Automatica 191 (2026) 113110

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Data-informativity conditions for structured linear systems with implications for dynamic networks☆

Paul M.J. Van den Hof [a,*], Shengling Shi [b], Stefanie J.M. Fonken [c], Karthik R. Ramaswamy [d], Håkan Hjalmarsson [e], Arne G. Dankers [f]

[a] Control Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands
[b] Delft Center for Systems and Control, Delft University of Technology, The Netherlands
[c] Unit Water and Energy Transition, Flemish Institute for Technological Research (VITO), Mol, Belgium, and EnergyVille, Genk, Belgium
[d] ASML, Veldhoven, The Netherlands
[e] Department of Decision and Control Systems, Royal Institute of Technology, Sweden
[f] Department of Electrical and Software Engineering, University of Calgary, Canada

ARTICLE INFO

Article history:
Received 20 November 2024
Received in revised form 12 January 2026
Accepted 17 April 2026
Available online 10 June 2026

Keywords:
System identification
Identifiability
Dynamic networks
Interconnected systems
Multivariable systems

ABSTRACT

When estimating a single subsystem (module) in a linear dynamic network with a prediction error method, a data-informativity condition needs to be satisfied for arriving at a consistent module estimate. This concerns a condition on input signals in the constructed, possibly MIMO (multiple input multiple output) predictor model being persistently exciting, which is typically guaranteed if the input spectrum is positive definite for a sufficient number of frequencies. Generically, the condition can be formulated as a path-based condition on the graph of the network model. The current condition has two elements of possible conservatism: (a) rather than focussing on the full MIMO model, one would like to be able to focus on consistently estimating the target module only, and (b) structural information, such as structural zero elements in the interconnection structure or known subsystems, should be taken into account. In this paper relaxed conditions for data-informativity are derived addressing these two issues, leading to relaxed path-based conditions on the network graph. This leads to experimental conditions that are less strict, i.e. require a smaller number of external excitation signals. Additionally, the new expressions for data-informativity in identification are shown to be closely related to earlier derived conditions for (generic) single module identifiability.

© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

While the dynamic systems that are currently addressed in engineering and science, become more and more complex, large-scale and interconnected, the attention for modeling, monitoring and control aspects of interconnected systems, like e.g. dynamic networks, is growing. For the resulting control problems this has led to the development of decentralized, distributed and multi-agent type of control systems. On the modeling side, the development of analysis tools for systems over graphs has to be mentioned, and we have seen the development of a field called data-driven modeling (or system identification) in dynamic networks. This paper is situated in this latter area, that will be overviewed next.

Identification of/in dynamic networks has been developed in a framework where dynamic systems are interconnected in a particular interconnection structure (topology), including cycles, that may be known a priori. In such a setting several different problems have been considered. Following the early contributions on uniqueness of network representations (Gonçalves & Warnick, 2008) and topology estimation (Materassi & Innocenti, 2010; Materassi & Salapaka, 2012), a framework for network identification in a prediction error setting is introduced in Van den Hof et al. (2013), while subspace methods for network identification have been developed in Haber and Verhaegen (2014), Verhaegen et al. (2022). Alternatively, there are network identification approaches based on conditional independence and dynamic factor models, see e.g. Zorzi and Chiuso (2017), Veedu et al. (2022).

☆ This work was funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the European Research Council can be held responsible for them. This work was also supported by VINNOVA Competence Center AdBIOPRO 2016-05181, and by the Swedish Research Council through contracts [2019-04956, 2025-04833]. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Mattia Zorzi under the direction of Editor Alessandro Chiuso.

* Corresponding author.
E-mail addresses: p.m.j.vandenhof@tue.nl (P.M.J. Van den Hof), shengling.shi@tudelft.nl (S. Shi), stefanie.fonken@vito.be (S.J.M. Fonken), k.r.ramaswamy89@gmail.com (K.R. Ramaswamy), hjalmars@kth.se (H. Hjalmarsson), arne.dankers2@ucalgary.ca (A.G. Dankers).

https://doi.org/10.1016/j.automatica.2026.113110
0005-1098/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

The vast literature on network identification problems in a prediction error setting, ranges from the problem of identifying the full network for a given topology (Yuan et al., 2011; Weerts et al., 2018b; Fonken et al., 2022), and the analysis and synthesis of identifiability conditions (Weerts et al., 2018a; Hendrickx et al., 2019; van Waarde et al., 2020; Legat & Hendrickx, 2020; Cheng et al., 2022; Shi et al., 2023; Mapurunga et al., 2024), to the problem of identifying the dynamics of a single module/link/subsystem in the network (Van den Hof et al., 2013; Dankers et al., 2016; Linder & Enqvist, 2017; Gevers et al., 2018; Everitt et al., 2018; Materassi & Salapaka, 2020; Ramaswamy & Van den Hof, 2021), which includes the problem of determining where to measure and where to excite the network, and the related data-informativity questions (Bombois et al., 2023).

In this paper we will focus on this latter single module identification problem, in a setting where some dynamic modules in the network may be known a priori, while a single target module will need to be estimated consistently. We focus on the so-called *local direct method* (Ramaswamy & Van den Hof, 2021) that is known to be able to exploit both measured external excitation signals and unmeasured disturbance signals for estimation. It has been shown (Ramaswamy & Van den Hof, 2021) that it can be necessary to build a predictor model that has multiple inputs and multiple outputs (MIMO), in order to arrive at consistent estimates of a single target module. This is due to the necessity to deal with *confounding variables*, i.e., correlated disturbances between inputs and outputs of an estimation problem, a situation that can typically occur in dynamic networks. The resulting MIMO predictor model then will need to satisfy *data-informativity* conditions in order to arrive at consistent module estimates. The objective of this paper is to formulate *path-based conditions*, dependent on the a priori known graph of the network only, that guarantee data-informativity, and therefore will provide the conditions for consistent module estimates. These path-based conditions will refer to the locations of external signals (measured excitation signals and unmeasured disturbance signals) that enter the network.

For specifying the role of a data-informativity condition, we need to go back to their origin, developed for classical (open-loop type of) identification problems in a prediction error setting. It is most often formulated as a condition on the input spectrum matrix being positive definite at a sufficient number of frequencies (Söderström & Stoica, 1989). The condition is typically applied to situations where the inputs are general quasi-stationary signals with continuous spectral densities. For other types of inputs, as e.g. periodic signals, the concerned condition is conservative and replaced by conditions on the selection of a sufficient number of frequency components for each input and orthogonality of the signals over the different inputs (Pintelon & Schoukens, 2012). For particular black box model structures with finite order model sets, the spectrum condition can be further detailed, as shown e.g., in Gevers et al. (2008). The available theory for data informativity analysis seems to be mainly directed towards the situation of unstructured MIMO black box models, meaning that all entries in the MIMO model are parametrized. A relevant question that pops up is then: what are the consequences for the conditions for data-informativity if particular entries in the MIMO model are structurally zero, or non-zero but a priori known and therefore not necessarily parametrized? This latter situation is typically present in a dynamic network situation.

In the current literature on dynamic network identification, the data-informativity conditions so far have been taken from a general non-structured MIMO estimation problem, i.e., requiring $\Phi_{w_\mathcal{D}}(\omega) \succ 0$ for a sufficient number of frequencies ω, with $w_\mathcal{D}$ the set of inputs nodes in the predictor model. For a MISO setting, the data-informativity conditions have also been considered in Bombois et al. (2023) in the form of rank conditions on particular system-dependent matrices. For a generic satisfaction of the data-informativity conditions, very effective use has been made of a result of van der Woude (1991) and Hendrickx et al. (2019) where it is shown that rank conditions on transfer functions can generically be verified by path-based conditions on the graph of the network. This has led to the integration of path-based conditions in the selection of inputs and outputs of the predictor model to warrant that data-informativity conditions can be satisfied generically (Van den Hof & Ramaswamy, 2020; Van den Hof et al., 2023).

In all these approaches, the data-informativity conditions for black box, fully parametrized MIMO models have been exploited. However, in the considered network problem, structural information on the multivarible predictor model might be readily available (e.g. its topology, indicating which nodes are connected with which other nodes), and rather than consistently estimating the full MIMO predictor model, we can focus on consistently estimating the single target module only.

The question that we address in this paper is therefore: in the setting of the mentioned local direct method (Ramaswamy & Van den Hof, 2021), can we exploit structural information (topology) of the network, in the form of a priori known entries in the predictor model, to relax the data-informativity conditions and to focus these conditions on estimating a single target module only. In comparison with the current data-informativity literature, these are two innovative steps to be taken: including the structural information, and addressing a single target module only.

We will analyze the situation by first addressing the structural information in a standard (open-loop type of) estimation problem in Sections 3–5, including a presentation of the modeling setup and the notion of data-informativity. After presenting an example in Section 6, we will then formulate the new results in terms of path-based conditions on the network graph in Section 7. With these new results, we can address the single module identification problem in dynamic networks in Section 8, with some examples collected in Section 9. The proofs of the most important results are collected in an Appendix. We start by presenting a very brief example.

2. A motivating example

In order to give a flavor of the type of results that will be achieved in this paper, we present an example network provided in Fig. 1, where linear time-invariant systems G_{jk} are interconnected in a network structure, and the node variables w_ℓ are the output signals of the corresponding summation points. v_k are non-measured stationary stochastic disturbance signals, v_1 and v_2 being mutually correlated, and x_i are possible locations for adding external signals. Our objective is to identify the target module G_{12} from data w_ℓ, $\ell = 1, \ldots, 6$, and possibly present measured x_i. In the current theory of the local direct method, a consistent estimation of G_{12}, can be achieved if the four external signals x_1, x_2, x_3 and x_5 are all present, with x_1 and x_2 being measured. Our new results, derived in this paper, will show that the presence of signals x_3 and x_5 only, with x_5 being measured, will be sufficient, thereby considerably reducing the experimental burden. This example will be treated in full detail in Section 9.2.

3. The modeling setup

For properly addressing the data-informativity problem, we start by considering a multivariable (open-loop) data-generating dynamic system described by

$$y(t) = G^0(q)u(t) + H^0(q)e(t) \tag{1}$$

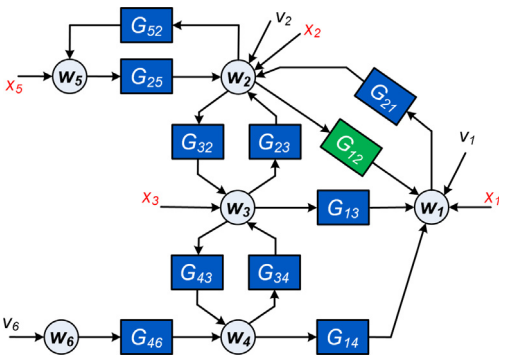


Fig. 1. Six node example with target module $G_{12}(q)$ and correlated disturbances v_1, v_2.

where output $y(t) \in \mathbb{R}^p$, input $u(t) \in \mathbb{R}^m$, the rational transfer function $G^0(\cdot) \in \mathbb{R}^{p\times m}(\cdot)$, q the shift operator $qu(t) = u(t+1)$, $e(t) \in \mathbb{R}^p$ a multivariate white-noise process, and $H^0(\cdot) \in \mathbb{R}^{p\times p}(\cdot)$ a rational noise model, being monic, i.e. $\lim_{z\to\infty} H(z) = I$, stable and stably invertible. For simplicity we will assume that $G^0(q)$ is stable too.

The data-generating system is modeled through a parametrized model $(G(q,\theta), H(q,\theta))$, with parameters $\theta \in \Theta \subset \mathbb{R}^n$. The model's one-step ahead predictor can be defined as $\hat{y}(t|t-1;\theta) := \bar{\mathbb{E}}[y(t)|y^{t-1}, u^t]$ where u^t refers to the past of signal u up to time instant t, and $\bar{\mathbb{E}}$ is the generalized expectation operator $\bar{\mathbb{E}} := \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} \mathbb{E}$ for quasi-stationary signals (Ljung, 1999).

This leads to the classical expression for the one-step-ahead predictor:

$$\hat{y}(t|t-1;\theta) = \underbrace{\begin{bmatrix}(I - H(q,\theta)^{-1}) & H(q,\theta)^{-1}G(q,\theta)\end{bmatrix}}_{W(q,\theta)} \underbrace{\begin{bmatrix}y(t)\\ u(t)\end{bmatrix}}_{z(t)} \tag{2}$$

with predictor filter $W(q,\theta)$.

In line with the corresponding definitions in the prediction error literature (Ljung, 1999, Definition 8.1), we can now define the notion of data-informativity for the related predictor model.

Definition 1. Consider a set of signals contained in z and a predictor model $\hat{y}(t|t-1;\theta) = W(q,\theta)z(t)$, as in (2) for a parametrized set of models $\mathcal{M} := (G(q,\theta), H(q,\theta))_{\theta\in\Theta}$. Then a quasi-stationary data set $Z^\infty := \{z(t)\}_{t=0,\dots,\infty}$ is *informative enough with respect to the model set* $\mathcal{M}$ if, for any two predictor models $W(q,\theta_1)$ and $W(q,\theta_2)$ in the model set, the equation

$$\bar{\mathbb{E}}[\|(W(q,\theta_1) - W(q,\theta_2))z(t)\|^2] = 0 \tag{3}$$

implies that $W(e^{i\omega},\theta_1) \equiv W_2(e^{i\omega},\theta_2)$ for almost all ω, where $\|\cdot\|$ is the standard Euclidean norm. □

Data-informativity guarantees that one can distinguish between different models in a model set, and consequently, that one can identify unique models on the basis of an experimental data set of infinite length. It is a necessary condition for the consistent estimation of models. If we do not constrain the order of the models in the model set $\mathcal{M}$, but allow $\mathcal{M}$ to consist of all linear time-invariant models of any order, then a typical sufficient condition for data-informativity is the condition that $\Phi_z(\omega) \succ 0$ for almost all ω. For open-loop situations where the signals u and e are uncorrelated, this can be further relaxed to the condition

$$\Phi_u(\omega) \succ 0 \quad \text{for almost all } \omega, \tag{4}$$

being referred to as the property of input signal u being persistently exciting (Ljung, 1999).

4. Data-informativity conditions for consistent estimation

The data-informativity conditions play a central role in the proof of consistency of estimates of $G^0(q)$ and $H^0(q)$ when applying an (asymptotic) prediction error identification criterion

$$\theta^* = \arg\min_\theta \bar{\mathbb{E}}\varepsilon^T(t,\theta)Q\varepsilon(t,\theta) \tag{5}$$

with the one-step-ahead prediction error

$$\begin{aligned}\varepsilon(t,\theta) &:= y(t) - \hat{y}(t|t-1;\theta),\\ &= H(q,\theta)^{-1}[y(t) - G(q,\theta)u(t)],\end{aligned} \tag{6}$$

and Q any positive definite weighting matrix. Without loss of generality we will use $Q = I$ in the sequel, while for the consistency analysis we will assume that $(G^0, H^0) \in \mathcal{M}$.

Substituting (1) into (6) leads to

$$\begin{aligned}\varepsilon(t,\theta) = H(q,\theta)^{-1}[G^0(q) - G(q,\theta)]u(t) +\\ + H(q,\theta)^{-1}[H^0(q) - H(q,\theta)]e(t) + e(t)\end{aligned} \tag{7}$$

which can alternatively be written as

$$\varepsilon(t,\theta) = H(q,\theta)^{-1}\Delta M(q,\theta)\kappa(t) + e(t), \tag{8}$$

with $\Delta M(q,\theta) := [\Delta G(q,\theta) \;\; \Delta H(q,\theta)]$, $\Delta G(q,\theta) := G^0(q) - G(q,\theta)$, $\Delta H(q,\theta) := H^0(q) - H(q,\theta)$ and $\kappa(t) := [u^T(t) \;\; e^T(t)]^T$. The proof of consistency (Ljung, 1978) comes down to showing that

a. $\Delta M(q,\theta) = 0$ achieves the global minimum of the cost function in (5), and
b. this minimum is unique, i.e. it provides a single point estimate $(G(q,\theta^*), H(q,\theta^*))$.

If the transfer function $\Delta H(q,\theta)$ is strictly proper, which is guaranteed by the assumption that $H^0(q)$ and $H(q,\theta)$ both are monic, then it is straightforward to show that

$$\theta^* = \arg\min_\theta \bar{\mathbb{E}}\|H(q,\theta)^{-1}\Delta M(q,\theta)\kappa(t)\|^2. \tag{9}$$

With this expression it is obvious that condition (a) above is satisfied, reaching the minimum value of 0 of the cost function in (9) for $\Delta M = 0$. For the uniqueness condition (b), we need to formulate a condition under which the following implication holds:

$$\{\bar{\mathbb{E}}\|H(q,\theta)^{-1}\Delta M(q,\theta)\kappa(t)\|^2 = 0\} \Longrightarrow \{\Delta M(q,\theta) = 0\}.$$

Since $H(q,\theta)$ is stable and stably invertible, for any quasi-stationary signal $x(t)$ it holds that

$$\{\bar{\mathbb{E}}\|H(q,\theta)^{-1}x(t)\|^2 = 0\} \Longleftrightarrow \{\bar{\mathbb{E}}\|x(t)\|^2 = 0\}$$

so that we can rewrite the required implication for condition (b) as

$$\{\bar{\mathbb{E}}\|\Delta M(q,\theta)\kappa(t)\|^2 = 0\} \Longrightarrow \{\Delta M(q,\theta) = 0\}. \tag{10}$$

Using Parseval's relation the left hand side can be written as:

$$\frac{1}{2\pi}\text{trace}\{\int_{-\pi}^{\pi} \Delta M(e^{i\omega},\theta)\Phi_\kappa(\omega)\Delta M^T(e^{-i\omega},\theta)d\omega\} = 0 \tag{11}$$

and this expression implies $\Delta M(q,\theta) = 0$ if $\Phi_\kappa(\omega) \succ 0$ in a sufficient number of frequencies. If no conditions are imposed on the order of $M(q,\theta)$ then the typically used condition is to require $\Phi_\kappa(\omega) \succ 0$ at almost all ω. In the open-loop case where u and e are uncorrelated, this reduces to the persistent excitation condition (4).

5. Data informativity conditions for structured systems

5.1. Data-informativity for a single row/output

We will now consider the situation that the multivariable system G^0 is structured, in the sense that there is prior knowledge on the presence of structural zeros in the transfer matrix G^0. In order to analyze this situation, we rewrite the expression (11) as

$$\frac{1}{2\pi}\sum_{\ell=1}^{p}\{\int_{-\pi}^{\pi} \Delta M_\ell(e^{i\omega},\theta)\Phi_\kappa(\omega)\Delta M_\ell^T(e^{-i\omega},\theta)d\omega\} = 0 \tag{12}$$

with $M_\ell(e^{i\omega},\theta)$ the ℓ-th row of $M(e^{i\omega},\theta)$.

Additionally we define $\kappa^{[\ell]}$ as the subset of signals in κ corresponding to the column numbers of the nonzero entries in the row $M_\ell(q,\theta)$. In other words: $\kappa^{[\ell]}$ are those signals u and e that in the parametrized model are inputs to nonzero maps to y_ℓ, with y_ℓ the ℓ-th component of output vector y. Then the following result can be formulated.

Proposition 1. *Consider the model setup as described in Section 3. If the rows in the model $M_\ell(q,\theta)$ are independently parametrized, then $M(q,\theta) = M^0(q)$ is the unique minimum of the cost function in (5) if, for $\ell = 1,\dots,p$, $\Phi_{\kappa^{[\ell]}}(\omega) \succ 0$ for almost all ω.*

Proof. Since the cost function (12) is written as a sum of p components, with each of the components ≥ 0, the implication (10) can now be applied to each component separately. Uniqueness of the minimum is then guaranteed if for each individual component the corresponding spectral density is positive definite. □

By considering $\kappa^{[\ell]}$ rather then κ, we take account of the structural zero entries in row $M_\ell(q,\theta)$. Note that if one of the rows in $M(q,\theta)$ has no structural zero entries, then there is no added value in the above result, since it will reduce to the condition $\Phi_\kappa(\omega) \succ 0$ for the non-structured case.

However there is an important second consequence of the handling of structure. In the situation that we are actually only interested in consistent estimation of row ℓ in the model $M(q,\theta)$, i.e. related to the particular output y_ℓ, then the following result can be formulated.

Proposition 2. *Consider the model setup as described in Section 3. If the parameters in row $M_j(q,\theta)$ are independent of the parameters in the other rows of $M(q,\theta)$, then $M_j(q,\theta) = M_j^0(q)$ is unique in the minimum of the cost function in (5) if $\Phi_{\kappa^{[j]}}(\omega) \succ 0$ for almost all ω.*

Proof. Since the cost function (12) is actually a summation over p nonnegative terms, the term $\ell = j$ can be minimized independent of the other terms if there are no shared parameters. Uniqueness of the minimum for the term related to $\ell = j$ is then achieved when the corresponding spectrum related to the considered row of $\Delta M(e^{i\omega},\theta)$ is positive definite. □

This result provides a data-informativity condition for the consistent estimation of the model terms in $M_j(q,\theta)$ irrespective of possible consistency of the other terms in the model $M(q,\theta)$. As a result it is a less conservative condition than the original (non-structured) result $\Phi_\kappa(\omega) \succ 0$.

A (prediction error) identification algorithm that would estimate a model by minimizing a scalar cost function of the prediction error (6), would, under the data-informativity condition of Proposition 2, typically lead to a situation where part of the model parameters are uniquely determined and estimated consistently, while other parameters are non-unique. For the concerned optimization algorithms, this requires the appropriate handling of this situation in terms of dealing with singular Jacobian matrices during the optimization. Currently applied optimization algorithm, as e.g. implemented in MATLAB's System Identification Toolbox, can handle this effectively.

One may wonder why it would be interesting, when identifying a MIMO model, to focus on the consistency of only a part of the model. This problem particularly appears when identifying a single module in a dynamic network, as further elaborated in subsequent sections.

Remark 1. While the vector $\kappa^{[j]}$ is composed of the signals related to (the columns of) nonzero terms in the corresponding model $[G(q,\theta)\ H(q,\theta)]_{j*}$, the theory actually holds for signals related to (the columns of) *parametrized* terms in the corresponding model. In other words, terms in the model that are known a priori (like controllers in a network), can be excluded from $\kappa^{[j]}$, as they do not serve as inputs to terms that are parametrized and need to be identified.

5.2. Including a network structure in the generation of input signals

The result of Proposition 2 can be further worked out if, besides structure in the considered MIMO system, we also consider structure in the generation of the input signal u. As a further step towards the consideration of dynamic networks, we will first consider the situation that the input signal u has been generated in a structured way, including dependencies among the different components in u.

In line with the notation used in dynamic networks (see e.g. Van den Hof et al., 2013), we will consider the input u to be constructed as a multivariable filtered version of an external n_x-dimensional vector signal x, such that

$$u(t) = P(q)u(t) + Q(q)x(t) \tag{13}$$

with P a $m \times m$ proper and stable transfer function matrix, with $(I - P(q))$ invertible, P being hollow, meaning that $[P(q)]_{kk} = 0$ for all k, and Q being composed of a (sub)set of columns of the $m \times m$ identity matrix. Furthermore, and without loss of generality, it is assumed that $\Phi_x(\omega)$ is diagonal and $\Phi_x(\omega) \succ 0$ for all ω. An example of a network graph that generates an input vector u is shown in Fig. 2, in the particular situation of a target module G_{j1}.

Note that in this representation x can be either a user-controlled measured signal or an unmeasured disturbance process.

The reason for writing the input signal in this specific form, is that it allows us to specify dependencies among input signals in a graph-based format, which is particularly attractive when generalizing our framework from an open-loop situation to a dynamic network situation.

Definition 2. Consider a network graph $\mathcal{G}$ with vertex set $\mathcal{V} := \{u_k\}_{k=1,\dots,m} \cup \{x_k\}_{k=1,\dots,n_x}$ and directed edge/link set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$:

- $(u_k, u_\ell) \in \mathcal{E}$ if and only if $[P]_{\ell k} \neq 0$;
- $(x_k, u_\ell) \in \mathcal{E}$ if and only if $[Q]_{\ell k} \neq 0$.

Vertex u_ℓ is an out-neighbor of u_k if $(u_k, u_\ell) \in \mathcal{E}$. □

Definition 3. A vertex set $\mathcal{V}_{\mathcal{D}^c}$ is called a disconnecting set from $\mathcal{V}_1$ to $\mathcal{V}_2$, if upon removal of $\mathcal{V}_{\mathcal{D}^c}$ from the graph, there is no directed path from $\mathcal{V}_1$ to $\mathcal{V}_2$. It is a minimum disconnecting set if it has the minimum cardinality among all such disconnecting sets. □

Note that $\mathcal{V}_{\mathcal{D}^c}$ may contain elements of $\mathcal{V}_1$ and/or $\mathcal{V}_2$.

We recall that $\kappa(t) := \begin{bmatrix} u(t) \\ e(t) \end{bmatrix}$ and that $\kappa^{[j]}(t)$ is the selection of components in $\kappa(t)$ that correspond to parametrized entries in

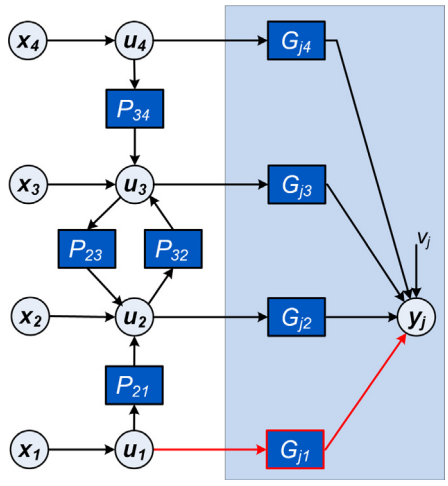


Fig. 2. Five node example with target module $G_{j1}(q)$, and dependence structure among the input components u as in (13), where v_j is the jth component of $H^0(q)e(t)$ in (1).

the columns of $M_j(q,\theta)$. In line with this decomposition of $\kappa(t)$, we write $\kappa^{[j]}$ as $\kappa^{[j]} = \begin{bmatrix} \kappa_u^{[j]} \\ \kappa_e^{[j]} \end{bmatrix}$. This implies that the signals in $\kappa_u^{[j]}$ are the inputs to parametrized modules in the MISO model with output y_j. Note that the input signals to modules that are known a priori are not present in $\kappa^{[j]}$.

As a next step we are going to decompose the input signals in $\kappa_u^{[j]}$ in three different components. In doing so we adopt the notation that for an m-dimensional node vector u, and $\mathcal{D}^c \subset \{1,\dots,m\}$ an index set, $u_{\mathcal{D}^c}$ is a vector of those components of u that are defined by the index set $\mathcal{D}^c$. We will use this general notation throughout the remaining part of the paper.

Definition 4. Let $\{u_{\mathcal{D}^c}\}$ be defined as a disconnecting set from $\{u_i\}$ to the other components in $\{\kappa_u^{[j]}\}$, under the constraints that $i \notin \mathcal{D}^c$.[1] □

The interpretation of this disconnecting set is that it disconnects the input u_i of the target module from all other inputs to parametrized modules. To illustrate this, consider the situation as depicted in Fig. 2. The disconnecting set $\{u_2\}$ disconnects $\{u_1\}$ from $\{u_3\}$. As a result, the dependency of u_3 on u_1 is fully covered by the signal u_2. This phenomenon is closely related to the concept of blocked parallel paths from u_1 to y_j, i.e. parallel to the target module G_{j1} (Dankers et al., 2016), i.e. u_2 blocks all connections from u_1 to y_j that are not passing through G_{j1}.

As we are focusing on estimating the entry $G_{ji}(q)$ of our MIMO system, we focus on its input u_i and determine a decomposition of the vector $\kappa_u^{[j]}$ as follows:

$$\kappa_u^{[j]} = \begin{bmatrix} u_i \\ u_{\mathcal{D}^c} \\ u_{\mathcal{T}} \end{bmatrix} \tag{14}$$

where $u_{\mathcal{T}}$ is a remainder set, after defining $u_{\mathcal{D}^c}$. In the example of Fig. 2, $u_i = u_1$, $u_{\mathcal{D}^c} = u_2$ and $u_{\mathcal{T}} = \begin{bmatrix} u_3 \\ u_4 \end{bmatrix}$.

The full input vector u is then composed of the elements in $\kappa_u^{[j]}$ complemented with the input signals $u_{\mathcal{F}}$ related to a priori known terms in $G(q,\theta)_{j*}$. In the sequel the following set of external signals will appear to be instrumental.

[1] $u_{\mathcal{D}^c}$ will obviously be dependent on i and j. In the remaining text we will focus on one fixed choice of i and j, avoiding the use of more complex notation.

Definition 5. Let $x_{\mathcal{T}^*}$ be defined as those components in vector signal x, that in the graph of the input signal u, have a path to $\{u_{\mathcal{T}}\}$ without passing through $\{u_{\mathcal{D}^c}\}$. □

On the basis of expression (13) for u, we can then remove (immerse) the signals $u_{\mathcal{F}}$ from the equations and derive an expression for $u_{\mathcal{T}}$ as follows.

Proposition 3. *Consider the expression (13) for u and consider the decomposition of u as specified in (14). Then there exist rational transfer function matrices $T_s(q)$ and $R_s(q)$ such that $u_{\mathcal{T}}$ can be written as:*

$$u_{\mathcal{T}}(t) = T_s(q)u_{\mathcal{D}^c}(t) + R_s(q)x_{\mathcal{T}^*}(t). \tag{15}$$

Proof. See Appendix.

Note that for external signals that pass through $\{u_{\mathcal{D}^c}\}$ before reaching $\{u_{\mathcal{T}}\}$, their contribution to $u_{\mathcal{T}}$ is covered by the transfer function $T_s(q)$ in (15).

We can now formulate the following result, which is a further sharpening of the result of Proposition 2.

Theorem 4. *Consider the situation of Proposition 2, where the input signals are generated according to (13), the input signal decomposition (14), and $x_{\mathcal{T}^*}$ according to Definition 5.*

- *Let the columns of $M_j(q,\theta)$ corresponding to the input signals $u_{i\cup\mathcal{D}^c}$ be parametrized independently from the columns corresponding to the input signals $u_{\mathcal{T}}$.*
- *Let $u_{i\cup\mathcal{D}^c}{}^{\perp x_{\mathcal{T}^*}}$ be the projection[2] of signals $u_{i\cup\mathcal{D}^c}$ onto the orthogonal complement of $x_{\mathcal{T}^*}$.*

Then $G_{ji}(q,\theta) = G_{ji}^0(q)$ is unique in the minimum of the cost function in (5) if

$$\Phi_{\mu^{[j]}}(\omega) \succ 0 \text{ for almost all } \omega \tag{16}$$

with

$$\mu^{[j]} := \begin{bmatrix} u_{i\cup\mathcal{D}^c}{}^{\perp x_{\mathcal{T}^*}} \\ \kappa_e^{[j]} \end{bmatrix}. \tag{17}$$

Proof. See Appendix. □

The interpretation of this result is as follows:

a. It is not the full input vector $\kappa_u^{[j]}$ of inputs that have a parametrized map to the output, that requires excitation. Only a subset of input signals, present in $u_{i\cup\mathcal{D}^c}$ need excitation.
b. For exciting this subset of inputs, all external signals x can be used, except for the x signals that are present in $x_{\mathcal{T}^*}$. For excitation of $u_{i\cup\mathcal{D}^c}$ we can use those signals x that in the graph of the input network reach $\{u_{i\cup\mathcal{D}^c}\}$ without passing through $\{u_{\mathcal{T}}\}$.
c. The white noise signals in $\kappa_e^{[j]}$ are trivially sufficiently exciting.

Remark 2. Since the white noise signals e in $\kappa_e^{[j]}$ are considered to be independent of x, an equivalent statement for (16) would be $\Phi_{u_{i\cup\mathcal{D}^c}{}^{\perp x_{\mathcal{T}^*}}}(\omega) \succ 0$. However, for generalization of these results to the dynamic network case, we prefer the current expressions (16)–(17).

[2] Projection is considered with respect to the inner product defined by $\langle x_m, y_n\rangle := \bar{\mathbb{E}}[x_m(t)y_n(t)]$, while the projection of a vector signal x onto a vector signal y is defined as the projection of every component of vector x onto all components of vector y.

Remark 3. In this section, the node set $\{u_{\mathcal{D}^c}\}$ is defined as any disconnecting set from u_i to $\{\kappa_u^{[j]}\}\backslash\{u_i\}$. It may seem attractive to select it as a *minimum* disconnecting set to reduce the number of required excitation signals, but since $\mathcal{D}^c$ also influences the set $x_{\mathcal{T}^*}$ such a choice does not always lead to the smallest number of excitation signals to be added.

Interpretation (a.) above confirms the earlier insights of Dankers et al. (2016), that for consistently estimating a target module G_{ji} it would be sufficient to measure input u_i and those inputs that block all parallel paths from u_i to y_j This has been formalized in the Parallel Path and Loop condition (Dankers et al., 2016), or alternatively phrased as a separating/disconnecting set condition (Dankers et al., 2016; Shi et al., 2022).

The current results show that we do not need to exclude the remaining inputs from the predictor model, as was done in Dankers et al. (2016). We can include the inputs $u_{\mathcal{T}}$ in our predictor model, but the concerned input signals do not require additional excitation. Their effect on output y_j is modeled, but for this we do not require a consistent estimate of the concerned transfer functions that map the input signals to y_j. In this setting the current result improves on the results of Van den Hof and Ramaswamy (2020), Van den Hof et al. (2023) where persistence of excitation of all inputs was chosen in order to guarantee data-informativity.

6. Example

We illustrate the result of Theorem 4 through the four-input one-output example as indicated in Fig. 2. In the considered situation, G_{j1} is the target module to be estimated, while the predictor model is given by:

$$\varepsilon(t,\theta) = H(q,\theta)^{-1}[y_j(t) - \sum_{k=1}^{4} G_{jk}(q,\theta)u_k(t)]. \tag{18}$$

First we consider the situation that all four modules $G_{jk}(q,\theta)$, $k = 1,\dots,4$ are parametrized. A minimum disconnecting set $\{u_{\mathcal{D}^c}\}$ from input $u_i = u_1$ to all other inputs of parametrized modules is $\{u_2\}$. In an alternative interpretation it is clear that u_2 blocks all parallel paths (i.e. parallel to the target module G_{j1}) from u_i to output y_j.

The signal $u_{i\cup\mathcal{D}^c}$ is composed of $u_i = u_1$, $u_{\mathcal{D}^c} = u_2$, while $u_{\mathcal{T}} = [u_3 \;\; u_4]^T$. Possible excitation or noise signals x_3 and x_4 will serve as $x_{\mathcal{T}^*}$. Since u and e are considered to be independent, and e is white noise, the condition (16) can be replaced by a positive definite spectrum of $u_{\{1\cup2\}}{}^{\perp\{x_3,x_4\}}$.

For evaluating this spectrum we can write u_1 and u_2 as function of the external signals affecting the system:

$$u_1 = \sum_{k=1}^{4} F_{1k}(q)x_k(t); \qquad u_2 = \sum_{k=1}^{4} F_{2k}(q)x_k(t). \tag{19}$$

The vector signal $u_{\{1\cup2\}}{}^{\perp\{x_3,x_4\}}$ can then be written as

$$u_{\{1\cup2\}}{}^{\perp\{x_3,x_4\}} = \underbrace{\begin{bmatrix} F_{11}(q) & F_{12}(q) \\ F_{21}(q) & F_{22}(q) \end{bmatrix}}_{F(q)} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

A positive definite spectrum of $u_{\{1\cup2\}}{}^{\perp\{x_3,x_4\}}$ is then obtained for almost all ω if the matrix $F(z)$ has full row rank over the field of rational functions, and the spectrum of $[x_1 \;\; x_2]^T$ is positive definite (Van den Hof & Ramaswamy, 2020).

This shows that excitation of inputs u_1 and u_2 is sufficient for data-informativity, while excitation of u_3 and u_4 is not necessary.

Moreover the signals x_3 and x_4 cannot be used to excite (u_1, u_2) since they pass through $\{u_{\mathcal{T}}\}$ before reaching $\{u_2\}$(= $\{u_{\mathcal{D}^c}\}$), and therefore they belong to $x_{\mathcal{T}^*}$. To understand why x_3 and x_4 cannot be used for excitation of (u_1, u_2) we consider the prediction error

$$\varepsilon(t,\theta) = H(q,\theta)^{-1}\left[\sum_{k=1}^{4} \Delta G_{jk}(q,\theta)u_k(t) + H^0(q)e(t)\right].$$

Writing u_k as a function of the variables x, according to the situation of Fig. 2, this can be rewritten as

$$\varepsilon(t,\theta) = H(q,\theta)^{-1}\left[\sum_{k=1}^{4} T_k(q,\theta)x_k(t) + H^0(q)e(t)\right]$$

with

$$\begin{aligned}
T_1(q,\theta) &= \Delta G_{j1} + \Delta G_{j2}\frac{P_{21}}{1-P_{32}P_{23}} + \Delta G_{j3}\frac{P_{32}P_{21}}{1-P_{32}P_{23}} \\
T_2(q,\theta) &= \Delta G_{j2}\frac{1}{1-P_{32}P_{23}} + \Delta G_{j3}\frac{P_{32}}{1-P_{32}P_{23}} \\
T_3(q,\theta) &= \Delta G_{j2}\frac{P_{23}}{1-P_{32}P_{23}} + \Delta G_{j3}\frac{1}{1-P_{32}P_{23}} \\
T_4(q,\theta) &= \Delta G_{j2}\frac{P_{23}P_{34}}{1-P_{32}P_{23}} + \Delta G_{j3}\frac{P_{34}}{1-P_{32}P_{23}} + \Delta G_{j4}
\end{aligned}$$

where $\Delta G_{jk} := G^0_{jk}(q) - G_{jk}(q,\theta)$, for $k = 1,\dots,4$. The minimum of the cost function $\bar{\mathbb{E}}\varepsilon^T(t,\theta)\varepsilon(t,\theta)$ is achieved if the T_k-dependent terms are zero. Presence of x_k implies that $T_k(q,\theta) = 0$ in the minimum of the cost function. If x_1 and x_2 are present and persistently exciting, then T_1 and T_2 are forced to be zero, which implies that $\Delta G_{j1} = 0$, and thus a consistent estimate of G^0_{j1} is guaranteed. This follows from the observation that $T_1(q,\theta) = \Delta G_{j1} + P_{21}T_2(q,\theta)$. If we would replace x_2 by presence of x_3, a similar result does not exist. Introducing x_4 shows that $T_4(q,\theta)$ gets involved in the expression and a new unknown term ΔG_{j4} is introduced also. This explains that forcing T_1 and T_4 to be zero is not sufficient for arriving at $\Delta G_{j1} = 0$, or in other words: exciting x_1 and x_4 does not lead to a consistent estimate of the target module.

In the case that module $G_{j2}(q)$ is not parametrized, but known a priori, there are two options for choosing (minimum) disconnecting sets:

- When selecting $u_{\mathcal{D}^c} = u_2$, it follows that $u_{\mathcal{T}} = [u_3 \;\; u_4]^T$, and as a result x_3 and x_4 cannot be used for excitation of (u_1, u_2), and therefore we need excitations x_1 and x_2.
- When selecting $u_{\mathcal{D}^c} = u_3$, it follows that $u_{\mathcal{T}} = u_4$, and as a result x_4 cannot be used for excitation of (u_1, u_3). This implies that external signals (x_1, x_2, x_3) can be used for realizing a positive definite spectrum of $u_{\{1\cup3\}}{}^{\perp x_{\mathcal{T}^*}}$. This is guaranteed by having an excitation x_1 plus either x_2 or x_3.

The result can be verified by setting $\Delta G_{j2} = 0$ in the above expressions for $T_k(q,\theta)$. Then $\Delta G_{j3} = 0$ will follow from the presence of either x_2 or x_3, while ΔG_{j1} then follows through the presence of x_1.

Note that in case multiple disconnecting sets are possible, each one of them provides a set of sufficient conditions for data-informativity. Not satisfying the conditions for one particular choice of disconnecting set, does not necessarily lead to the conclusion that the data-informativity conditions are not satisfied.

In the case that G_{j4} is not parametrized and known a priori, the (minimum) disconnecting set from u_1 to $\{u_2, u_3\}$ is given by u_2. Then $u_{\mathcal{T}} = u_3$, while the external signals $x_{\mathcal{T}^*}$ that affect $u_{\mathcal{T}}$ are given by $x_{\mathcal{T}^*} = [x_3 \;\; x_4]^T$. This implies that (only) signals (x_1, x_2) are available for realizing a positive definite spectrum of $u_{\{1\cup2\}}{}^{\perp x_{\mathcal{T}^*}}$. Verification of this through the expressions of $T_k(q,\theta)$,

shows that with $\Delta G_{j4} = 0$, the linear combination of the two terms in T_3 and T_4 are different from the linear combinations of ΔG_{j2} and ΔG_{j3} in the expressions for T_1 and T_2. As a result the original result with all modules parametrized follows as the solution.

7. Path-based interpretation

As shown in earlier work (Van den Hof & Ramaswamy, 2020), a spectrum condition for data-informativity like (16), can be guaranteed to hold generically, if particular path-based conditions on the network graph are satisfied. As was shown to hold in a related identifiability studies in Hendrickx et al. (2019), rank conditions on particular transfer functions in a dynamic network can generically be translated into conditions on vertex disjoint path connections between sets of nodes in the network graph. Two directed paths are called vertex disjoint if they do not share any vertex, including the starting and ending vertices. They can be calculated on the basis of the network graph through a max-flow min-cut algorithm as the Ford-Fulkerson algorithm. For handling this result we copy the formulation from Van den Hof et al. (2023), replacing κ by $\mu^{[j]}$ from (17):

Proposition 5 (*Van den Hof et al., 2023*). *The spectrum condition $\Phi_{\mu^{[j]}} \succ 0$ for almost all ω, holds generically*[3] *if in the graph of the network there exist $\dim(\mu^{[j]})$ vertex disjoint paths from the node sets $\{x, e\}$ to $\{\mu^{[j]}\}$, and the signals in x are mutually independent and persistently exciting.*

Actually, it should be sufficient to require only those x signals to be persistently exciting that serve as starting nodes in the vertex disjoint paths. When specifying the above result for the particular situation treated in Section 5, we obtain the following.

Corollary 6. *The graph-based condition in Proposition 5 is satisfied if there exist $\dim(u_{\mathcal{D}^c}) + 1$ vertex disjoint paths between the node sets $\{x \backslash x_{\mathcal{T}^*}\}$ and $\{u_{i\cup\mathcal{D}^c}\}$.*

Proof. Since u and e are independent, the spectrum condition $\Phi_{\mu^{[j]}}(\omega) \succ 0$ can be recast into the expression $\Phi_{u_{i\cup\mathcal{D}^c}{}^{\perp x_{\mathcal{T}^*}}}(\omega) \succ 0$. Together with the fact that by assumption $\Phi_x(\omega) \succ 0$ for all ω, this directly leads to the result. □

We can illustrate this result for the example in Section 6. In the situation where all G-modules are parametrized it was found that $\{u_{\mathcal{D}^c}\} = \{u_2\}$, $\{u_{i\cup\mathcal{D}^c}\} = \{u_1, u_2\}$ and $\{x \backslash x_{\mathcal{T}^*}\} = \{x_1, x_2\}$. Therefore, two vertex disjoint paths between $\{x_1, x_2\}$ and $\{u_1, u_2\}$, are sufficient to guarantee (generic) data-informativity. This condition is satisfied if both x_1 and x_2 are present.

8. Single module identification in dynamic networks

8.1. Single module data-informativity results

In this section we are going to make the step from the isolated open-loop situation as considered before, to the situation where the considered MIMO/MISO system can be part of a dynamic network, including possible feedback loops. Dynamic networks are interconnections of dynamic systems and can be represented in different forms, one of them being the so-called module representation (Van den Hof et al., 2013), defined by

$$w(t) = G^0(q)w(t) + H^0(q)e(t) + u(t) \tag{20}$$

with w an L-dimensional vector, $G^0(q)$ a hollow rational transfer function matrix, i.e. with zeros on the diagonal entries, e an L-dimensional vector of white noise processes, $H^0(q)$ the rational disturbance model, and $u(t) = R^0 \cdot r(t)$, accounting for the effect of measured external excitation signals r on the network, with R^0 a binary $L \times K$ matrix and r a K-dimensional vector of excitation signals. A non-zero element $G^0_{k\ell}(q)$ in $G^0(q)$ is referred to as a *module*.

It is further assumed that the network is stable and well-posed (Van den Hof et al., 2013). In this paper it is also assumed, for simplicity, that all elements in $G^0(q)$ are strictly proper.

When targeting on the identification of a single module G^0_{ji} in this network, several approaches are available for arriving at a consistent module estimate. An indirect identification method is suitable when there is a sufficient number of u signals present in the network, see e.g (Van den Hof et al., 2013; Dankers et al., 2016; Gevers et al., 2018). When noise signals e are required to provide sufficient excitation in the network for consistent estimation of the target module, a direct method is the prime approach, see e.g. Van den Hof et al. (2013), Dankers et al. (2016), Ramaswamy and Van den Hof (2021). In this paper we will focus on the direct method.

In this method a predictor model is constructed composed of a set of predictor inputs $w_{\mathcal{D}}$ and a set of predicted outputs $w_{\mathcal{Y}}$, on the basis of which the following predictor model equation can be formulated:

$$w_{\mathcal{Y}}(t) = \bar{G}^0(q)w_{\mathcal{D}}(t) + \bar{H}^0(q)\xi_{\mathcal{Y}}(t) + \bar{T}^0(q)u(t). \tag{21}$$

This actually concerns the description of part of the network (20), where the equations for the nodes outside of $w_{\mathcal{Y}}$ have been decoupled and removed, while the network nodes that are not present in $w_{\mathcal{Y}\cup\mathcal{D}}$ have been removed (immersed) from the description. The noise contribution $\bar{H}^0(q)\xi_{\mathcal{Y}}(t)$, represents the (correlated) noise that affects the output $w_{\mathcal{Y}}$, with $\bar{H}^0(q)$ being monic, stable and stably invertible, and $\xi_{\mathcal{Y}}$ a vector white noise process. In general, the noise model in (21) is the result of a spectral factorization, after manipulating the network Eq. (20); that is why the noise term $\xi_{\mathcal{Y}}$ is typically different from $e_{\mathcal{Y}}$. The matrix $\bar{T}^0(q)$ represents the effect of external excitation signals on $w_{\mathcal{Y}}$. Note that not all excitation signals that are present in the network will appear in the term $\bar{T}^0(q)u(t)$. Excitation signals that enter the network on nodes $w_{\mathcal{D}}$ will generally contribute to $w_{\mathcal{Y}}$ through their respective input term $w_{\mathcal{D}}$. The matrix $\bar{T}^0(q)$ can have both dynamic terms and fixed terms, the latter situation, e.g., occurring when an excitation signal directly enters the network on a node in $w_{\mathcal{Y}}$.

The single module identification problem is then formulated as the problem of identifying a single module $\bar{G}^0_{ji}(q)$, defined as the mapping from w_i to w_j in (21), on the basis of data $\{w_{\mathcal{D}\cup\mathcal{Y}}(t), u(t)\}_{t=0,\dots,N-1}$, under the condition that $\bar{G}^0_{ji}(q) = G^0_{ji}(q)$, i.e. the original module in the network. Estimation is done on the basis of the one-step ahead prediction error

$$\varepsilon(t,\theta) = \bar{H}(q,\theta)^{-1}\left[w_{\mathcal{Y}}(t) - \bar{G}(q,\theta)w_{\mathcal{D}}(t) - \bar{T}(q,\theta)u(t)\right] \tag{22}$$

and a quadratic type of cost function as in (5).

In order to arrive at a consistent estimate (Ljung, 1999) of the target module, a set of conditions has been formulated in previous works:

a. In order to guarantee module invariance, i.e. $\bar{G}^0_{ji}(q) = G^0_{ji}(q)$, a parallel path and loop (PPL) condition[4] is assumed to be satisfied, which influences the selection of predictor inputs $w_{\mathcal{D}}$, see Dankers et al. (2016).

[3] In line with its use in Hendrickx et al. (2019), this concerns a property that holds for all possible dynamic modules in the network, except possibly for a set of measure 0.

[4] The PPL condition (Dankers et al., 2016) requires that each path in the network from w_i to w_j excluding the path through G_{ji}, and every loop around w_j, pass through a measured node that is included in the predictor input.

b. In order to cope with *confounding variables*[5] multivariable noise models can be used to model the correlated disturbances, leading to a selection of an appropriate set of node signals in $w_{\mathcal{Y}}$ and $w_{\mathcal{D}}$, see Ramaswamy and Van den Hof (2021).
c. Conditions on data-informativity, see Van den Hof et al. (2023).
d. Technical conditions on the presence of delays in the network, see Ramaswamy and Van den Hof (2021).
e. A system-in-the-model set condition, implying that the chosen model set is sufficiently flexible to include the true system dynamics.

In the research so far, the mentioned data-informativity conditions have been formulated for consistent estimation of the full "subnetwork" $\bar{G}^0(q)$. However, in view of the results presented in Section 5, these conditions can now be further relaxed, by focusing on consistent estimation of $\bar{G}^0_{j\star}(q)$ only, i.e. the jth row of $\bar{G}^0(q)$, corresponding to the output of the target module, and actually on element $\bar{G}^0_{ji}(q) = G^0_{ji}(q)$ only, while taking account of the structurally zero and possible fixed (non-parametrized) entries in this row.

For the formulation of the relaxed conditions, we denote

$$M(q,\theta) = \begin{bmatrix} \bar{G}(q,\theta) & \bar{H}(q,\theta) & \bar{T}(q,\theta) \end{bmatrix} \tag{23}$$

and $\Delta M(q,\theta) = \begin{bmatrix} \Delta\bar{G}(q,\theta) & \Delta\bar{H}(q,\theta) & \Delta\bar{T}(q,\theta) \end{bmatrix}$, with $\Delta\bar{G}(q,\theta) := \bar{G}^0(q) - \bar{G}(q,\theta)$ etcetera, where the resulting prediction error can be written as

$$\varepsilon(t,\theta) = \bar{H}(q,\theta)^{-1}\Delta M(q,\theta)\kappa(t) \tag{24}$$

with

$$\kappa = \begin{bmatrix} w_{\mathcal{D}} \\ \xi_{\mathcal{Y}} \\ u \end{bmatrix}. \tag{25}$$

In accordance with the reasoning in Section 5 we focus on output node w_j and therefore on the jth row $M_j(q,\theta)$ of $M(q,\theta)$, evaluating under which conditions the power of the signal $\Delta M_j(q,\theta)\kappa^{[j]}$ is equal to zero, where

$$\kappa^{[j]} = \begin{bmatrix} w^{[j]}_{\mathcal{D}} \\ \xi^{[j]}_{\mathcal{Y}} \\ u^{[j]} \end{bmatrix} \tag{26}$$

being those terms in κ that correspond with a *parametrized* term in the corresponding column of the row $M_j(q,\theta)$.

Definition 6. Let $\{w_{\mathcal{D}^c}\}$ be defined as a disconnecting set from w_i to the other components in $\{w^{[j]}_{\mathcal{D}}\}$, under the constraint that $i \notin \mathcal{D}^c$.

Note that Definition 6 is similar to Definition 4, but the consequence is slightly different, in the sense that, because of the network structure, there can be nodes in $w_{\mathcal{D}^c}$ now that are not necessarily in-neighbors of w_j. Similar to the situation of Section 5, we define

- $w_{\mathcal{T}}$ as the node vector defined by $w_{\mathcal{T}} := w^{[j]}_{\mathcal{D}\setminus\{w_i\cup w_{\mathcal{D}^c}\}}$, i.e. the set of input nodes to parametrized modules, excluding w_i, that are not in the disconnecting set.
- $w^{[j]}_{\mathcal{F}}$ as the node vector with nodes that are inputs to a non-parametrized (known) module in the jth row of the predictor model.

[5] A confounding variable is an unmeasured variable that has paths to both the input and output of an estimation problem (Pearl, 2000).

We can now formulate a companion result for Proposition 3 for the network situation.

Definition 7. Let $x_{\mathcal{T}^*}$ be defined as those components in vector signal $x = vec(u,e)$, that in the graph of the network, have a path to $\{w_{\mathcal{T}}\}$ without passing through $\{w_{\mathcal{D}^c}\}$.

Proposition 7. *Consider the network Eq. (20) for w and consider the node vectors w_i, $w_{\mathcal{T}}$, $w_{\mathcal{D}^c}$ and $x_{\mathcal{T}^*}$ as defined above. Then there exist rational transfer function matrices $T_s(q)$ and $R_s(q)$ such that $w_{\mathcal{T}}$ can be written as:*

$$w_{\mathcal{T}}(t) = T_s(q)w_{\mathcal{D}^c}(t) + R_s(q)x_{\mathcal{T}^*}(t). \tag{27}$$

Proof. The proof of this result is dual to the proof of Proposition 3, with all u-node signals replaced by w-node signals. □

This provides the right conditions for showing the data-informativity result for single module identification.

Theorem 8. *Consider the network (20), predictor model (21)–(22), the signal vector $\kappa^{[j]}$ (26) and the node vectors w_i, $w_{\mathcal{D}^c}$, $w_{\mathcal{T}}$, $x_{\mathcal{T}^*}$ as defined above. If*

- *row $M_j(q,\theta)$ is parametrized independently from the other rows in $M(q,\theta)$, and*
- *the columns in $M_j(q,\theta)$ corresponding to the inputs $w_{i\cup\mathcal{D}^c}$ are parametrized independently from the other columns in $M_j(q,\theta)$,*

then $\bar{G}_{ji}(q,\theta) = \bar{G}^0_{ji}(q)$ is unique in the minimum of the quadratic cost function (5) applied to (22), if

(a) $\Phi_{\eta^{[j]}}(\omega) \succ 0$, for almost all ω

with

$$\eta^{[j]} = \begin{bmatrix} w_{i\cup\mathcal{D}^c}{}^{\perp\chi} \\ \xi^{[j]}_{\mathcal{Y}} \end{bmatrix} \tag{28}$$

and

$$\chi = \begin{bmatrix} x_{\mathcal{T}^*} \\ u^{[j]} \end{bmatrix}. \tag{29}$$

Proof. See Appendix. □

One of the main differences with the situation of Section 5 is that the two components in the vector $\eta^{[j]}$ are no longer independent. Because of possible feedback loops present in the network, the concerned signals can be correlated. An implication of the spectrum condition is that $w_{i\cup\mathcal{D}^c}$ should be persistently excited, through the presence of external signals that exclude χ and $\xi^{[j]}_{\mathcal{Y}}$, while $\xi^{[j]}_{\mathcal{Y}}$ should be persistently exciting in itself. This latter condition is due to the fact that these signals are inputs to parametrized terms in the noise model, and therefore cannot be used for exciting the elements in G.

The external signals χ that are excluded from serving as excitation sources for estimating $G^0_{ji}(q)$, now also include all u-signals that are input to a parametrized module towards the target output w_j, i.e. the inputs that correspond to parametrized terms in $[\bar{T}(q,\theta)]_{j*}$, the jth row of $\bar{T}(q,\theta)$ in (22). Note however that, different from the term $\xi^{[j]}_{\mathcal{Y}}$, the signals in $u^{[j]}$ do not require excitation, as they are not part of the vector signal $\eta^{[j]}$.

Following the reasoning in Section 7 we can now also formulate the path-based conditions for generically satisfying the positive definite spectrum condition in Theorem 8.

Proposition 9. *Consider the situation of Theorem 8, and consider the selections of external signals, defined by:*

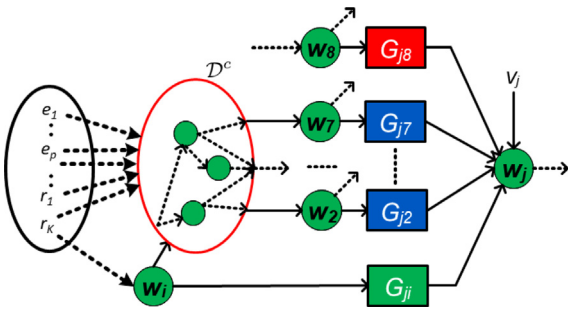


Fig. 3. Network situation with target module G_{ji}. Module G_{j8} is known and therefore not parametrized. The disconnecting set $\{w_{\mathcal{D}^c}\}$ is disconnecting from node w_i to the remaining inputs to parametrized modules. $w_{\mathcal{D}^c}$ and w_i need sufficient excitation from external signals, either e- or $r-$ signals, while external signals that have a parametrized link to the output w_j are excluded. Note that $u = R^0 r$ with R^0 a binary matrix.

- $e^{\perp\mathcal{Y}}$: *All e signals in the network that do not have an unknown path*[6] *to* $w_{\mathcal{Y}}$, *and*
- $u^{\perp j}$: *All u-signals in the network that do not have a parametrized link to* w_j *in the predictor model.*

Then the data informativity condition (a) in Theorem 8 is generically satisfied if there exist $dim(\mathcal{D}^c) + 1$ *vertex disjoint paths from* $\{e^{\perp\mathcal{Y}}, u^{\perp j}\}\backslash\{x_{\mathcal{T}^*}\}$ *to* $\{w_{i\cup\mathcal{D}^c}\}$, *and the signals in* $\{e^{\perp\mathcal{Y}}, u^{\perp j}\}$ *are mutually independent and persistently exciting.*

Proof. See Appendix. □

The result shows that for appropriately exciting $w_{i\cup\mathcal{D}^c}$, we can use all external signals that in the predictor model do not have a parametrized link to the output. External signals that do have such a parametrized link are required for exciting the concerned entries in $\bar{H}(q,\theta)$ and $\bar{T}(q,\theta)$ respectively, and therefore are not available for exciting the node inputs for estimating $G^0_{ji}(q)$. From the remaining set of external signals, only those external signals can be used that reach the predictor input signals through the set of nodes $\{w_{i\cup\mathcal{D}^c}\}$, without passing through $\{w_{\mathcal{T}}\}$.

An illustration of this situation is provided in Fig. 3.

Remark 4. The selection $e^{\perp\mathcal{Y}}$ as formulated in Proposition 9, can in some specific cases of predictor models, be slightly enlarged, in particular when the number of e nodes with an unknown path to $e_{\mathcal{Y}}$ exceeds $dim(w_{\mathcal{Y}})$. This holds for predictor models that allow for predictor inputs that have a confounding variable with an output signal, a so-called nonempty set $\{w_{\mathcal{B}}\}$ (Ramaswamy & Van den Hof, 2021). A further specification of this will be presented elsewhere.

8.2. Discussion

The new single module data-informativity results presented in Theorem 8 and Proposition 9, are relaxed versions of the data-informativity conditions presented in Van den Hof et al. (2023), where the target was to consistently estimate the full (unstructured) predictor model. It has been shown in Van den Hof et al. (2023) that for satisfying those data-informativity conditions it may be necessary to adjust the selection of predictor inputs and outputs. In other words, for a particular selection of predictor inputs and outputs, it is not guaranteed that the data-informativity can be met by adding a sufficient number of external excitation signals. That is different in the current results, where the data-informativity conditions can always be satisfied by adding a sufficient number of external excitation signals at appropriate locations.

Corollary 10. *Consider the situation of Theorem 8. Then for any predictor model that satisfies the structural conditions for module invariance and confounding variables handling,*[7] *there exists a disconnecting set* $\{w_{\mathcal{D}^c}\}$ *such that the data-informativity condition for consistent estimation of the target module* $\bar{G}^0_{ji}(q)$ *can be satisfied, by adding a sufficient number of external excitation signals to the predictor inputs* $w_{i\cup\mathcal{D}^c}$.

When estimating a single module $G^0_{ji}(q)$ in a dynamic network, the user has the option now to require two different sets of conditions for data-informativity, either the single module conditions of the current paper, or the full predictor model conditions of Van den Hof et al. (2023). While the single module conditions are more relaxed, this situation also comes at a cost. When validating an estimated module, as part of a consistently estimated predictor model, model validation can be applied to the full predictor model, for validating all of the present modules. This can e.g. be done by considering a validation data set. However when only the target module is estimated consistently, validation of this module using a validation data set, would require that in this validation data set, the node signals that are input to parametrized modules in the predictor model, are exactly the same as in the estimation data set. This is caused by the fact that in the consistent single module estimation case, the estimated modules, different from the target module, will become dependent on the particular input node signals, present in the estimation data set.

So far we have been dealing with the data-informativity conditions for the direct method of estimating predictor models in a dynamic network, see Ramaswamy and Van den Hof (2021). There are alternative methods, as e.g. the multi-step method of Fonken et al. (2023), where data-informativity conditions are also more relaxed. Construction of predictor models for both the direct method and the multi-step method, including the different types of data-informativity conditions, has been implemented in the MATLAB app and toolbox SYSDYNET (Van den Hof et al., 2024).

One of the prime consequences of the results presented in this paper is in the relation with single module identifiability. Identifiability conditions for single modules in a dynamic network have been studied in Weerts et al. (2018a), Hendrickx et al. (2019), Shi et al. (2022, 2023). These conditions are typically not dependent on a particular identification algorithm. At the same time, identification algorithms have been studied that provide consistent target module estimates under specific conditions, among which data-informativity conditions. One would expect that if the single module identifiability conditions are satisfied, that there exists an estimation algorithm that can provide the consistent target module estimate. However, so far this step has only been achieved for a particular situation, namely when the single module identifiability conditions are satisfied by external excitation signals r only. In that situation an indirect method can be used to estimate the target module consistently, see e.g. Van den Hof et al. (2013), Dankers et al. (2016), Gevers et al. (2018). However when both r-excitation and external e-signals are required for satisfying the single module identifiability conditions, there was a gap with the related data-informativity conditions of the estimation algorithms.

In order to understand how the current results close this gap, we consider the identifiability result as formulated in Theorem 4 and Corollary 1 of Shi et al. (2022), adapted here to the situation of single module identifiability. We also consider a model

[6] An "unknown path" is a path that contains a link with unknown dynamics that is parametrized in the predictor model.

[7] See conditions a. and b. in Section 8.1.

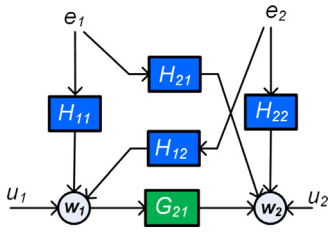


Fig. 4. Two node example with correlated disturbances and target module $G_{21}(q)$.

set $\mathcal{M}$ of a full network that satisfies the basic assumptions as formulated in Shi et al. (2022), including the assumption that all modules in $G(q,\theta)$ are strictly proper.

Theorem 11 (*Shi et al., 2022*). *Let $\mathcal{X}_j$ be composed of those components in $x = vec(u,e)$ that in the model set $\mathcal{M}$ do not have a parametrized link to w_j. Then G_{ji} is generically identifiable in $\mathcal{M}$ from (w,r) if there exists a disconnecting set $\{w_{\mathcal{D}^c}\}$ from $\{\mathcal{X}_j \cup w_i\}$ to $\{w_{\mathcal{D}}^{[j]}\}\backslash\{w_i\}$, subject to $i \notin \mathcal{D}^c$, such that there exist $dim(w_{\mathcal{D}^c})+1$ vertex disjoint paths from $\mathcal{X}_j$ to $\{w_{i\cup\mathcal{D}^c}\}$.*

Now it appears that the conditions for generic single module data-informativity, as formulated in Proposition 9 are extremely closely related to the generic single module identifiability conditions of Theorem 11. While in Proposition 9 external signals in set $\{x_{\mathcal{T}^*}\}$ are excluded from excitation, this similar situation is covered in Theorem 11 by requiring that the disconnecting set $\{w_{\mathcal{D}^c}\}$ should also be a disconnecting set from $\mathcal{X}_j$ to $\{w_{\mathcal{D}}^{[j]}\}\backslash\{w_i\}$. I.e. external signals that reach $\{w_{\mathcal{D}}^{[j]}\}\backslash\{w_i\}$ without passing through $\{w_{\mathcal{D}^c}\}$ are excluded. This set is exactly covered by $\{w_{\mathcal{T}}\}$.

The single difference between the conditions in the two results, is in terms of the white noise signals e that are available for excitation. While in Theorem 11 only white noise signals with a parametrized link to w_j are excluded, in Proposition 9 these are the white noise signals with a parametrized path to $w_{\mathcal{Y}}$. In the situation of a single output (MISO) predictor model these sets are exactly the same. However in the multi-output case, there is apparently some conservatism in the direct identification method. Choosing a multiple output predictor model for identification of a single module is typically done for coping with confounding variables, see e.g. Ramaswamy and Van den Hof (2021). Increasing the number of outputs apparently comes at the cost of more strict data-informativity conditions. This is in line with the data-informativity conditions for consistency of full predictor models, as presented in Ramaswamy and Van den Hof (2021), Van den Hof et al. (2023). The multi-step method (Fonken et al., 2023) does not seem to suffer from this conservatism, due to an alternative handling of confounding variables.

9. Examples

9.1. Two-node example

We consider the 2-node example as depicted in Fig. 4 (Van den Hof & Ramaswamy, 2020) with target module G_{21}.

Because of the confounding variables that occur due to the correlated disturbances, we include node signal w_1 in the output, leading to a predictor model $w_1 \to (w_1, w_2)$. In the original reasoning of the local direct method, data-informativity is analyzed for the full SIMO model, leading to the situation that neither u_1 nor u_2 can be used for excitation of input w_1. This is due to the fact that in modeling output w_1 in the predictor model, and since w_2 is not an input signal, both u_1 and u_2 appear in the predictor

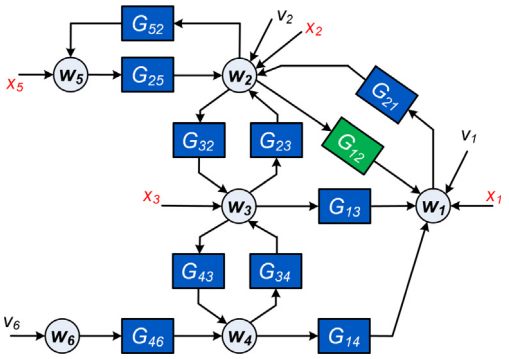


Fig. 5. Six node example with target module $G_{12}(q)$ and correlated disturbances v_1, v_2.

expression for w_1 with unknown dynamic terms, reflecting the sensitivity of the feedback loop between w_1 and w_2. The solution was then to move to a 2-input 2-output model $(w_1, w_2) \to (w_1, w_2)$ which then satisfies the data-informativity conditions when both u_1 and u_2 are present (Van den Hof et al., 2023). In the analysis of the current paper, we consider the predictor model $w_1 \to (w_1, w_2)$, but for data-informativity only need to account for the mapping $w_1 \to w_2$. In this case $w_i = w_1$, $w_j = w_2$, $\mathcal{T} = \emptyset$ and $\mathcal{D}^c = \emptyset$. Since in the predictor model for predicting w_2, u_1 is not present, and u_2 appears with a constant term 1, it follows that $u^{[j]} = \emptyset$, and therefore both u_1 and u_2 can be used to excite w_1. As a result, data-informativity is guaranteed generically if either u_1 or u_2 is present.

9.2. Six-node example

We consider the 6-node example with target module G_{21}, presented in Section 2, while for ease of reference we repeat its figure as Fig. 5. Given that the target module is G_{12} we start composing the predictor model with input w_2 and output w_1. Because of the parallel path connection between w_2 and w_1 through w_3, we add w_3 as predictor input. And in order to handle the confounding variable that is present through the correlated disturbances v_1, v_2, we add w_2 to the predictor output. This leads to a predictor model $(w_2, w_3) \to (w_1, w_2)$, that satisfies the structural conditions for consistently estimating the target module. With this choice, $w_i = w_2$ and $w_{\mathcal{D}^c} = w_3$.

When analyzing the data-informativity conditions as developed in this paper, the vector signal composed of (w_2, w_3) would need appropriate excitation. Let us first consider possible locations (x_5, x_3, x_6) for unmeasured disturbance signals. Possible disturbance signals $x_5 = e_5$ and $x_6 = e_6$ will not provide excitation for (w_2, w_3) as these disturbance signals have unmeasured paths to one of the output signals w_2 or w_1. Therefore their white noise sources are part of $e^{\mathcal{Y}}$, and cannot be used for excitation. A disturbance signal $x_3 = e_3$ will lead to excitation for the input signal w_3.

An external excitation signal $x_5 = u_5$, will have an unknown path to output w_2, but in the predictor model equation for the target output w_1 will not appear as input to an unknown term. As a result u_5 will belong to $u^{\perp 1}$ and therefore is available for exciting w_2. The conclusion is that two vertex disjoint paths are present between u_5 and $(u_3$ or $e_3)$ on the one side and $\{w_2, w_3\}$ on the other side, thereby satisfying the conditions for target module data-informativity of G_{12}.

When applying the earlier results for full predictor model data-informativity, as presented in Van den Hof et al. (2023), the result would be that the present predictor model cannot satisfy the data-informativity conditions by adding a sufficient number

of external signals only. Relying on the result of Proposition 5 in Van den Hof et al. (2023), we would need to extend the predictor model to $(w_1, w_2, w_3, w_5) \rightarrow (w_1, w_2)$, thus adding w_5 as input node signal to be measured, and also adding w_1 to the predictor input. This is based on the result of the mentioned Proposition, indicating that each loop in the network around nodes that appear both in the input and the output, should pass through an input node. Sufficient conditions for (full predictor model) data-informativity would then be that w_5 and w_3 are excited by either a noise signal (e_3, e_5) or an excitation signals (u_3, u_5), and that both w_1 and w_2 are excited by an external excitations signals u_1 and u_2. This illustrates how single module data-informativity can lead to significantly more relaxed conditions for instrumenting the network with sensors and actuators.

10. Conclusions

It has been shown that data-informativity conditions for MIMO models can be relaxed if we are only interested in identifying a single target module (transfer function) in the MIMO model, while using a direct prediction error method. Only a subset of input signals require excitation, where the subset is constructed on the basis of a disconnecting set property. Starting from an open-loop situation, the results have been formulated for generally interconnected dynamic networks, and graph-based conditions have been developed that guarantee satisfaction of the conditions in a generic sense. The resulting path-based data-informativity conditions are shown to be closely related to earlier derived conditions for generic identifiability of single modules in a dynamic network. This implies that in specified situations, satisfaction of the generic single module identifiability conditions guarantees that the local direct method provides an algorithm for consistently estimating the target module.

Acknowledgments

The authors gratefully acknowledge discussions with Xavier Bombois that contributed to the results of this paper.

Appendix A. Proof of Proposition 3

With the decomposition of input signal u according to (14), and adding the inputs $u_{\mathcal{F}}$ to non-parametrized terms, we can write the expression (13) for u by reordering its rows as (leaving out time arguments (t) for brevity):

$$\begin{bmatrix} u_i \\ u_{\mathcal{D}^c} \\ u_{\mathcal{T}} \\ u_{\mathcal{F}} \end{bmatrix} = \begin{bmatrix} 0 & T_{12}(q) & T_{13}(q) & T_{14}(q) \\ T_{21}(q) & T_{22}(q) & T_{23}(q) & T_{24}(q) \\ T_{31}(q) & T_{32}(q) & T_{33}(q) & T_{34}(q) \\ T_{41}(q) & T_{42}(q) & T_{43}(q) & T_{44}(q) \end{bmatrix} \begin{bmatrix} u_i \\ u_{\mathcal{D}^c} \\ u_{\mathcal{T}} \\ u_{\mathcal{F}} \end{bmatrix} + R_x x. \quad \text{(A.1)}$$

The components $u_{\mathcal{F}}$ can be removed (immersed) from the description through Gaussian elimination, leading to the expression

$$\begin{bmatrix} u_i \\ u_{\mathcal{D}^c} \\ u_{\mathcal{T}} \end{bmatrix} = \begin{bmatrix} 0 & \tilde{T}_{12}(q) & \tilde{T}_{13}(q) \\ \tilde{T}_{21}(q) & \tilde{T}_{22}(q) & \tilde{T}_{23}(q) \\ \tilde{T}_{31}(q) & \tilde{T}_{32}(q) & \tilde{T}_{33}(q) \end{bmatrix} \begin{bmatrix} u_i \\ u_{\mathcal{D}^c} \\ u_{\mathcal{T}} \end{bmatrix} + \tilde{T}_x(q) x. \quad \text{(A.2)}$$

The components of x that appear in the equation for u_k, $k \in \{i \cup \mathcal{D}^c \cup \mathcal{T}\}$ are then the x components that are directly added to the concerned nodes in the original system representation, plus all those components x that only pass through nodes in $\{u_{\mathcal{F}}\}$ before they reach $\{u_k\}$. This is a direct result of Gaussian elimination (immersion) of the nodes $\{u_{\mathcal{F}}\}$ (Dankers et al., 2016).

Because of the disconnecting set property of $\{u_{\mathcal{D}^c}\}$, it follows that $\tilde{T}_{31}(q) = 0$. As a result, the third block row of the expression can be rewritten as:

$$\begin{aligned} u_{\mathcal{T}}(t) = {} & (I - \tilde{T}_{33}(q))^{-1} \tilde{T}_{32}(q) u_{\mathcal{D}^c}(t) + \\ & + (I - \tilde{T}_{33}(q))^{-1} \begin{bmatrix} 0 & 0 & I \end{bmatrix} \tilde{T}_x(q) x(t). \end{aligned} \quad \text{(A.3)}$$

The components of x that appear in this equation, are the x-components that are directly added to $u_{\mathcal{T}}$, plus the components of x that are directly added to $u_{\mathcal{F}}$ and only pass through nodes in $\{u_{\mathcal{F}}\}$ before they reach $\{u_{\mathcal{T}}\}$. Differently formulated: the x-components that affect the equation for $u_{\mathcal{T}}$ are those x components that directly enter nodes $\{u_{\mathcal{T} \cup \mathcal{F}}\}$ and that have a path to $\{u_{\mathcal{T}}\}$ without passing through $\{u_{i \cup \mathcal{D}^c}\}$. Since any path from u_i to $\{u_{\mathcal{T}}\}$ has to pass through $\{u_{\mathcal{D}^c}\}$, it is sufficient to check the passing through $\{u_{\mathcal{D}^c}\}$. □

Appendix B. Proof of Theorem 4

We consider the expression $\Delta \bar{M}_j(q, \theta) \kappa^{[j]}(t)$, where $\bar{M}_j$ is constructed from M_j by removing the non-parametrized terms, and we decompose the expression as

$$\Delta \bar{M}_j(q, \theta) \kappa^{[j]}(t) = \Delta \bar{M}_j^u(q, \theta) \kappa_u^{[j]}(t) + \Delta \bar{M}_j^e(q, \theta) \kappa_e^{[j]}(t). \quad \text{(B.1)}$$

With an appropriate decomposition of $\Delta \bar{M}_j^u(q, \theta)$, corresponding to (14) according to:

$$\Delta \bar{M}_j^u(q, \theta) = \begin{bmatrix} \Delta G_{ji}(q, \theta) & \Delta_j^a(q, \theta) & \Delta_j^b(q, \theta) \end{bmatrix} \quad \text{(B.2)}$$

as

$$\Delta \bar{M}_j^u(q, \theta) \kappa_u^{[j]} = \begin{bmatrix} \Delta G_{ji}(q, \theta) & \Delta_j^a(q, \theta) & \Delta_j^b(q, \theta) \end{bmatrix} \begin{bmatrix} u_i \\ u_{\mathcal{D}^c} \\ u_{\mathcal{T}} \end{bmatrix}. \quad \text{(B.3)}$$

Since according to (15), $u_{\mathcal{T}}$ can be written as a function of $u_{\mathcal{D}^c}$ and $x_{\mathcal{T}^*}$, substituting this in the above expression provides

$$\begin{bmatrix} \Delta G_{ji}(q, \theta) & \Delta_j^a(q, \theta) + \Delta_j^b(q, \theta) T_s(q) & \Delta_j^b(q, \theta) R_s(q) \end{bmatrix} \cdot \begin{bmatrix} u_i \\ u_{\mathcal{D}^c} \\ x_{\mathcal{T}^*} \end{bmatrix}.$$

Denote $v := \begin{bmatrix} u_i \\ u_{\mathcal{D}^c} \end{bmatrix}$. Then we can decompose v according to $v = v_{x_{\mathcal{T}^*}} + v_{\perp x_{\mathcal{T}^*}}$ and using the fact that $\Delta_j^b(q, \theta)$ is parametrized independently from ΔG_{ji} and Δ_j^a, we can rewrite $\Delta \bar{M}_j^u(q, \theta) \kappa_u^{[j]}$ as

$$= \begin{bmatrix} \Delta G_{ji}(q, \theta) & \Delta_j^c(q, \theta) & | & \Delta_j^b(q, \theta) R_s(q) \end{bmatrix} \begin{bmatrix} v_{x_{\mathcal{T}^*}} + v_{\perp x_{\mathcal{T}^*}} \\ x_{\mathcal{T}^*} \end{bmatrix}$$

where $\Delta_j^c(q, \theta)$ replaces the term $\Delta_j^a(q, \theta) + \Delta_j^b(q, \theta) T_s(q)$. The expression for $\Delta \bar{M}_j^u(q, \theta) \kappa_u^{[j]}$ is now written as a summation of two terms that are mutually uncorrelated: one driven by $x_{\mathcal{T}^*}$ and $v_{x_{\mathcal{T}^*}}$ and one driven by its orthogonal complement. According to (B.1), for evaluating $\Delta \bar{M}_j(q, \theta) \kappa^{[j]}(t)$ a third term is added, given by $\Delta \bar{M}_j^e(q, \theta) \kappa_e^{[j]}$, leading to

$$\begin{aligned} \Delta \bar{M}_j(q, \theta) \kappa^{[j]}(t) = {} & \begin{bmatrix} \Delta G_{ji}(q, \theta) & \Delta_j^c(q, \theta) & | & \Delta_j^b(q, \theta) R_s(q) \end{bmatrix} \\ & \cdot \left(\begin{bmatrix} v_{x_{\mathcal{T}^*}} \\ x_{\mathcal{T}^*} \end{bmatrix} + \begin{bmatrix} v_{\perp x_{\mathcal{T}^*}} \\ 0 \end{bmatrix} \right) + \Delta \bar{M}_j^e(q, \theta) \kappa_e^{[j]}. \end{aligned}$$

Since both terms $v_{\perp x_{\mathcal{T}^*}}$ and $\kappa_e^{[j]}$ are orthogonal to $x_{\mathcal{T}^*}$, it follows now that $\bar{\mathbb{E}} \| \Delta \bar{M}_j(q, \theta) \kappa^{[j]}(t) \|^2 = 0$ implies $\begin{bmatrix} \Delta G_{ji}(q, \theta) & \Delta_j^c(q, \theta) & \Delta \bar{M}_j^e(q, \theta) \end{bmatrix} = 0$ under the condition that $\Phi_{\mu^{[j]}}(\omega) \succ 0$ for almost all ω with $\mu^{[j]}$ given by

$$\mu^{[j]} = \begin{bmatrix} v_{\perp x_{\mathcal{T}^*}} \\ \kappa_e^{[j]} \end{bmatrix},$$

as specified in (17). This implies $\Delta G_{ji}(q,\theta) = 0$ in the minimum of the quadratic cost function. □

Appendix C. Proof of Theorem 8

The proof follows a reasoning that is similar to the proof of Theorem 4. For completeness we include it here.

We consider the expression $\Delta\bar{M}_j(q,\theta)\kappa^{[j]}(t)$, where $\bar{M}_j$ is constructed from M_j by removing the non-parametrized terms, and we decompose the expression as

$$\Delta\bar{M}_j(q,\theta)\kappa^{[j]}(t) = \Delta\bar{M}_j^w(q,\theta)w_{\mathcal{D}}^{[j]}(t) + \Delta\bar{M}_j^{\xi}(q,\theta)\xi_{\mathcal{Y}}^{[j]}(t) + \Delta\bar{M}_j^u(q,\theta)u^{[j]}(t). \tag{C.1}$$

As a consequence of the definition of $w_{\mathcal{D}^c}$ and the resulting $w_{\mathcal{T}}$, following Definition 6, the term $\Delta\bar{M}_j^w(q,\theta)w_{\mathcal{D}}^{[j]}$ can be written as

$$\Delta\bar{M}_j^w(q,\theta)w_{\mathcal{D}}^{[j]} = \begin{bmatrix}\Delta\bar{G}_{ji}(q,\theta) & \Delta_j^a(q,\theta) & \Delta_j^b(q,\theta)\end{bmatrix}\begin{bmatrix} w_i \\ w_{\tilde{\mathcal{D}}^c} \\ w_{\mathcal{T}}\end{bmatrix}$$

with $\{w_{\tilde{\mathcal{D}}^c}\} = \{w_{\mathcal{D}^c}\} \cap \{w_{\mathcal{D}}^{[j]}\}$. By complementing $w_{\tilde{\mathcal{D}}^c}$ with the remaining components of $w_{\mathcal{D}^c}$ and adding 0 entries to the accompanying $\Delta_j^a(q,\theta)$, this can be rewritten as

$$\Delta\bar{M}_j^w(q,\theta)w_{\mathcal{D}}^{[j]} = \begin{bmatrix}\Delta\bar{G}_{ji}(q,\theta) & \tilde{\Delta}_j^a(q,\theta) & \Delta_j^b(q,\theta)\end{bmatrix}\begin{bmatrix} w_i \\ w_{\mathcal{D}^c} \\ w_{\mathcal{T}}\end{bmatrix}.$$

Since according to (27), $w_{\mathcal{T}}$ can be written as a function of $w_{\mathcal{D}^c}$ and $x_{\mathcal{T}^*}$, substituting this in the above expression provides

$$\begin{bmatrix}\Delta\bar{G}_{ji}(q,\theta) & \tilde{\Delta}_j^a(q,\theta) + \Delta_j^b(q,\theta)T_s(q) & \Delta_j^b(q,\theta)R_s(q)\end{bmatrix}\cdot \begin{bmatrix} w_i \\ w_{\mathcal{D}^c} \\ x_{\mathcal{T}^*}\end{bmatrix}.$$

Denote $\nu := \begin{bmatrix} w_i \\ w_{\mathcal{D}^c}\end{bmatrix}$ and $\chi := \{x_{\mathcal{T}^*} \cup u^{[j]}\}$. Then we can decompose ν according to $\nu = \nu_\chi + \nu_{\perp\chi}$ and using the fact that $\Delta_j^b(q,\theta)$ is parametrized independently from ΔG_{ji} and $\tilde{\Delta}_j^a$, we can rewrite $\Delta\bar{M}_j^w(q,\theta)w_{\mathcal{D}}^{[j]}$ as

$$= \begin{bmatrix}\Delta G_{ji}(q,\theta) & \Delta_j^c(q,\theta) & | & \Delta_j^b(q,\theta)R_s(q)\end{bmatrix}\cdot\begin{bmatrix}\nu_\chi + \nu_{\perp\chi} \\ x_{\mathcal{T}^*}\end{bmatrix}$$

The expression for $\Delta\bar{M}_j^u(q,\theta)w_{\mathcal{D}}^{[j]}$ is now written as a summation of two terms that are mutually uncorrelated: one driven by $x_{\mathcal{T}^*}$ and ν_χ and one driven by a term that is orthogonal to this. For evaluating $\Delta\bar{M}_j(q,\theta)\kappa^{[j]}(t)$ two additional terms are added, given by $\Delta\bar{M}_j^{\xi}(q,\theta)\xi_{\mathcal{Y}}^{[j]}$ and $\Delta\bar{M}_j^u(q,\theta)u^{[j]}$.

The first of these two terms is possible correlated with $\nu_{\perp\chi}$, while the second is not.

It follows now that $\bar{\mathbb{E}}\|\Delta\bar{M}_j(q,\theta)\kappa^{[j]}(t)\|^2 = 0$ implies $\begin{bmatrix}\Delta\bar{G}_{ji}(q,\theta) & \Delta_j^c(q,\theta) \mid \Delta\bar{M}_j^{\xi}(q,\theta)\end{bmatrix} = 0$ under the condition that $\Phi_{\eta^{[j]}}(\omega) \succ 0$ for almost all ω with $\eta^{[j]}$ given by

$$\eta^{[j]} = \begin{bmatrix}\nu_{\perp\chi} \\ \xi_{\mathcal{Y}}^{[j]}\end{bmatrix},$$

as specified in (28). This implies that $\Delta\bar{G}_{ji}(q,\theta) = 0$ in the minimum of the quadratic cost function. □

Appendix D. Proof of Proposition 9

When applying the graph-based results of Proposition 5 to the results of Theorem 8, the graph-based condition becomes to having $dim(\eta^{[j]})$ vertex disjoint paths from all external signals $\{u, e\}$ to $\{\eta^{[j]}\}$. With the expression (28) for $\eta^{[j]}$ this becomes: $dim(\eta^{[j]})$ vertex disjoint paths from $\{u, e\}\backslash\{\chi\}$ to $\{w_{i\cup\mathcal{D}^c}, \xi_{\mathcal{Y}}^{[j]}\}$.

The white noise signals $\xi_{\mathcal{Y}}^{[j]}$ are typically constructed after spectral factorization of the disturbance process affecting $w_{\mathcal{Y}}$. If in the predictor model, the noise model $H(q,\theta)$ is parametrized fully, this implies that every noise signal $e_\ell \in e^{\mathcal{Y}}$, that in the network has a path to a node $w_{\mathcal{Y}}$ with unknown dynamics (i.e. the dynamics in the path is not fixed a priori), will potentially contribute to $\xi_{\mathcal{Y}}^{[j]}$. Note that this does not only refer to noise signals e that have a direct link to $w_{\mathcal{Y}}$ in the network; it also involves noise terms e on unmeasured nodes that are not part of the predictor model. If $H(q,\theta)$ has a structured parametrization, like e.g. a block diagonal form, then we can restrict the class of e-signals that should be excluded from the subset of $e_{\mathcal{Y}}$ to those signals e_ℓ that have a path to an output $w_k \in w_{\mathcal{Y}}$ of which the disturbance component is correlated to the disturbance component of w_j.

In the vertex disjoint path condition, this set of signals $e^{\mathcal{Y}}$ is then required to establish the positive definite spectrum condition for $\xi_{\mathcal{Y}}^{[j]}$, which is satisfied by definition. The remaining term then becomes: $dim(\mathcal{D}^c) + 1$ vertex disjoint paths between $\{u, e\}\backslash\{e^{\mathcal{Y}}, \chi\}$ and $\{w_{i\cup\mathcal{D}^c}\}$. Or equivalently $dim(\mathcal{D}^c) + 1$ vertex disjoint paths between $\{u\backslash u^{[j]}, e\backslash e^{\mathcal{Y}}\}\backslash\{x_{\mathcal{T}^*}\}$ and $\{w_{i\cup\mathcal{D}^c}\}$, which proves the result.

Appendix E. Proof of Corollary 10

External excitation signals that are directly entering input nodes $w_{\mathcal{D}}$ will contribute to $u^{\perp j}$, except when there is a confounding variable between the concerned input and $w_{\mathcal{Y}}$. The input node is then classified to belong to set $\{w_{\mathcal{B}}\}$, Van den Hof et al. (2023). For the data-informativity conditions, one can always construct a disconnecting set $\{w_{\mathcal{D}^c}\}$ consisting of nodes that are out-neighbors of w_i. For satisfying the structural conditions a. and b., nodes in $\{w_{\mathcal{B}}\}$ are not allowed to be out-neighbors of w_i. As a result there always exist a disconnecting set that does not contain nodes in $\{w_{\mathcal{B}}\}$, and external excitation on these nodes will always appear in $u^{\perp j}$.

References

Bombois, X., Colin, K., Van den Hof, P. M. J., & Hjalmarsson, H. (2023). On the informativity of direct identification experiments in dynamical networks. *Automatica*, *148*(110742).

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), 692–705.

Dankers, A. G., Van den Hof, P. M. J., Heuberger, P. S. C., & Bombois, X. (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. J. M., Ramaswamy, K. R., & Van den Hof, P. M. J. (2022). A scalable multi-step least squares method for network identification with unknown disturbance topology. *Automatica*, *141*(110295).

Fonken, S. J. M., Ramaswamy, K. R., & Van den Hof, P. M. J. (2023). Local identification in dynamic networks using a multi-step least squares method. In *Proc. 62th IEEE conf. on decision and control* (pp. 431–436). Singapore: IEEE.

Gevers, M., Bazanella, A., & Miskovic, L. (2008). Informative data: how to get just sufficiently rich? In *Proc. 47th IEEE conf. decision and control* (pp. 1962–1967). Cancun, Mexico: IEEE.

Gevers, M., Bazanella, A. S., & Vian da Silva, G. (2018). A practical method for the consistent identification of a module in a dynamical network. *IFAC-PapersOnLine*, *51–15*, 862–867, Proc. 18th IFAC Symp. System Identif. (SYSID2018).

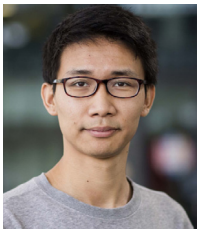
Gonçalves, J., & Warnick, S. (2008). Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks. *IEEE Transactions on Automatic Control*, *53*(7), 1670–1674.

Haber, A., & Verhaegen, M. (2014). Subspace identication of large-scale interconnected systems. *IEEE Transactions on Automatic Control*, *59*(10), 2754–2759.

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.
Legat, A., & Hendrickx, J. M. (2020). Local network identifiability with partial excitation and measurement. In *Proc. 59th IEEE conf. on decision and control* (pp. 4342–4347). Jeju Island, Republic of Korea: IEEE.
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. (1978). Convergence analysis of parametric identification methods. *IEEE Transactions on Automatic Control*, *23*, 770–783.
Ljung, L. (1999). *System identification: Theory for the user*. Englewood Cliffs, NJ: Prentice-Hall.
Mapurunga, E., Gevers, M., & Bazanella, A. S. (2024). Identifiability of dynamic networks: the essential role of dources and dinks. *IEEE Transactions on Automatic Control*, *69*(2), 1303–1308.
Materassi, D., & Innocenti, G. (2010). Topological identification in networks of dynamical systems. *IEEE Transactions on Automatic Control*, *55*(8), 1860–1871.
Materassi, D., & Salapaka, M. (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.
Pearl, J. (2000). *Causality: Models, reasoning, and inference*. New York: Cambridge University Press.
Pintelon, R., & Schoukens, J. (2012). *System identification: a frequency domain approach*. John Wiley & Sons.
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), 5237–5252.
Shi, S., Cheng, X., & Van den Hof, P. M. J. (2022). Generic identifiability of subnetworks in a linear dynamic network: the full measurement case. *Automatica*, *137*(110093).
Shi, S., Cheng, X., & Van den Hof, P. M. J. (2023). Single module identifiability in linear dynamic networks with partial excitation and measurement. *IEEE Transactions on Automatic Control*, *68*(1), 285–300.
Söderström, T., & Stoica, P. (1989). *System Identification*. Hemel Hempstead, UK: Prentice Hall.
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 den Hof, P. M. J., & Ramaswamy, K. R. (2020). Path-based data-informativity conditions for single module identification in dynamic networks. In *Proc. 59th IEEE conf. on decision and control* (pp. 4354–4359). Jeju Island, Republic of Korea: IEEE.
Van den Hof, P. M. J., Ramaswamy, K. R., & Fonken, S. J. M. (2023). Integrating data-informativity conditions in predictor models for single module identification in dynamic networks. *IFAC-PapersOnLine*, *56–2*, 2377–2382, Proc. 22nd IFAC World Congress, 9-14 July, 2023, Yokohama, Japan.
Van den Hof, P. M. J., Shi, S., Weerts, H. H. M., Cheng, X., Ramaswamy, K. R., Dankers, A. G., Dreef, H. J., Fonken, S. J. M., Steentjes, T. R. V., & Meijer, J. B. T. (2024). SYSDYNET - A MATLAB app and toolbox for dynamic network identification. *IFAC-PapersOnLine*, *56–2*, 574–579, Proc. 20th IFAC Symp. System Identification, 17-18 July 2024, Boston, MA, USA.
van der Woude, J. W. (1991). A graph-theoreric characterization for the rank of the transfer matrix of a structured system. *Mathematics of Control, Signals, and Systems*, *4*, 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.
Veedu, M. S., Doddi, H., & Salapaka, M. V. (2022). Topology learning of linear dynamical systems with latent nodes using matrix decomposition. *IEEE Transactions on Automatic Control*, *67*(11), 5746–5761.
Verhaegen, M., Yu, C., & Sinquin, B. (2022). *Data-driven identification of networks of dynamic systems*. Cambridge, UK: Cambridge University Press.
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.
Yuan, Y., Stan, G.-B., Warnick, S., & Gonçalves, J. (2011). Robust dynamical network structure reconstruction. *Automatica*, *47*, 1230–1235.
Zorzi, M., & Chiuso, A. (2017). Sparse plus low rank network identification: a nonparametric approach. *Automatica*, *76*, 355–366.



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 Life Fellow, Honorary Member of the Hungarian Academy of Sciences, ACA Fellow 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, 2022–2024), and an Associate Editor and Editor of Automatica (1992–2005). In the triennium 2017–2020 he served as Vice-President of IFAC.



Shengling Shi is an Assistant Professor with the Delft Center for Systems and Control (DCSC), Delft University of Technology (TU Delft), The Netherlands. He was a postdoctoral researcher at the Department of Chemical Engineering, Massachusetts Institute of Technology, USA, and at DCSC, TU Delft. He received the M.Sc. degree (with distinction) in 2017 and the Ph.D. degree in 2021, both from the Control Systems Group, Eindhoven University of Technology, The Netherlands. His research interests include system identification, machine learning, optimal control, and their applications to biological and bio-inspired systems.



Stefanie J.M. Fonken received her B.Sc. degree in Mechatronic Engineering in 2014 from Fontys University of Applied Sciences in Eindhoven. After her Bachelor she worked as a Machine Engineer at Bosch Transmission Technology for 3 years. In 2020, she obtained the M.Sc. degree in Systems and Control (with distinction cum laude) from Eindhoven University of Technology (TU/e). She continued to work temporarily as a Ph.D. candidate on the System Identification in Dynamic Networks (SYSDYNET) project in the Control Systems group of TU/e. She focused on multi-step methods and data-informativity conditions for single module identification in dynamic networks. Since 2025, she has been working as an R&D Expert Machine Learning for (Thermal) Energy Systems at EnergyVille/VITO.



Karthik Raghavan was born in Chennai, India in 1989. He received his Bachelor of Engineering in Electrical and Electronics (with distinction) in 2011 from Anna University, India. He worked as Control & Automation engineer at Larsen & Toubro for 4 years after his bachelors. During the period of his work, he completed his MBA degree. In 2017, he received his M.Sc. degree in Systems and Control (with great appreciation) at Eindhoven University of Technology. During the period 2017-2022, Karthik worked on his Ph.D. research and was awarded the Ph.D. degree (with honors cum laude) from Eindhoven University of Technology in May 2022. Since March 2022 he is employed by ASML, Veldhoven, The Netherlands, currently at ASML Research as a Mechatronics and Control Researcher. His research interests are leveraging Artificial Intelligence and machine learning for data-driven modeling of complex interconnected systems, system identification, and dynamic network identification.



Håkan Hjalmarsson received the M.S. degree in Electrical Engineering in 1988, and the Licentiate degree and the Ph.D. degree in Automatic Control in 1990 and 1993, respectively, all from Linköping University, Sweden. He has held visiting research positions at California Institute of Technology, Louvain University and at the University of Newcastle, Australia. He has served as an Associate Editor for Automatica (1996-2001), and IEEE Transactions on Automatic Control (2005-2007) and been Guest Editor for European Journal of Control, Control Engineering Practice and IFAC Journal of Systems and Control. He is Professor at the Department of Decision and Control Systems, School of Electrical Engineering and Computer Science, KTH, Stockholm, Sweden and also affiliated with the Competence Centre for Advanced BioProduction by Continuous Processing, AdBIOPRO. He is an IEEE Fellow and past Chair of the IFAC Coordinating Committee CC1 Systems and Signals. In 2001 he received the KTH award for outstanding contribution to undergraduate education. He was General Chair for the IFAC Symposium on System Identification held in 2018. His research interests include system identification, learning of dynamical systems for control, process modeling control and estimation in communication networks.



Arne Dankers is currently an Assistant Professor at the University of Calgary. His research interests include network system identification with applications in monitoring and control. He obtained his Ph.D. at the Delft University of Technology in 2014. After receiving his Ph.D. he worked at Hifi Engineering applying network system identification to monitoring of oil pipelines for leaks and causes of leaks. He also worked at Attabotics on data driven methods to optimize throughput in automated warehousing.