \bar{w}_j$.

As the number of required excitation signals depends on the cardinality of the disconnecting set, a minimum disconnecting set is desired. Additionally, based on [Corollary 1](#), a minimum disconnecting set \mathcal{D} subject to $\bar{w}_j \cap \mathcal{D} = \emptyset$ needs to be found. As standard graphical algorithms for computing minimum disconnecting sets do not take into account any constraint, we redefine the disconnecting set to make standard algorithms applicable.

Proposition 5. *Given a model set \mathcal{M} and any subset $\bar{x}_j \subseteq x_j$, \mathcal{D} is a minimum disconnecting set from $\bar{w}_j \cup \bar{x}_j$ to $w_j \setminus \bar{w}_j$ subject to $\bar{w}_j \cap \mathcal{D} = \emptyset$ if and only if \mathcal{D} is a minimum disconnecting set from $\mathcal{N}_{\bar{w}_j}^+ \cup \bar{x}_j$ to $w_j \setminus \bar{w}_j$.*

Proof. We will first show that for any vertex set \mathcal{D} subject to $\bar{w}_j \cap \mathcal{D} = \emptyset$, \mathcal{D} is a disconnecting set from $\mathcal{N}_{\bar{w}_j}^+ \cup \bar{x}_j$ to $w_j \setminus \bar{w}_j$ if

and only if it is also a disconnecting set from $\bar{w}_j \cup \bar{x}_j$ to $w_j \setminus \bar{w}_j$. The “only if” part holds because if \mathcal{D} intersects with all paths from $\mathcal{N}_{\bar{w}_j}^+ \cup \bar{x}_j$ to $w_j \setminus \bar{w}_j$, then it also intersects with the paths from $\bar{w}_j \cup \bar{x}_j$ to $w_j \setminus \bar{w}_j$. For the “if” part, since $\bar{w}_j \cap \mathcal{D} = \emptyset$ and \mathcal{D} intersects with all the paths from $\bar{w}_j \cup \bar{x}_j$ to $w_j \setminus \bar{w}_j$, those paths from \bar{w}_j to $w_j \setminus \bar{w}_j$ have to intersect with \mathcal{D} at their internal vertices or the ending vertices. Since the first internal vertices of the paths belong to set $\mathcal{N}_{\bar{w}_j}^+$, then \mathcal{D} is also a disconnecting set from $\mathcal{N}_{\bar{w}_j}^+ \cup \bar{x}_j$ to $w_j \setminus \bar{w}_j$.

Having the above result, the proposition is proved by showing that a minimum disconnecting set \mathcal{D} from $\mathcal{N}_{\bar{w}_j}^+ \cup \bar{x}_j$ to $w_j \setminus \{w_i\}$ does not contain \bar{w}_j , because if it does, it remains a disconnecting set after \bar{w}_j is excluded, which contradicts the minimality of \mathcal{D} . ■

The above result shows that a minimum disconnecting set \mathcal{D} from $\mathcal{N}_{\bar{w}_j}^+$ to $w_j \setminus \bar{w}_j$ can be computed for the synthesis problem, which now is an unconstrained problem and thus can be solved by standard graphic algorithms, e.g. the Ford–Fulkerson algorithm ([Schrijver, 2003](#)). Having established a choice of \mathcal{D} , the following synthesis result can be derived from condition (b) in [Theorem 4](#).

Corollary 2. *Consider a network model set \mathcal{M} that satisfies [Assumptions 2, 3](#) and [5](#). Given any minimum disconnecting set \mathcal{D} from $\mathcal{N}_{\bar{w}_j}^+$ to $w_j \setminus \bar{w}_j$, assigning distinct excitation signals to every vertex in $\mathcal{D} \cup \bar{w}_j$ leads to generic identifiability of $G_{j\bar{w}_j}$ in \mathcal{M} from (w, r) .*

Proof. Let x_j^a denote the set of allocated signals, and $x_j^a \subseteq x_j$ holds in the obtained model set after allocation, i.e. x_j^a has no unknown directed edge to w_j . As these signals are allocated directly at $\mathcal{D} \cup \bar{w}_j$, \mathcal{D} is a $\mathcal{N}_{\bar{w}_j}^+ \cup x_j^a - w_j \setminus \{w_i\}$ and thus a $\bar{w}_j \cup x_j^a - w_j \setminus \{w_i\}$ disconnecting set, based on [Proposition 5](#). In addition, (11) holds with $\bar{x}_j = x_j^a$. Thus, [Theorem 4\(b\)](#) is satisfied with $\bar{x}_j = x_j^a$ and the given \mathcal{D} . ■

The above result provides the basis for a synthesis approach that first computes a disconnecting set \mathcal{D} and then allocates external signals to directly excite $\mathcal{D} \cup \bar{w}_j$. This approach is valid even if a non-minimum disconnecting set is considered. However, it does not consider the initially present signals x_j^0 , and thus may allocate redundant signals. Additionally, excitation signals do not necessarily have to be directly allocated at the vertices in $\mathcal{D} \cup \bar{w}_j$, but could also reach those vertices through vertex disjoint paths. To make use of x_j^0 and to explore the freedom to allocate the signals, a more comprehensive method is introduced in [Algorithm 1](#).

Given a network model set with the target modules $G_{j\bar{w}_j}$ and the pre-existing external signals x_j^0 , [Algorithm 1](#) computes a minimum disconnecting set \mathcal{D} from $\mathcal{N}_{\bar{w}_j}^+ \cup x_j^0$ to $w_j \setminus \bar{w}_j$ first and then removes the vertices in $\mathcal{D} \cup \bar{w}_j$ that are already excited by x_j^0 via vertex disjoint paths. Then the algorithm allocates additional excitation signals to excite the remaining vertices in $\mathcal{D} \cup \bar{w}_j$ through vertex disjoint paths. The validity of [Algorithm 1](#) is shown in the following result.

Theorem 6. *Given a network model set \mathcal{M} that satisfies [Assumptions 2, 3, 5](#). In the returned model set of [Algorithm 1](#), $G_{j\bar{w}_j}$ is generically identifiable from (w, r) .*

Proof. From step 1 to 2 in [Algorithm 1](#), by construction, if $|\mathcal{P}| = |\mathcal{D} \cup \bar{w}_j|$, (11) holds with $\bar{x}_j = x_j^0$, and thus the modules are generically identifiable in the original model set \mathcal{M} . When

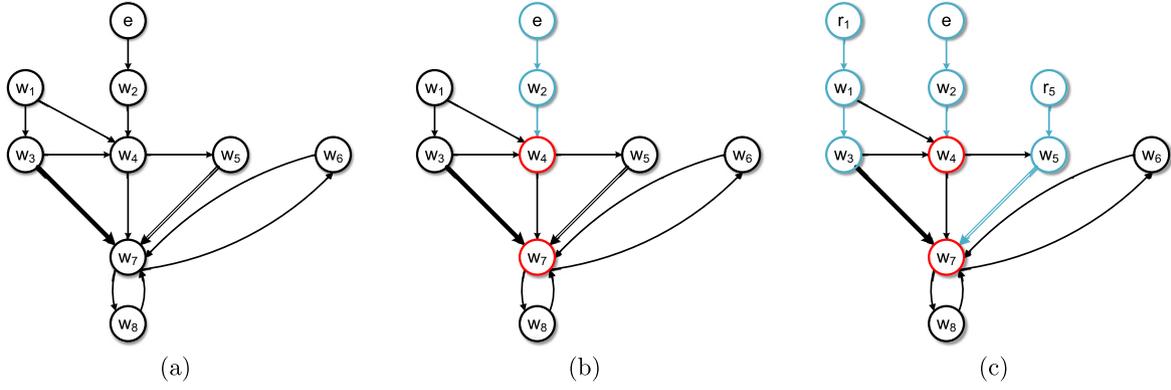


Fig. 4. An example of allocating signals for generic identifiability of G_{73} (thick line) using Algorithm 1 with a known module G_{75}^0 (double-line edge). Starting from the network model in (a), a disconnecting set (red vertices) is computed in (b). Since there already exists an external signal e , which has a path to w_4 that belongs to the disconnecting set, we only need to add r_1 and r_5 as in (c), which achieves generic identifiability of G_{73} . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Algorithm 1 Signal allocation for local modules

INPUT: A model set \mathcal{M} with graph \mathcal{G} , the target modules $G_{j\bar{\mathcal{W}}_j}$, and a set of initial external signals \mathcal{X}_j^0 ;

OUTPUT: A new model set \mathcal{M}_{out} with its graph \mathcal{G}_{out}

- 1: Compute a minimum disconnecting set \mathcal{D} from $\mathcal{N}_{\bar{\mathcal{W}}_j}^+ \cup \mathcal{X}_j^0$ to $\mathcal{W}_j \setminus \bar{\mathcal{W}}_j$;
- 2: Based on Lemma 7, compute a set \mathcal{P} that contains the maximum number of vertex disjoint paths from \mathcal{X}_j^0 to $\mathcal{D} \cup \bar{\mathcal{W}}_j$, while the paths are internally vertex disjoint with $\mathcal{D} \cup \bar{\mathcal{W}}_j$;
- 3: Let $\bar{\mathcal{D}} \subseteq \mathcal{D} \cup \bar{\mathcal{W}}_j$ denote all the ending vertices of the paths in \mathcal{P} ;
- 4: **if** $|\mathcal{P}| < |\mathcal{D} \cup \bar{\mathcal{W}}_j|$ **then**
- 5: Find the largest set $\bar{\mathcal{W}} \subseteq \mathcal{W}$ such that \mathcal{D} is a disconnecting set from $\bar{\mathcal{W}}$ to $\mathcal{W}_j \setminus \bar{\mathcal{W}}_j$;
- 6: Build a subgraph $\bar{\mathcal{G}} \subseteq \mathcal{G}$ by removing all vertices and edges of the paths in \mathcal{P} ;
- 7: Find a set $\mathcal{W}_{exp} \subseteq \bar{\mathcal{W}}$ such that in $\bar{\mathcal{G}}$, there are $|\mathcal{D} \cup \bar{\mathcal{W}}_j| - |\mathcal{P}|$ vertex disjoint paths from \mathcal{W}_{exp} to $(\mathcal{D} \cup \bar{\mathcal{W}}_j) \setminus \bar{\mathcal{D}}$;
- 8: In \mathcal{G} , assign distinct excitation signals to every vertex in \mathcal{W}_{exp} , which leads to a new model set \mathcal{M}_{out} with a new graph \mathcal{G}_{out} and a new R matrix;
- 9: Return \mathcal{M}_{out} with the graph \mathcal{G}_{out} ;
- 10: **else**
- 11: $\mathcal{M}_{out} \leftarrow \mathcal{M}$ and $\mathcal{G}_{out} \leftarrow \mathcal{G}$;
- 12: Return \mathcal{M}_{out} with the graph \mathcal{G}_{out} .
- 13: **end if**

$|\mathcal{P}| < |\mathcal{D} \cup \bar{\mathcal{W}}_j|$, based on Theorem 4, we need to allocate extra $|\mathcal{D} \cup \bar{\mathcal{W}}_j| - |\mathcal{P}|$ signals, such that: (i) there are $|\mathcal{D} \cup \bar{\mathcal{W}}_j| - |\mathcal{P}|$ vertex disjoint paths from these signals to $(\mathcal{D} \cup \bar{\mathcal{W}}_j) \setminus \bar{\mathcal{D}}$, and these paths are also vertex disjoint with \mathcal{P} ; (ii) \mathcal{D} remains a disconnecting set from the added signals to $\mathcal{W}_j \setminus \bar{\mathcal{W}}_j$. Then we can find in the algorithm, steps 5 guarantees (ii), and steps 3, 6, 7 and 8 guarantee (i). This concludes the proof. ■

Remark 1. In Algorithm 1, the signals do not need to be allocated directly at the vertices in $\mathcal{D} \cup \bar{\mathcal{W}}_j$. It suffices when they reach these vertices through vertex disjoint paths from excitation locations elsewhere in the network.

Algorithm 1 and the synthesis approach in Corollary 2 guarantee a minimum number of allocated signals when $\mathcal{X}_j^0 = \emptyset$, as a minimum disconnecting set is used. While a minimum number

of allocated signals cannot be guaranteed when $\mathcal{X}_j^0 \neq \emptyset$, the following bound on the number of additionally required signals can be derived.

Corollary 3. Given a network model set \mathcal{M} that satisfies Assumptions 2, 3, 5. Let \mathcal{X}_j^0 denote the set of initial external signals that have no unknown directed edge to w_j . Let \mathcal{D} be a minimum disconnecting set from $\mathcal{N}_{\bar{\mathcal{W}}_j}^+ \cup \mathcal{X}_j^0$ to $\mathcal{W}_j \setminus \bar{\mathcal{W}}_j$, and let c denote the number of excitation signals available for allocation. Then the number of additional excitation signals c is sufficient to make $G_{j\bar{\mathcal{W}}_j}$ generically identifiable if $c \geq |\mathcal{D} \cup \bar{\mathcal{W}}_j| - b_{\mathcal{X}_j^0 \rightarrow \mathcal{D} \cup \bar{\mathcal{W}}_j}$.

Algorithm 1 is illustrated in the following example.

Example 5. In Fig. 4(a), the problem is to allocate excitation signals such that G_{73} becomes generically identifiable. $\mathcal{X}_7^0 = \{e\}$ is the only external signal that is initially present. Firstly, a disconnecting set from $\mathcal{X}_7^0 \cup \{w_3\} = \{w_3, e\}$ to the other in-neighbors of w_7 through unknown edges, i.e. $\mathcal{W}_7 \setminus \{w_3\} = \{w_4, w_6, w_8\}$, is constructed as $\mathcal{D} = \{w_4, w_7\}$, indicated in Fig. 4(b). Based on Theorem 4, generic identifiability of G_{73} requires three vertex disjoint paths from external signals to $\mathcal{D} \cup \{w_3\} = \{w_3, w_4, w_7\}$, while \mathcal{D} remains a disconnecting set from the external signals to $\mathcal{W}_7 \setminus \{w_3\}$. Following step 2 in Algorithm 1, we find a path $e \rightarrow w_4$ from \mathcal{X}_7^0 to $\mathcal{D} \cup \{w_3\}$ (colored blue in Fig. 4(b)). Thus we only need to allocate extra excitation signals which have two vertex disjoint paths to $\{w_3, w_7\}$, and the two paths should be vertex disjoint with path $e \rightarrow w_4$. As in step 5, the set of potential locations to allocate excitation signals is given by $\bar{\mathcal{W}} = \{w_1, w_2, w_3, w_4, w_5, w_7\}$, which satisfies the requirement that \mathcal{D} remains a disconnecting set from $\bar{\mathcal{W}}$ to $\mathcal{W}_7 \setminus \{w_3\}$. Then we choose $\mathcal{W}_{exp} = \{w_1, w_5\} \subseteq \bar{\mathcal{W}}$ to be excited by r_1 and r_5 , as shown in Fig. 4(c). Based on Theorem 4, G_{73} is indeed generically identifiable in Fig. 4(c): There exists a disconnecting set $\mathcal{D} = \{w_4, w_7\}$ from $\mathcal{X}_7 = \{r_1, r_5, e\}$ to $\mathcal{W}_7 \setminus \{w_3\} = \{w_4, w_6, w_8\}$ such that $b_{\mathcal{X}_7 \rightarrow \mathcal{D} \cup \{w_3\}} = 3$, i.e., the blue paths in Fig. 4(c).

When generic identifiability of $\{G_{j\bar{\mathcal{W}}_j} | j \in \{1, \dots, L\}\}$ is of interest, Algorithm 1 can be applied recursively for each j . If generic identifiability of the full network is concerned, a compact graphical approach for signal allocation is presented in Cheng, Shi, and Van den Hof (2019) and Cheng et al. (2022), where effective use is made of the interdependence of the L different subproblems.

7. Indirect identification method

In the previous sections, graphical tools have been developed to verify identifiability and achieve identifiability via signal allocation. However, the identifiability results do not point directly to an identification method to actually estimate the modules of interest from data. In this section we will show that, when the identifiability conditions in [Theorem 4](#) are satisfied, in a particular situation the identifiability results directly point to an identification method for estimating $G_{j\bar{\mathcal{W}}_j}$. The particular situation is that the identifiability conditions are satisfied on the basis of a set of external signals $\bar{\mathcal{X}}_j$ that contains r signals only. We will first show how the target modules can be reconstructed on the basis of transfer functions between measured signals only, by combining the identifiability result and [Theorem 5](#). For simplicity, we consider the setting where all the known transfer operators in the network are fixed zeros.

Corollary 4. Consider a model set \mathcal{M} that has all the known transfer operators being zeros and satisfies [Assumptions 2, 3](#). If condition (b) of [Theorem 4](#) is satisfied with $\bar{\mathcal{X}}_j$ having no directed edge to w_j , then it holds that

$$G_{j\bar{\mathcal{W}}_j}(q, \theta) = T_{j\bar{\mathcal{X}}_j}(q, \theta) \begin{bmatrix} T_{\bar{\mathcal{W}}_j\bar{\mathcal{X}}_j}(q, \theta) \\ T_{\mathcal{D}\bar{\mathcal{X}}_j}(q, \theta) \end{bmatrix}^\dagger \begin{bmatrix} I_{|\bar{\mathcal{W}}_j|} \\ \mathbf{0} \end{bmatrix}, \quad (13)$$

where $(\cdot)^\dagger$ denotes the right inverse and exists for almost all θ , and $I_{|\bar{\mathcal{W}}_j|}$ denotes an identity matrix of size $|\bar{\mathcal{W}}_j|$.

Proof. When all the known transfer operators in \mathcal{M} are zeros, [Assumption 5](#) is not needed, and \mathcal{W}_j contains all the internal signals having directed edges to w_j . If $\bar{\mathcal{X}}_j$ has no edge to w_j , the j th row in $(I - G)T_{\mathcal{W}\mathcal{X}} = X$ can be reformulated as

$$\begin{bmatrix} G_{j\bar{\mathcal{W}}_j} & G_{j(\mathcal{W}_j \setminus \bar{\mathcal{W}}_j)} \end{bmatrix} \begin{bmatrix} T_{\bar{\mathcal{W}}_j\bar{\mathcal{X}}_j} \\ T_{(\mathcal{W}_j \setminus \bar{\mathcal{W}}_j)\bar{\mathcal{X}}_j} \end{bmatrix} = T_{j\bar{\mathcal{X}}_j}. \quad (14)$$

If condition (b) of [Theorem 4](#) holds, K exists so that $T_{(\mathcal{W}_j \setminus \bar{\mathcal{W}}_j)\bar{\mathcal{X}}_j} = KT_{\mathcal{D}\bar{\mathcal{X}}_j}$ according to [Theorem 5](#). Combining the above decomposition and (14) leads to (13), where the right inverse exists generically due to (11) and (8). ■

The expression (13) shows an immediate opportunity to estimate $G_{j\bar{\mathcal{W}}_j}$ based on a selected set of measured internal signals. When $\bar{\mathcal{X}}_j$ contains only measured excitation signals r , the transfer functions on the right hand side of (13) can all be estimated consistently through a standard identification method, based on the measured signals in $\bar{\mathcal{X}}_j \cup \bar{\mathcal{W}}_j \cup \mathcal{D} \cup \{w_j\}$.

The above method is a generalization of the ‘‘classical’’ indirect method of identification, as developed for closed-loop systems ([Ljung, 1987](#)) and dynamic networks ([Gevers et al., 2018](#)), where all the in-neighbors of w_j are measured and directly excited for identifying the target modules. The method based on [Corollary 4](#) generalizes this situation by allowing to measure and excite the internal signals determined by disconnecting sets. This generalization can lead to an experimental setup with fewer actuators and sensors for identifying the target modules than required by the classical method.

Remark 2. When $\bar{\mathcal{X}}_j$ only contains r signals, it is shown in this section that $G_{j\bar{\mathcal{W}}_j}$ can be estimated consistently based on the signals in $\mathcal{D} \cup \bar{\mathcal{W}}_j \cup \{w_j\}$. It is then straightforward that under these conditions $G_{j\bar{\mathcal{W}}_j}$ is generically identifiable in an appropriately chosen model set, without requiring to measure all the internal signals. This establishes a particular generic identifiability result with unmeasured internal signals, as also addressed in [Bazanella et al. \(2019\)](#).

If the set $\bar{\mathcal{X}}_j$ contains both r and e -signals, then the reasoning underlying the indirect approach fails, and one can apply a direct method for consistently identifying the target module, see [Ramaswamy and Van den Hof \(2021\)](#).

8. Duality

It is not hard to extend the results of this work from the full measurement setting to the full excitation setting in [Bazanella et al. \(2017\)](#) and [Hendrickx et al. \(2019\)](#), where all internal signals are excited with a subset unmeasured, and only r signals are used as excitation sources for identifiability. In this case, the goal is to ensure the uniqueness of a subset of out-going edges of w_i , given the mapping from the r signals to the measured internal signals. Then the path-based condition as in [Theorem V.I of Hendrickx et al. \(2019\)](#) can be obtained, and a disconnecting-set-based result can be obtained similarly.

However, exploiting the noise signals as excitation sources for identifiability, as done in [Proposition 1](#), is not trivial when there are unmeasured internal signals. This requires to handle the noise information in a submatrix of the power spectrum $\Phi_{\bar{v}}$ of the disturbance, instead of the full matrix $\Phi_{\bar{v}}$ as in this work. This more complex situation is addressed in the follow up work ([Shi et al., 2021b](#)).

9. Conclusion

In this work, generic identifiability of a subnetwork, i.e. a subset of modules, in linear dynamic networks has been investigated in the setting where all internal signals are measured. Algebraic conditions are obtained for both generic and global identifiability. Then a path-based condition for generic identifiability is obtained by extending the results in [Hendrickx et al. \(2019\)](#), in order to handle the model-set type of identifiability, while including disturbance inputs in the network and allowing network modules to be fixed/known. This path-based condition is then equivalently reformulated into a novel graphical condition based on disconnecting sets. With this new condition, the synthesis problem is tackled, which aims to allocate the minimum number of excitation signals to achieve generic identifiability of the modules. In addition, the new condition allows us to make a further step towards a new indirect identification method for estimating the subnetwork. The next step in this research is the handling of the more complex situation of partial measurement and partial excitation. This topic is addressed in the follow up work ([Shi et al., 2021b](#)). The Matlab software for conducting identifiability analysis based on [Theorem 3](#) and synthesis according to [Algorithm 1](#) can be found in [Shi, Cheng, and Van den Hof \(2021a\)](#).

Appendix

A.1. Proof for [Proposition 3](#)

The following result is instrumental in the proof.

Lemma 4. Consider a model set \mathcal{M} that satisfies [Assumptions 3 and 5](#). Then for any subsets $\bar{\mathcal{W}} \subseteq \mathcal{W}$ and $\bar{\mathcal{X}} \subseteq \mathcal{X}$, $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}$ has full rank for almost all $\theta \in \Theta$ if and only if $F(\bar{\mathcal{W}}, \bar{\mathcal{X}})$, as defined in (6), has structural full rank.

The above result suggests that testing the generic rank of $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}$ only needs to test the sparsity pattern of $F(\bar{\mathcal{W}}, \bar{\mathcal{X}})$. Then due to the connection between graph \mathcal{G} and F 's sparsity pattern, [Lemma 4](#) allows us to obtain [Proposition 3](#). To prove [Lemma 4](#), two preliminary results on structural rank are presented. Consider a block

matrix whose entries are either zeros or distinct indeterminates:

$$M = \begin{bmatrix} A - I & B \\ C & D \end{bmatrix}, \tag{15}$$

where A is hollow, A and D are of dimensions $L \times L$ and $m \times m$, respectively. Here L or m is allowed to be zero.

Lemma 5. *If M in (15) is not structural full rank, then M can be permuted as*

$$\begin{bmatrix} \bar{A} & 0 \\ \bar{C} & \bar{D} \end{bmatrix}, \tag{16}$$

where \bar{A} has dimension $k_1 \times (k_1 - 1)$ for some $k_1 \geq 1$.

Proof. The non-zero structure of M can be characterized by a graph $\bar{\mathcal{G}} := (\bar{\mathcal{V}}, \bar{\mathcal{E}})$ with $\bar{\mathcal{V}} := \bar{\mathcal{W}} \cup \bar{\mathcal{X}} \cup \bar{\mathcal{Y}}$, where $\mathcal{W} = \{w_1, \dots, w_L\}$, $\bar{\mathcal{X}} = \{x_1, \dots, x_m\}$, and $\mathcal{Y} = \{y_1, \dots, y_m\}$ correspond to the rows/columns of A , the columns of B and the rows of C , respectively. Besides, a directed edge $(j, i) \in \bar{\mathcal{E}} := \bar{\mathcal{V}} \times \bar{\mathcal{V}}$ if and only if M_{ij} is non-zero. When M is not structural full rank, it follows from van der Woude (1991) that $b_{\bar{\mathcal{X}} \rightarrow \bar{\mathcal{W}}} < |\bar{\mathcal{X}}| = m$ in $\bar{\mathcal{G}}$. Then, from Theorem 1, there exists a $\bar{\mathcal{X}} - \bar{\mathcal{W}}$ disconnecting set \mathcal{D} with $|\mathcal{D}| = m - 1$ in $\bar{\mathcal{G}}$. Note that with \mathcal{D} and based on Lemma 1, we can divide $\bar{\mathcal{V}}$ into three disjoint sets \mathcal{D} , \mathcal{S} and \mathcal{P} with $|\mathcal{S}| + |\mathcal{P}| = L + m + 1$. Moreover, there is no edge from \mathcal{S} to \mathcal{P} , where $\mathcal{S} \subseteq \bar{\mathcal{W}} \cup \bar{\mathcal{X}}$, $\mathcal{P} \subseteq \bar{\mathcal{W}} \cup \bar{\mathcal{Y}}$. Thus, from the definition of $\bar{\mathcal{E}}$, we can find a permutation of M in the form of (16) with a zero block, whose rows and columns correspond to \mathcal{P} and \mathcal{S} , respectively. Furthermore, the column dimension of \bar{A} is computed as $L + m - |\mathcal{S}| = |\mathcal{P}| - 1$, which completes the proof. ■

Lemma 6. *Let the non-zero entries of M in (15) be divided into two disjoint sets \mathcal{P}_1 and \mathcal{P}_2 . If $\det(M)$ does not depend on the entries in \mathcal{P}_1 , then $\det(M) = (-1)^j \prod_{i=1}^l \det(A_i)$, for some $l \geq 1, j \in \{0, 1\}$, where A_i is a square submatrix of M that contains only non-zero entries in \mathcal{P}_2 .*

Proof. When $\mathcal{P}_1 = \emptyset$, the proof is trivial. Now suppose $\mathcal{P}_1 \neq \emptyset$, and consider the cofactor expansion formula of $\det(M)$, which contains a term $b \det(\bar{B})$ with $\det(\bar{B})$ the cofactor of the nonzero entry $b \in \mathcal{P}_1$. Since $\det(M)$ does not depend on b , we obtain $\det(\bar{B}) = 0$, i.e., \bar{B} is not structural full rank. Then it follows from Lemma 5 that M can be permuted as

$$\begin{bmatrix} \star & \begin{bmatrix} \bar{B}_{11} & 0 \\ \bar{B}_{21} & \bar{B}_{22} \end{bmatrix} \\ b & \star \end{bmatrix} := \begin{bmatrix} \bar{A} & 0 \\ \bar{C} & \bar{D} \end{bmatrix},$$

with square matrices \bar{A} and \bar{D} , and \bar{C} containing b . Then, we have $\det(M) = \pm \det(\bar{A}) \det(\bar{D})$, which does not depend on b . Here, the sign \pm depends on the permutation. If \bar{A} or \bar{D} contains other entries in \mathcal{P}_1 , the above analysis can be applied recursively, which proves the lemma. ■

With the previous lemmas, Lemma 4 can be achieved.

Proof. Without loss of generality, consider $|\bar{\mathcal{W}}| \leq |\bar{\mathcal{X}}|$, and thus $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}$ having full generic rank implies that its square submatrix $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}_1}$ has full rank generically. Consider \bar{F} defined as (7) for $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}_1}$, and denote $|\bar{\mathcal{X}}_1| = m$. As \bar{F} is structural full rank if and only if $F(\bar{\mathcal{W}}, \bar{\mathcal{X}}_1)$ (and thus $F(\bar{\mathcal{W}}, \bar{\mathcal{X}})$) has structural full rank, and \bar{F} has generically full rank if and only if $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}_1}$ has generically full rank. Thus, the proof aims to prove that \bar{F} is structural full rank if and only if it is generically full rank. Let a_{ij} denote the (i, j)

entry of \bar{F} . The ‘‘if’’ part is clear as being structural full rank is a necessary condition for being full rank generically. For the ‘‘only if’’ part, consider now all the summands in the Leibnitz formula for $\det(\bar{F})$. If \bar{F} is structural full rank, without loss of generality, $\det(\bar{F})$ contains a term according to a permutations $\bar{\sigma}$ of $[1, \dots, L + m]$ as

$$a \prod_{i=p+1}^{L+m} a_{\bar{\sigma}(i)i} \neq 0, \tag{17}$$

for some $p \geq 0$, where $\sigma(i)$ denotes the i th index in the permutation, and a is the product of all parametrized entries in the term and also contains the maximum number of parametrized entries among the other permutations of $[1, \dots, L + m]$. In addition, the summation of all terms with the common factor a in $\det(\bar{F})$ equals

$$a \det(\bar{F}_{\{\bar{\sigma}(p+1), \dots, \bar{\sigma}(L+m)\}\{p+1, \dots, L+m\}}). \tag{18}$$

Denote the above submatrix of \bar{F} as F_1 . Then \bar{F} is proved to be generically full rank by showing that the term (18) is non-zero and cannot be canceled by the other terms in $\det(\bar{F})$ generically as follows. As (17) contains the maximum number of parametrized entries compared to the other permutations, $\det(F_1)$ does not depend on the parametrized entries when considered as a polynomial of F_1 's all non-zero entries. Based on Lemma 6 and the fact that C is a selection matrix, (18) can be reformulated as

$$(-1)^j a \prod_{i=1}^l \det(A_i), \tag{19}$$

for some $l \geq 1$ and $j \in \{0, 1\}$, where A_i contains only known transfer functions and is square submatrix of F_1 and thus of $F(\bar{\mathcal{W}}, \bar{\mathcal{X}}_1)$ defined in (6). As (19) contains the non-zero term (17) after expansion, A_i is structural full rank and thus full rank by Assumption 5, for all i . This means that $\det(F_1)$ is non-zero and thus $\det(\bar{F})$ contains at least the non-zero term (18). As each parametrized entry has independent parameters by Assumption 3, the non-zero term (18) cannot be canceled in $\det(\bar{F})$ by the other terms without the factor a . Thus, $\det(\bar{F})$ is a non-constant analytic function of θ . Then following the property of analytic functions and the analysis in Lemma V.2 of Hendrickx et al. (2019), \bar{F} and thus $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}$ have full rank generically. ■

Finally, Proposition 3 can be proved based on Lemma 4.

Proof. The proof is analogous to the proof of Proposition V.1 in Hendrickx et al. (2019): For generically $\text{rank}(T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}) \geq b_{\bar{\mathcal{X}} \rightarrow \bar{\mathcal{W}}}$, consider a subgraph of the network containing all the vertices but only the edges in a set of maximum number vertex disjoint paths from $\bar{\mathcal{X}}$ to $\bar{\mathcal{W}}$. Let $\bar{\mathcal{X}}_1 \subseteq \bar{\mathcal{X}}$ and $\bar{\mathcal{W}}_1 \subseteq \bar{\mathcal{W}}$ denote the starting and ending vertices of the vertex disjoint paths, respectively. The obtained subgraph's structure can then be encoded by matrices A and B with only zeros and ones, where $A_{ji} = 1$ and $B_{kn} = 1$ if and only if G_{ji} and X_{kn} denote the edges in the subgraph, respectively. Then following the same analysis of Proposition V.1 in Hendrickx et al. (2019), we can show that $C(I - A)^{-1}B_{\bar{\mathcal{W}}_1\bar{\mathcal{X}}_1}$ is a permutation matrix and thus has full rank that equals $b_{\bar{\mathcal{X}} \rightarrow \bar{\mathcal{W}}}$, where C is a selection matrix that extracts rows of $(I - A)^{-1}$ corresponding to $\bar{\mathcal{X}}_1$. Then following Lemma 3 similarly, it holds that $[(A - I)_{\bar{\mathcal{W}}(\bar{\mathcal{W}} \setminus \bar{\mathcal{W}}_1)} \quad B_{\bar{\mathcal{W}}_1\bar{\mathcal{X}}_1}]$ is full rank, and thus $F(\bar{\mathcal{W}}_1, \bar{\mathcal{X}}_1)$ defined in (6) is structural full rank. Then based on Lemma 4, $T_{\bar{\mathcal{W}}_1\bar{\mathcal{X}}_1}$ is generically full rank that equals $b_{\bar{\mathcal{X}} \rightarrow \bar{\mathcal{W}}}$, and thus $\text{rank}(T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}) \geq b_{\bar{\mathcal{X}} \rightarrow \bar{\mathcal{W}}}$ generically.

For $\text{rank}(T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}) \leq b_{\bar{\mathcal{X}} \rightarrow \bar{\mathcal{W}}}$ generically, a minimum $\bar{\mathcal{X}} - \bar{\mathcal{W}}$ disconnecting set can be considered as in Hendrickx et al. (2019), which leads to permuted network matrices with block zeros as also explored later in (21). Then the proof can be achieved

similarly as in Hendrickx et al. (2019). Note that Hendrickx et al. (2019) require its Lemma V.2 to ensure the invertibility of certain submatrix $I - G_{\mathcal{V}_1 \mathcal{V}_1}$, which is guaranteed by Assumption 1(b) in this work and thus the lemma is not needed. ■

A.2. Proof of Theorem 4

Theorem 4 is established by two graphical results.

Lemma 7. *In a simple directed graph, given a set \mathcal{P} of vertex disjoint paths from vertex set \mathcal{V}_1 to a vertex set \mathcal{V}_2 , there exists a set \mathcal{P}_{new} of vertex disjoint paths from \mathcal{V}_1 to \mathcal{V}_2 such that $|\mathcal{P}_{new}| = |\mathcal{P}|$ and paths in \mathcal{P}_{new} are internally vertex disjoint⁴ with $\mathcal{V}_1 \cup \mathcal{V}_2$.*

Proof. We prove the lemma by showing that there always exists a \mathcal{P}_{new} by modifying the paths in \mathcal{P} . Let $w_i \rightarrow w_j$ be an arbitrary path in \mathcal{P} which contains an internal vertex in \mathcal{V} , then we can always replace $w_i \rightarrow w_j$ by its subpath which contains a starting vertex in \mathcal{V}_1 and an ending vertex in \mathcal{V}_2 , while the other vertices in the subpath are not in \mathcal{V} . This includes the special case that the obtained subpath has no internal vertex. Applying the above modification to all the paths in \mathcal{P} , which contain internal vertices in \mathcal{V} , leads to \mathcal{P}_{new} . ■

Based on the above result, there always exists a set of the maximum number of vertex disjoint paths from \mathcal{V}_1 and \mathcal{V}_2 , which are internally vertex disjoint with $\mathcal{V}_1 \cup \mathcal{V}_2$. We denote such a set by $\mathcal{B}_{\mathcal{V}_1 \rightarrow \mathcal{V}_2}$. With this notation, an instrumental graphical result can be obtained.

Lemma 8. *Consider a simple directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with any two vertex sets $\mathcal{V}_1, \mathcal{V}_2 \subseteq \mathcal{V}$ and a subset $\bar{\mathcal{V}}_2 \subseteq \mathcal{V}_2$. The following statements are equivalent:*

- (1) $b_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2} = |\bar{\mathcal{V}}_2|$ and $b_{\mathcal{V}_1 \rightarrow \mathcal{V}_2} = b_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2} + b_{\mathcal{V}_1 \rightarrow \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2}$;
- (2) there exists a $\mathcal{V}_1 - \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2$ disconnecting set \mathcal{D} such that

$$b_{\mathcal{V}_1 \rightarrow \mathcal{D} \cup \bar{\mathcal{V}}_2} = |\mathcal{D}| + |\bar{\mathcal{V}}_2|; \quad (20)$$
- (3) there exist a $\bar{\mathcal{V}}_1 \subseteq \mathcal{V}_1$ and a $\bar{\mathcal{V}}_1 - \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2$ disconnecting set \mathcal{D} such that $b_{\bar{\mathcal{V}}_1 \rightarrow \bar{\mathcal{V}}_2 \cup \mathcal{D}} = |\mathcal{D}| + |\bar{\mathcal{V}}_2|$.

Proof. We first prove that (1) holds if and only if (2) holds. If (2) holds, (20) shows that there exist $|\mathcal{D}|$ and $|\bar{\mathcal{V}}_2|$ vertex disjoint paths from \mathcal{V}_1 to \mathcal{D} and from \mathcal{V}_1 to $\bar{\mathcal{V}}_2$, respectively. Then according to Lemma 7, there exist $\mathcal{B}_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2}$ with cardinality $|\bar{\mathcal{V}}_2|$ and $\mathcal{B}_{\mathcal{V}_1 \rightarrow \mathcal{D}}$ with cardinality of $|\mathcal{D}|$, and the paths in the above two sets are vertex disjoint. A new set $\mathcal{B}_{\mathcal{D} \rightarrow \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2}$ can also be introduced and is vertex disjoint with both $\mathcal{B}_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2}$ and $\mathcal{B}_{\mathcal{V}_1 \rightarrow \mathcal{D}}$, because if not vertex disjoint, a path from \mathcal{V}_1 to $\mathcal{V}_2 \setminus \bar{\mathcal{V}}_2$ will exist and do not intersect with \mathcal{D} , contradicting \mathcal{D} as a disconnecting set. Then by linking a subset of paths in $\mathcal{B}_{\mathcal{V}_1 \rightarrow \mathcal{D}}$ and all paths in $\mathcal{B}_{\mathcal{D} \rightarrow \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2}$, we can obtain a set of vertex disjoint paths from \mathcal{V}_1 to $\mathcal{V}_2 \setminus \bar{\mathcal{V}}_2$. Since \mathcal{D} is a $\mathcal{V}_1 - \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2$ disconnecting set and $|\mathcal{B}_{\mathcal{V}_1 \rightarrow \mathcal{D}}| = |\mathcal{D}|$, the obtained set of paths forms a set $\mathcal{B}_{\mathcal{V}_1 \rightarrow \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2}$ with the maximum number of vertex disjoint paths, and its cardinality equals $|\mathcal{B}_{\mathcal{D} \rightarrow \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2}|$. In addition, the obtained $\mathcal{B}_{\mathcal{V}_1 \rightarrow \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2}$ is also vertex disjoint with $\mathcal{B}_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2}$. Since $b_{\mathcal{V}_1 \rightarrow \mathcal{V}_2} \leq b_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2} + b_{\mathcal{V}_1 \rightarrow \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2}$ always holds, $\mathcal{B}_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2}$ with cardinality $|\bar{\mathcal{V}}_2|$ and $\mathcal{B}_{\mathcal{V}_1 \rightarrow \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2}$ together forms a set of maximum number of vertex disjoint paths from \mathcal{V}_1 to \mathcal{V}_2 , which proves (1).

⁴ A path is internally vertex disjoint with a set of vertices \mathcal{V} , if the internal vertices of the path are not in \mathcal{V} .

If (1) holds, let \mathcal{D} be a minimum $\mathcal{V}_1 - \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2$ disconnecting set, and we have $\mathcal{B}_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2}$ with cardinality $|\bar{\mathcal{V}}_2|$ and $\mathcal{B}_{\mathcal{V}_1 \rightarrow \mathcal{D}}$ with cardinality $|\mathcal{D}|$ which are vertex disjoint. Thus the above two sets together form a set of vertex disjoint paths from \mathcal{V}_1 to $\bar{\mathcal{V}}_2 \cup \mathcal{D}$, which leads to $b_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2 \cup \mathcal{D}} \geq |\mathcal{D}| + |\bar{\mathcal{V}}_2|$. As $b_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2 \cup \mathcal{D}}$ is upper bounded by $|\mathcal{D}| + |\bar{\mathcal{V}}_2|$, it then holds that $b_{\mathcal{V}_1 \rightarrow \bar{\mathcal{V}}_2 \cup \mathcal{D}} = |\mathcal{D}| + |\bar{\mathcal{V}}_2|$, which leads to (1) \iff (2).

Then we show that (2) is equivalent to (3). The implication (2) \implies (3) is straightforward by letting $\bar{\mathcal{V}}_1 = \mathcal{V}_1$. For (3) \implies (2), if (3) holds, then $\mathcal{D}_1 = \mathcal{D} \cup (\mathcal{V}_1 \setminus \bar{\mathcal{V}}_1)$ becomes a $\mathcal{V}_1 - \mathcal{V}_2 \setminus \bar{\mathcal{V}}_2$ disconnecting set, and $b_{\mathcal{V}_1 \rightarrow \mathcal{D}_1 \cup \bar{\mathcal{V}}_2} = |\mathcal{D}_1| + |\bar{\mathcal{V}}_2|$ since a single vertex can be regarded to have a path to itself, which concludes the proof. ■

Finally, Lemma 8 and Theorem 3 lead to Theorem 4.

A.3. Proof of Proposition 4

Before proving the proposition, we first prove that there exists a directed path from w_i to $\mathcal{W}_j \setminus \{w_i\}$ if and only if there exist a parallel path from w_i to w_j or a cycle around the output w_j . Note that due to model (1), \mathcal{G} is a simple graph, i.e. there is no self-loop such as (w_i, w_i) , and no parallel directed edges from one vertex to another vertex. For “if” part, if there exists a parallel path from w_i to w_j , this parallel path has to intersect with $\mathcal{W}_j \setminus \{w_i\}$. Then we can find a directed path from w_i to one vertex in $\mathcal{W}_j \setminus \{w_i\}$ as a subpath of the parallel path. If a cycle around w_j exists, it will also intersect with $\mathcal{W}_j \setminus \{w_i\}$, and thus the cycle contains a subpath from w_j to one vertex in $\mathcal{W}_j \setminus \{w_i\}$. Linking this subpath and the edge (w_i, w_j) leads to a path from w_i to $\mathcal{W}_j \setminus \{w_i\}$.

For “only if” part, for any directed path from w_i to $w_k \in \mathcal{W}_j \setminus \{w_i\}$, if the path does not contain edge (w_i, w_j) , then combining the path and the edge (w_k, w_j) will create a parallel path. If the path contains (w_i, w_j) , then combining the path and the edge (w_k, w_j) while excluding (w_i, w_j) will lead to a cycle around w_j . This concludes the relationship between the parallel paths, the cycles around the output and the paths from w_i to $\mathcal{W}_j \setminus \{w_i\}$.

Then based on the above result, the “only if” of the proposition is straightforward. For the “if” part, if we collect an internal vertex from each parallel path and a vertex from cycle around the output into \mathcal{D} , \mathcal{D} then must disconnect from w_i to $\mathcal{W}_j \setminus \{w_i\}$.

A.4. Proof of Theorem 5

According to Lemma 1, the disconnecting set separates all the vertices in the graph into three disjoint sets as $\mathcal{V} = \mathcal{S} \cup \mathcal{D} \cup \mathcal{P}$, while there is no directed edge from \mathcal{S} to \mathcal{P} . In addition, each set may contain both internal signals and external signals, i.e. $\mathcal{S} = \mathcal{S}_x \cup \mathcal{S}_w$, $\mathcal{D} = \mathcal{D}_x \cup \mathcal{D}_w$ and $\mathcal{P} = \mathcal{P}_x \cup \mathcal{P}_w$, and it holds that $\bar{\mathcal{X}} \subseteq \mathcal{S}_x \cup \mathcal{D}_x$ and $\bar{\mathcal{W}} \subseteq \mathcal{D}_w \cup \mathcal{P}_w$. Algebraically, the above statements mean that G can be permuted and partitioned as

$$G = \begin{bmatrix} G_{\mathcal{S}_w \mathcal{S}_w} & G_{\mathcal{S}_w \mathcal{D}_w} & G_{\mathcal{S}_w \mathcal{P}_w} \\ G_{\mathcal{D}_w \mathcal{S}_w} & G_{\mathcal{D}_w \mathcal{D}_w} & G_{\mathcal{D}_w \mathcal{P}_w} \\ 0 & G_{\mathcal{P}_w \mathcal{D}_w} & G_{\mathcal{P}_w \mathcal{P}_w} \end{bmatrix}. \quad (21)$$

X and $T_{\mathcal{W}\bar{\mathcal{X}}}$ can be permuted and partitioned similarly with block $X_{\mathcal{P}_w \mathcal{S}_x} = 0$, where, for example, $T_{\mathcal{S}_w \mathcal{D}_x}$ denotes the mapping from \mathcal{D}_x to \mathcal{S}_w . Based on the above partition and the equation $(I - G)T_{\mathcal{W}\bar{\mathcal{X}}} = X$, our goal is to find a proper matrix K such that $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}} = KT_{\mathcal{D}\bar{\mathcal{X}}}$.

Firstly, considering the division of the sets \mathcal{D} and $\bar{\mathcal{X}}$, the mapping $T_{\mathcal{D}\bar{\mathcal{X}}}$ can be re-written as

$$T_{\mathcal{D}\bar{\mathcal{X}}} = \begin{bmatrix} T_{\mathcal{D}_w\bar{\mathcal{X}}_S} & T_{\mathcal{D}_w\bar{\mathcal{X}}_D} \\ \mathbf{0} & I \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (22)$$

where $\bar{\mathcal{X}}_S = \bar{\mathcal{X}} \cap \mathcal{S}_X$ and $\bar{\mathcal{X}}_D = \bar{\mathcal{W}} \cap \mathcal{D}_X$, the identity matrix is the mapping $T_{\bar{\mathcal{X}}_D\bar{\mathcal{X}}_D}$. Note that the rows of the bottom block matrices in (22) correspond to the vertices in $\mathcal{D}_X \setminus \bar{\mathcal{X}}_D$. In addition, $T_{\bar{\mathcal{W}}\bar{\mathcal{X}}}$ can be written as

$$T_{\bar{\mathcal{W}}\bar{\mathcal{X}}} = \begin{bmatrix} T_{\bar{\mathcal{W}}_p\bar{\mathcal{X}}} \\ T_{\bar{\mathcal{W}}_D\bar{\mathcal{X}}} \end{bmatrix}, \quad (23)$$

where $\bar{\mathcal{W}}_p = \bar{\mathcal{W}} \cap \mathcal{P}_w$ and $\bar{\mathcal{W}}_D = \bar{\mathcal{W}} \cap \mathcal{D}_w$. Thus, it is clear that

$$T_{\bar{\mathcal{W}}_D\bar{\mathcal{X}}} = C \begin{bmatrix} T_{\mathcal{D}_w\bar{\mathcal{X}}_S} & T_{\mathcal{D}_w\bar{\mathcal{X}}_D} \end{bmatrix}, \quad (24)$$

where C is a selection matrix that extracts the rows of $\begin{bmatrix} T_{\mathcal{D}_w\bar{\mathcal{X}}_S} & T_{\mathcal{D}_w\bar{\mathcal{X}}_D} \end{bmatrix}$ corresponding to $\bar{\mathcal{W}}_D$. In addition, from the permuted matrices and the equation $(I - G)T_{\mathcal{W}\mathcal{X}} = X$, it holds that $T_{\mathcal{P}_w\mathcal{S}_X} = (I - G_{\mathcal{P}_w\mathcal{P}_w})^{-1}G_{\mathcal{P}_w\mathcal{D}_w}T_{\mathcal{D}_w\mathcal{S}_X}$, where $I - G_{\mathcal{P}_w\mathcal{P}_w}(z)$ is invertible and inversely proper because $I - G_{\mathcal{P}_w\mathcal{P}_w}(z)$ is proper and the network is well-posed, i.e. $\lim_{z \rightarrow \infty} \det(I - G_{\mathcal{P}_w\mathcal{P}_w}(z)) \neq 0$. The above equation leads to

$$T_{\bar{\mathcal{W}}_p\bar{\mathcal{X}}_S} = \bar{K}T_{\mathcal{D}_w\bar{\mathcal{X}}_S}, \quad (25)$$

where $\bar{K} = [(I - G_{\mathcal{P}_w\mathcal{P}_w})^{-1}]_{\bar{\mathcal{W}}_p\star}G_{\mathcal{P}_w\mathcal{D}_w}$. Then combining the above equation with (22), (23) and (24) leads to

$$T_{\bar{\mathcal{W}}\bar{\mathcal{X}}} = \begin{bmatrix} T_{\bar{\mathcal{W}}_p\bar{\mathcal{X}}_S} & T_{\bar{\mathcal{W}}_p\bar{\mathcal{X}}_D} \\ T_{\bar{\mathcal{W}}_D\bar{\mathcal{X}}_S} & T_{\bar{\mathcal{W}}_D\bar{\mathcal{X}}_D} \end{bmatrix} = \begin{bmatrix} \bar{K} & T_{\bar{\mathcal{W}}_p\bar{\mathcal{X}}_D} - \bar{K}T_{\mathcal{D}_w\bar{\mathcal{X}}_D} \\ C & \mathbf{0} \end{bmatrix} \\ \times \begin{bmatrix} T_{\mathcal{D}_w\bar{\mathcal{X}}_S} & T_{\mathcal{D}_w\bar{\mathcal{X}}_D} \\ \mathbf{0} & I \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = KT_{\mathcal{D}\bar{\mathcal{X}}}.$$

The formulation of the above K matrix can be further simplified. Based on the permuted matrices and the equation $(I - G)T_{\mathcal{W}\mathcal{X}} = X$, it holds $(I - G_{\mathcal{P}_w\mathcal{P}_w})T_{\mathcal{P}_w\mathcal{D}_X} - G_{\mathcal{P}_w\mathcal{D}_w}T_{\mathcal{D}_w\mathcal{D}_X} = X_{\mathcal{P}_w\mathcal{D}_X}$. Thus, we can conclude that

$$K = \begin{bmatrix} \bar{K} & [(I - G_{\mathcal{P}_w\mathcal{P}_w})^{-1}]_{\bar{\mathcal{W}}_p\star}X_{\mathcal{P}_w\bar{\mathcal{X}}_D} \\ C & \mathbf{0} \end{bmatrix}, \quad (26)$$

where \bar{K} is defined in (25); C is defined in (24) and its rows correspond to $\bar{\mathcal{W}}_D$; the columns of the last block column in (26) correspond to $\mathcal{D}_X \setminus \bar{\mathcal{X}}_D$. Note that certain blocks in K may disappear depending on if the corresponding set of signals is empty. ■

References

- Adebayo, J., Southwick, T., Chetty, V., Yeung, E., Yuan, Y., Gonçalves, J., et al. (2012). Dynamical structure function identifiability conditions enabling signal structure reconstruction. In *Proc. 51st IEEE Conf. Decis Control* (pp. 4635–4641). IEEE.
- Bazanella, A. S., Gevers, M., & Hendrickx, J. M. (2019). Network identification with partial excitation and measurement. In *Proc. 58th IEEE Conf. Decis Control* (pp. 5500–5506). IEEE.
- Bazanella, A. S., Gevers, M., Hendrickx, J. M., & Parraga, A. (2017). Identifiability of dynamical networks: which nodes need be measured? In *Proc. 56th IEEE Conf. Decis Control CDC*, (pp. 5870–5875). IEEE.
- Cheng, X., & Scherpen, J. M. A. (2021). Model reduction methods for complex network systems. *Annual Review of Control, Robotics, and Autonomous Systems*, 4, 425–453.
- Cheng, X., Shi, S., & Van den Hof, P. M. J. (2019). Allocation of excitation signals for generic identifiability of dynamic networks. In *Proc. 58th IEEE Conf. Decis Control* (pp. 5507–5512). IEEE.
- Cheng, X., Shi, S., & Van den Hof, P. M. J. (2022). Allocation of excitation signals for generic identifiability of linear dynamic networks. *IEEE Transactions on Automatic Control*, 67(2), To appear.

- Chiuseo, A., & Pilonetto, G. (2012). A Bayesian approach to sparse dynamic network identification. *Automatica*, 48(8), 1553–1565.
- Dankers, A. G. (2014). *System identification in dynamic networks*. (Ph.D. Dissertation), Delft University of Technology.
- Dankers, A. G., Van den Hof, P. M. J., Bombois, X., & Heuberger, P. S. C. (2016). Identification of dynamic models in complex networks with prediction error methods: Predictor input selection. *IEEE Transactions on Automatic Control*, 61(4), 937–952.
- Everitt, N., Bottegal, G., & Hjalmarsson, H. (2018). An empirical Bayes approach to identification of modules in dynamic networks. *Automatica*, 91, 144–151.
- Fonken, S., Ferizbegovic, M., & Hjalmarsson, H. (2020). Consistent identification of dynamic networks subject to white noise using weighted null-space fitting. *IFAC-PapersOnLine*, 53(2), 46–51, Proc. 21th IFAC World Congress.
- Gevers, M., & Bazanella, A. S. (2015). Identification in dynamic networks: identifiability and experiment design issues. In *Proc. 54th IEEE Conf. Decis Control CDC*, (pp. 4005–4010). IEEE.
- Gevers, M., Bazanella, A. S., & da Silva, G. V. (2018). A practical method for the consistent identification of a module in a dynamical network. *IFAC-PapersOnLine*, 51(15), 862–867.
- Gonçalves, J., & Warnick, S. (2008). Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks. *IEEE Transactions on Automatic Control*, 53(7), 1670–1674.
- Haber, A., & Verhaegen, M. (2014). Subspace identification of large-scale interconnected systems. *IEEE Transactions on Automatic Control*, 59(10), 2754–2759.
- Hayden, D., Chang, Y. H., Gonçalves, J., & Tomlin, C. J. (2016). Sparse network identifiability via compressed sensing. *Automatica*, 68, 9–17.
- Hendrickx, J. M., Gevers, M., & Bazanella, A. S. (2019). Identifiability of dynamical networks with partial node measurements. *IEEE Transactions on Automatic Control*, 64(6), 2240–2253.
- Ishizaki, T., Chakraborty, A., & Imura, J. (2018). Graph-theoretic analysis of power systems. *Proceedings of the IEEE*, 106(5), 931–952.
- Linder, J., & Enqvist, M. (2017). Identification of systems with unknown inputs using indirect input measurements. *International Journal of Control*, 90(4), 729–745.
- Ljung, L. (1987). *System identification: Theory for the user*. Prentice-hall.
- Materassi, D., & Salapaka, M. V. (2012). On the problem of reconstructing an unknown topology via locality properties of the Wiener filter. *IEEE Transactions on Automatic Control*, 57(7), 1765–1777.
- Materassi, D., & Salapaka, M. V. (2020). Signal selection for estimation and identification in networks of dynamic systems: A graphical model approach. *IEEE Transactions on Automatic Control*, 65(10), 4138–4153.
- Mesbahi, M., & Egerstedt, M. (2010). *Graph theoretic methods in multiagent networks*. Princeton University Press.
- Ramaswamy, K. R., & Van den Hof, P. M. J. (2021). A local direct method for module identification in dynamic networks with correlated noise. *IEEE Transactions on Automatic Control*, 66(11), 3237–3252.
- Sanandaji, B. M., Vincent, T. L., & Wakin, M. B. (2011). Exact topology identification of large-scale interconnected dynamical systems from compressive observations. In *Proc. Am Control Conf ACC*, (pp. 649–656). IEEE.
- Schrijver, A. (2003). Disjoint paths. In *Combinatorial optimization: Polyhedral and efficiency* (pp. 131–132). Springer Science & Business Media.
- Shi, S. (2021). *Topological aspects of linear dynamic networks: Identifiability and identification*. (Ph.D. Dissertation). Eindhoven Univ. Technology.
- Shi, S., Bottegal, G., & Van den Hof, P. M. J. (2019). Bayesian topology identification of linear dynamic networks. In *2019 18th European control conference, (ECC)* (p. 2814–2819).
- Shi, S., Cheng, X., & Van den Hof, P. M. J. (2021a). Identifiability with full measurement. Gitlab Repository, Gitlab, <https://gitlab.tue.nl/CS/SYSDYNET/Identifiability>.
- Shi, S., Cheng, X., & Van den Hof, P. M. J. (2021b). Single module identifiability in linear dynamic networks with partial excitation and measurement. *IEEE Trans. Autom. Control*, provisionally accepted. arXiv:2012.11414 [cs.LG].
- Shi, S., Cheng, X., & Van den Hof, P. M. J. (2021c). Supplementary material for “generic identifiability of subnetworks in a linear dynamic network: the full measurement case”. <https://research.tue.nl/en/publications/supplementary-material-for-generic-identifiability-of-subnetworks>, Eindhoven Univ. Technology.
- Shi, S., Cheng, X., & Van den Hof, P. M. J. (2020). Excitation allocation for generic identifiability of a single module in dynamic networks: A graphic approach. *IFAC-PapersOnLine*, 53(2), 40–45, Proc. 21th IFAC World Congress.
- Steffen, T. (2005). *Control reconfiguration of dynamical systems: Linear approaches and structural tests*, vol. 320. Springer Science & Business Media.
- Van den Hof, P. M. J., Dankers, A. G., Heuberger, P. S. C., & Bombois, X. (2013). Identification of dynamic models in complex networks with prediction error methods basic methods for consistent module estimates. *Automatica*, 49(10), 2994–3006.

- van der Woude, J. W. (1991). A graph-theoretic characterization for the rank of the transfer matrix of a structured system. *Mathematics of Control, Signals, and Systems*, 4(1), 33–40.
- van Waarde, H. J., Tesi, P., & Camlibel, M. K. (2020). Necessary and sufficient topological conditions for identifiability of dynamical networks. *IEEE Transactions on Automatic Control*, 65(11), 4525–4537.
- Weerts, H. H. M., Dankers, A. G., & Van den Hof, P. M. J. (2015). Identifiability in dynamic network identification. *IFAC-PapersOnLine*, 48–28, 1409–1414, Proc. 17th IFAC Symp. System Identif..
- Weerts, H. H. M., Linder, J., Enqvist, M., & Van den Hof, P. M. J. (2020). Abstractions of linear dynamic networks for input selection in local module identification. *Automatica*, 117(108975).
- Weerts, H. H. M., Van den Hof, P. M. J., & Dankers, A. G. (2018a). Identifiability of linear dynamic networks. *Automatica*, 89, 247–258.
- Weerts, H. H. M., Van den Hof, P. M. J., & Dankers, A. G. (2018b). Prediction error identification of linear dynamic networks with rank-reduced noise. *Automatica*, 98, 256–268.
- Weerts, H. H. M., Van den Hof, P. M. J., & Dankers, A. G. (2018c). Single module identifiability in linear dynamic networks. In *Proc. 57th IEEE Conf. Decis Control* (pp. 4725–4730). IEEE.
- Yu, C., & Verhaegen, M. (2018). Subspace identification of individual systems operating in a network (SI²ON). *IEEE Transactions on Automatic Control*, 63(4), 1120–1125.
- Zorzi, M., & Chiuso, A. (2017). Sparse plus low rank network identification: a nonparametric approach. *Automatica*, 76, 355–366.



Shengling Shi is a postdoctoral researcher at the Delft Center for Systems and Control, Delft University of Technology, the Netherlands. He received the B.Sc. degree from the Harbin Institute of Technology, China, in 2013 and the M.Sc. degree with distinction (great appreciation) from the Eindhoven University of Technology (TU/e), the Netherlands, in 2017. He received the Ph.D. degree from the Control Systems Group of TU/e in 2021. His research interests include system identification and model predictive control of networked systems.



Xiaodong Cheng is currently a Research Associate with the Control Group in the Department of Engineering at the University of Cambridge. He received his B.Sc. and M.Sc. degrees from the Northwestern Polytechnical University, Xi'an, China, in 2011 and 2014, respectively, and the Ph.D. degree with honours (cum laude) from the University of Groningen, The Netherlands, in November 2018. Then, he was appointed as a post-doctoral researcher with the Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands, from 2019 to 2020. His main research interests include system identification, model reduction, control in large-scale networked systems. He is the recipient of the Paper Prize Award from the IFAC Journal *Automatica* in the triennium 2017–2019 and the Outstanding Paper Award from IEEE Transactions on Control Systems Technology in 2020.



Paul M.J. Van den Hof received the M.Sc. and Ph.D. degrees in electrical engineering from Eindhoven University of Technology, Eindhoven, The Netherlands, in 1982 and 1989, respectively.

In 1986 he moved to Delft University of Technology, where he was appointed as Full Professor in 1999. From 2003 to 2011, he was founding co-director of the Delft Center for Systems and Control (DCSC). As of 2011, he is a Full Professor in the Electrical Engineering Department, Eindhoven University of Technology. His research interests include data-driven modeling, identification for control, dynamic network identification, and model-based control and optimization, with applications in industrial process control systems and high-tech systems. He holds an ERC Advanced Research grant for a research project on identification in dynamic networks. Paul Van den Hof is an IFAC Fellow and IEEE Fellow, Honorary Member of the Hungarian Academy of Sciences, and IFAC Advisor. He has been a member of the IFAC Council (1999–2005, 2017–2020), the Board of Governors of IEEE Control Systems Society (2003–2005), and an Associate Editor and Editor of *Automatica* (1992–2005). In the triennium 2017–2020 he served as Vice-President of IFAC.