Path-Based Conditions for the Identifiability of Non-additive Nonlinear Networks with Full Measurements

Renato Vizuete and Julien M. Hendrickx

***Abstract*—We analyze the identifiability of nonlinear networks with node dynamics characterized by functions that are non-additive. We consider the full measurement case (all the nodes are measured) in the path-independent delay scenario where all the excitation signals of a specific node have the same delay in the output of a measured node. Based on the notion of a generic nonlinear matrix associated with the network, we introduce the concept of generic identifiability and characterize the space of functions that satisfies this property. For directed acyclic graphs (DAGs) characterized by analytic functions, we derive a sufficient condition for identifiability based on vertex-disjoint paths from excited nodes to the in-neighbors of each node in the network. Furthermore, when we consider the class of polynomial functions, by using well-known results on algebraic varieties, we prove that the vertex-disjoint path condition is also necessary. Finally, we show that this identifiability condition is not necessary for the additive nonlinear model. Some examples are added to illustrate the results.**

***Index Terms*—Network analysis and control, system identification.**

I. Introduction

Identifiability is a fundamental property in the identification of networked systems [1], [2] that allows us to determine which nodes must be excited and measured to identify all the existing dynamics in the system. It is commonly assumed that the network topology is known, which allows obtaining conditions for identifiability based primarily on the structure of the network. When linear dynamics are located at the level of the edges, the identifiability conditions have been fully characterized in the case of full excitation or full measurement [3]–[5]. In the partial excitation and measurement case, several identifiability conditions have been derived in [6]–[9], but necessary and sufficient conditions are still not known.

For more complex interactions [10]–[13], most of the dynamics in the edges are nonlinear and the identifiability conditions turn out to be in many cases different, and sometimes weaker, than in the linear case due to the lesser risk of ambiguities created by linear superposition. For the full excitation case where nonlinear terms are present, identifiability conditions based on the measurement of sinks were obtained in [14], [15], which are weaker than the linear counterpart.

This work was supported by F.R.S.-FNRS via the *KORNET* project, and by the *SIDDARTA* Concerted Research Action (ARC) of the Fédération Wallonie-Bruxelles.

R. Vizuete and J. M. Hendrickx are with ICTEAM institute, UCLouvain, B-1348, Louvain-la-Neuve, Belgium. R. Vizuete is a FNRS Postdoctoral Researcher - CR. renato.vizueteharo@uclouvain.be, julien.hendrickx@uclouvain.be.

Previous identifiability conditions for nonlinear networks have been derived for additive models. In linear networks, this assumption is always satisfied due to the superposition principle, which restricts the type of networks that the dynamics can model. In the nonlinear case, a model that is not additive allows us to encompass richer behaviors in more complex networks. This type of model can be found in neural networks [16], opinion dynamics [13], interconnection of Wiener models [17], complex networks [18], among others. In this case, the output of a node is a function of the output of in-neighbors that cannot be necessarily separated into edge functions. Therefore, the identifiability problem is formulated in terms of node dynamics [19] instead of edge dynamics as in the additive case. This implies that the class of functions that have been previously considered in [14], [15], [20] can be extended since the presence of a static component can be included without ambiguity in the node dynamics.

In most networks, the excitation of nodes requires the use of actuators, while the measurement implies the use of sensors. Usually, the cost associated with an actuator is greater than the cost associated with a sensor. Moreover, the implementation of an actuator could involve a temporary disconnection of the network, while the use of a sensor typically has less impact. This implies that from a practical point of view, the full measurement case is more relevant for identification. However, in the linear case, identifiability in the full measurement case was not usually analyzed since the conditions were obtained from the full excitation case thanks to the symmetry of the model. Unfortunately, in [15], it was shown that this symmetry does not hold in the nonlinear case, which emphasizes the necessity of an independent analysis for the full measurement case.

In the full excitation case [14], [15], the identifiability conditions are sufficient and necessary for any function in the specific class. However, for some graph topologies, it is possible that some identifiability conditions are valid except for a particular choice of parameters in the dynamics. These types of results are known as *generic* [21], where the identifiability is guaranteed for almost all parameters such as in the linear case [4], [8], [9]. However, in the nonlinear case, the functions belong to a space of infinite dimension, and a proper characterization of this notion of genericity is fundamental to determine what a zero measure set encompasses [19].

A preliminary version of this paper was presented in [20] where a sufficient condition for the identifiability of DAGs was derived in terms of vertex-disjoint paths when the model

is additive nonlinear. Furthermore, it was shown that certain identifiability conditions might not hold for particular choices of functions in the edges, similarly to the generic identifiability in linear networks. In [20], a sufficient condition for the identifiability of DAGs with partial excitation and measurement was derived based on the notion of a generic nonlinear matrix $J_G(\mathbf{u}_G^k)$, where $\mathbf{u}_G^k \in \mathbb{R}^{|\mathcal{N}^e|}$. However, a formal characterization of the space of functions that satisfies this notion of *genericity* was missing. In this document, we provide a rigorous definition of genericity based on the nonlinear matrix $J_G(\mathbf{u}_G^k)$ that will play an important role in the proof of Theorem 1. By contrast, in this work, we derive similar identifiability conditions which are sufficient and necessary for a more general model that is non-additive. In addition, we fully characterize the notion of genericity that was missing in the previous work [20]. Finally, we show that the sufficient condition in [20] is not necessary for the additive model.

The remainder of this article is organized as follows. In Section II, we introduce the model where the unknown dynamics are located at the level of the nodes and we present the class of analytic functions that will be used in this work. In Section III, we introduce the path-independent delay case where all the excitation signals associated with a node that influences the output of a measured node, have the same delay. Based on this particular scenario, we present the notion of generic identifiability where a set of measure zero is fully characterized in terms of a finite number of parameters associated with the analytic functions. In Section IV, we analyze the identifiability problem in DAGs, and we show that a DAG is identifiable if there are vertex-disjoint paths from excited nodes to the in-neighbors of each node in the network. For the class of polynomial functions, we prove that the vertex-disjoint path condition is also necessary by using well-known results in algebraic varieties. This makes a connection between the linear case and the non-additive nonlinear case. In Section V, we show with a counterexample that the sufficient condition derived in this work is not necessary for the additive model analyzed in [20]. Finally, conclusions and future work are presented in Section VI.

II. Problem formulation

A. Model class

We consider a network characterized by a weakly connected digraph $G = (V, E)$ composed by a set of nodes $V = \{1, \dots, n\}$ and a set of edges $E \subseteq V \times V$. The output of each node i in the network is given by a nonlinear dynamics of the form:

$$y_i^k = u_i^{k-1} + \Phi_i(y_1^{k-m_{i,1}}, \dots, y_{M_i}^{k-m_{i,M_i}}), \quad \text{for } y_1, \dots, y_{M_i} \in \mathcal{N}_i, \quad (1)$$

where the superscripts of the inputs and outputs denote the corresponding values at the specific time instants, y_i^k is the state of the node i, u_i^{k-1} is an arbitrary external excitation signal, the delays $m_{i,j} \in \mathbb{Z}^+$ are finite, Φ_i is the node function associated with node i, and $\mathcal{N}_i$ is the set of in-neighbors of node i.

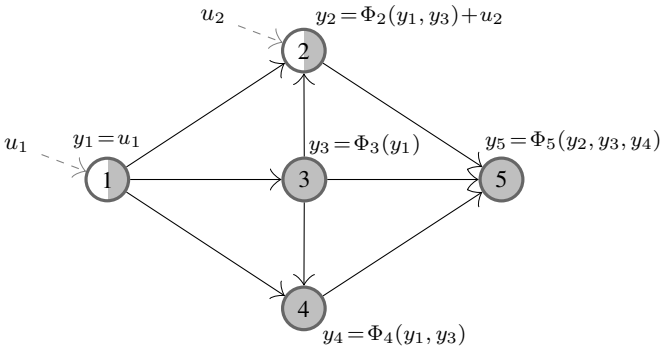

Fig. 1. Model of a network considered for the identification where all the nodes are measured (gray) and some nodes can be excited (white and gray). The dynamics of each node is determined by a nonlinear function Φ_i of the outputs of the in-neighbors.

If a node i is not excited, its corresponding excitation signal u_i is set to zero. The model (1) corresponds to a generalized version of the nonlinear model in [14], [20], where the node function Φ_i is not necessarily additive. Similarly to [20], we assume that the node function Φ_i depends only on a single past value of each in-neighbor $y_j^{m_{i,j}}$, where $m_{i,j}$ can take a value different from 1. Our goal is to analyze the effect of the non-additivity of the model determined by the function Φ_i, in the derivation of identifiability conditions, while the more complex case with dynamics depending on several past values of each in-neighbor is left for future work.

Unlike the additive model in [14], [15], [20], we will assume that the topology of the network G defines the arguments of the function Φ_i. If $(i, j) \in E$ then $\frac{\partial \Phi_i}{\partial y_j} \not\equiv 0$. We do not consider multiple edges between two nodes, since they would be indistinguishable and hence unidentifiable.

Assumption 1. *The graph G associated with the network is known, where the presence of an edge (i, j) implies $\frac{\partial \Phi_i}{\partial y_j} \not\equiv 0$.*

Assumption 1 implies that we know if a node function Φ_i is influenced by the output of a specific node y_j.

In this work, we restrict our attention to networks that do not contain any cycle (i.e., directed acyclic graphs), so that the function F_i associated with the measurement of a node i is of the form:

$$y_i^k = u_i^{k-1} + F_i(u_1^{k-2}, \dots, u_1^{k-M_1}, \dots, u_{n_i}^{k-2}, \dots, u_{n_i}^{k-M_{n_i}}), \quad 1, \dots, n_i \in \mathcal{N}^{e \to i}, \quad (2)$$

where $\mathcal{N}^{e \to i}$ is the set of excited nodes with a path to the node i. The function F_i determines the output of the node i through the dynamics (1) and the excitation signals of the nodes in $\mathcal{N}^{e \to i}$, and only depends on a finite number of inputs (i.e., $M_1, \dots, M_{n_i}$ are finite) due to the finite delays $m_{i,j}$ and the absence of cycles. For the identifiability analysis, we consider the notion of identifiability in system identification [22], where the objective is to determine if there exists a unique set of local dynamics Φ_i that leads to a global behavior given by the functions F_i. Hence, we make the following assumption for the derivation of the identifiability conditions.

Assumption 2. *If the node i is measured, the function F_i is known.*

Remark 1 (Functions Φ_i and F_i)**.** *Although Φ_i and F_i are functions associated with a node i, they are not the same since Φ_i determines the dynamics of i while F_i is the information obtained through the measurement of i. Notice that Φ_i is a function of the outputs of the in-neighbors and is independent of the set of excited nodes $\mathcal{N}^e$, while F_i is a function of the excitation signals and depends on $\mathcal{N}^e$.*

In this work, we will analyze the full measurement case where the set of measured nodes $\mathcal{N}^m$ is given by $\mathcal{N}^m = V$. This setting corresponds to networks where we have access to all the nodes, whose states can generally be measured through an appropriate device like a sensor. Our objective is to determine which nodes need to be excited to identify all the node functions Φ_i. In this way, our aim is to determine the possibility of identification, and not to derive identification algorithms or study identification methods. Fig. 1 presents the framework considered in this work where all the nodes are measured, some nodes are excited and the node functions Φ_i depend on the outputs of the in-neighbors.

B. Generic identifiability

We define the relationships between the measurements of the nodes and the functions Φ_i.

Definition 1 (Set of measured functions)**.** *Given a set of excited nodes $\mathcal{N}^e$, the partially ordered set of measured functions $(F(\mathcal{N}^e), \leq)$ associated with $\mathcal{N}^e$ is given by:*

$$F(\mathcal{N}^e) := \{F_i \mid i \in V\},$$

with $F_i \leq F_j$ if $i \leq j$.

Since the potential functions considered for the analysis of identifiability must in general satisfy some constraints depending on the type of network or application, we restrict the problem to a certain class of functions $\mathcal{F}$: the node functions belong to $\mathcal{F}$ and the identifiability is considered only among the functions belonging to $\mathcal{F}$.

We say that a a digraph G and a partially ordered set of node functions $(\{\Phi\}, \leq) = \{\Phi_i \in \mathcal{F} \mid i \in V\}$ with $\Phi_i \leq \Phi_j$ if $i \leq j$ generate $F(\mathcal{N}^e)$ if the functions $F_i \in F(\mathcal{N}^e)$ are obtained through (1). Since the ordering of the functions only depends on the labeling of the nodes, in the rest of the paper, we will refer to $F(\mathcal{N}^e)$ and $\{\Phi\}$ only as a set of measured functions and a set of node functions respectively.

While in the full excitation case [14], [15], the set of measured functions depends on the set of measured nodes, in the full measurement case, the set of measured functions encompasses all the functions F_i, and its dependence is expressed with respect to the set of excited nodes.

Unlike [14], [15] where the notion of *global identifiability* (i.e., for all functions)) was used to derive the identifiability conditions, in this work, we will introduce the notion of *generic identifiability* (i.e., for almost all functions).

Definition 2 (Generic Identifiability)**.** *Given a set of node functions $\{\Phi\} = \{\Phi_i \in \mathcal{F}\}$ that generates $F(\mathcal{N}^e)$ and any other set of node functions $\{\tilde{\Phi}\} = \{\tilde{\Phi}_i \in \mathcal{F}\}$ that generates $\tilde{F}(\mathcal{N}^e)$. A node function Φ_i is generically identifiable in a class $\mathcal{F}$ if $F(\mathcal{N}^e) = \tilde{F}(\mathcal{N}^e)$, implies that $\Phi_i = \tilde{\Phi}_i$ for any $\tilde{\Phi}_i$ and a finite number of parameters except possibly those lying on a zero measure set. A network G is generically identifiable in a class $\mathcal{F}$ if $F(\mathcal{N}^e) = \tilde{F}(\mathcal{N}^e)$, implies that $\{\Phi_i\} = \{\tilde{\Phi}_i\}$ for any $\{\tilde{\Phi}_i\}$.*

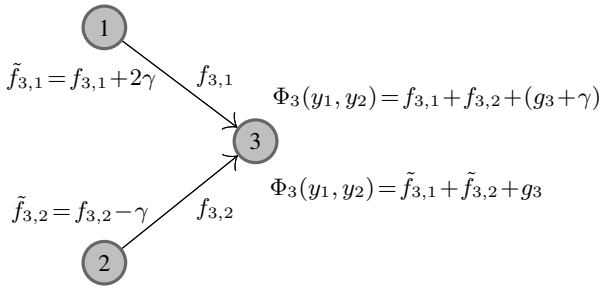

Fig. 2. Each node function Φ_i can be expressed as the sum of edge functions $f_{i,j}$ and a function g_i composed by products of the outputs of the in-neighbors. Unlike the additive model, the presence of a static component γ does not affect the identifiability.

The zero measure set of functions associated with the definition of generic identifiability will be clarified in Section III. Definition 2 implies that a network G is generically identifiable in a class $\mathcal{F}$ if all the node functions Φ_i are generically identifiable.

In this work, we will consider analytic functions with a Taylor series that converges to the function for all $x \in \mathbb{R}^m$ (possibly with a finite radius of convergence).

Definition 3 (Class of functions $\mathcal{F}_A$)**.** *Let $\mathcal{F}_A$ be the class of analytic functions $\Phi : \mathbb{R}^m \to \mathbb{R}$.*

Therefore, in the rest of this work we will assume that $\Phi_i \in \mathcal{F}_A$ for all $i \in V$. This implies that the function Φ_i is analytic and can be expressed as:

$$\Phi_i = \sum_{p_1,\ldots,p_{M_i}=0} a_{p_1,\ldots,p_{M_i}} (y_1^{k-m_{i,1}})^{p_1} \cdots (y_{M_i}^{k-m_{i,M_i}})^{p_{M_i}}. \tag{3}$$

For our convenience, we will separate the univariate part from the rest of the function, so that (3) can be rewritten as[1]:

$$\Phi_i = \sum_{j=1}^{M_i} f_{i,j}(y_j^{k-m_{i,j}}) + g_i(y_1^{k-m_{i,1}}, \ldots, y_{M_i}^{k-m_{i,M_i}}), \tag{4}$$

where the function g_i encompasses all the terms corresponding to the product of two or more variables $y_j^{k-m_{i,j}}$, and a possible constant term, but not the terms depending only on one variable $y_j^{k-m_{i,j}}$. Hence, identifiability of Φ_i is equivalent to identifiability of the functions g_i and $f_{i,j}$ for $j \in \mathcal{N}_i$. By contrast, the additive model is of the form [14], [15]:

$$y_i^k = u_i^{k-1} + \sum_{j \in \mathcal{N}_i} f_{i,j}(y_j^{k-m_{i,j}}), \tag{5}$$

[1]This is just a particular decomposition of the function Φ_i and does not represent the existence of functions $f_{i,j}$ in the edges like in the additive case [14], [15], [20].

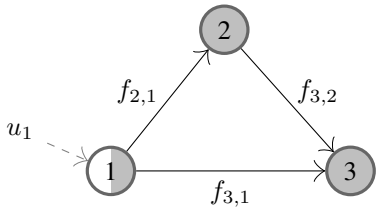

Fig. 3. DAG where a particular choice of delays $m_{2,1}$, $m_{3,2}$ and $m_{3,1}$ might generate a function F_3 with a path-independent delay where the information coming through $f_{3,1}$ and $f_{3,2}$ cannot be distinguished.

where $f_{i,j}$ is a nonlinear function associated with the edge (i,j), and the objective of the identifiability is to determine conditions to identify the functions $f_{i,j}$. If for some functions Φ_i we have $g_i = 0$ for all $i \in V$, then the network is additive. Also, the network becomes additive if the additional constraint $g_i = 0$ for all $i \in V$ is imposed. Unlike the additive model (5), a static component does not affect the identifiability since we are interested on the identifiability of the node function Φ_i and not of the particular decomposition into $f_{i,j}$ and g_i, such that a static component can be included on any of the functions $f_{i,j}$ or g_i without changing the function Φ_i. For instance, in the DAG in Fig. 2, the static component γ of the node function Φ_3 can be considered as part of g_3 or as part of $f_{3,1}$ and $f_{3,2}$ without modifying Φ_3. In our setting, based on the particular decomposition (4), it is assumed that the constant term is included in the function g_i.

III. Generic nonlinear matrix for $g_i = 0$

In this section, we will use the particular decomposition (4) and consider the simpler problem where all the functions $g_i = 0$ (i.e., the additive model). This simplification will be used as a mathematical tool to analyze the more general case of non-additive nonlinear networks. In the subsection III-B, we will introduce some definitions and results associated exclusively to the functions $f_{i,j}$ to explain the notion of *genericity* that does not depend on the functions g_i. The contents of this section will be used in the proof of the main results in Section IV.

A. Path-independent delay

In general, one of the main challenges in the identification of networks is to distinguish the information that arrives to a node i through different paths. For instance, let us consider the graph in Fig. 3 with functions $f_{2,1}(y_1^{k-m_{2,1}})$, $f_{3,2}(y_2^{k-m_{3,2}})$ and $f_{3,1}(y_1^{k-m_{3,1}})$. If $m_{3,1} = m_{2,1} + m_{3,2}$, the measurement of the node 3 provides the output:

$$\begin{aligned} y_3^k &= f_{3,1}(y_1^{k-m_{2,1}-m_{3,2}}) + f_{3,2}(y_2^{k-m_{3,2}}) \\ &= f_{3,1}(y_1^{k-m_{2,1}-m_{3,2}}) + f_{3,2}(f_{2,1}(y_1^{k-m_{2,1}-m_{3,2}})) \\ &= f_{3,1}(u_1^{k-m_{2,1}-m_{3,2}-1}) + f_{3,2}(f_{2,1}(u_1^{k-m_{2,1}-m_{3,2}-1})), \end{aligned} \tag{6}$$

which implies that the information coming through $f_{3,1}$ and $f_{3,2}$ cannot be distinguished for this particular choice of delays $m_{i,j}$.

Proposition 1. *For any DAG G, there exists a particular choice of delays $m_{i,j}$ of the functions $f_{i,j}$ such that all the variables u_j associated to a node j have the same delays in $F_i(u_1^{k-T_{i,1}}, \dots, u_{n_i}^{k-T_{i,n_i}})$.*

Proof. Since any DAG can be sorted in topological order [23], let us consider an arbitrary ordering of the DAG G. Without loss of generality, we relabel the nodes in V according to the topological ordering, and for each edge in E, we set the delay $m_{p,q}$ associated with the nonlinear function $f_{p,q}$ as $m_{p,q} = p - q$. Now, let us consider two arbitrary nodes i and j such that j is excited and reaches i through several paths. Since for a given path, the total delay is given by the sum of the delays of each edge in the path, we have that all paths from j to i present a total delay $i-j$ such that the function F_i associated with the measurement of the node i, will be a function only of the excitation signal $u_j^{k-i-j-1}$. □

Fig. 4 presents a DAG with a topological ordering where the delays $m_{i,j}$ associated with the functions $f_{i,j}$ are chosen according to the position of the nodes in the topological ordering so that the delay configuration in Proposition 1 is satisfied. Proposition 1 shows that for any DAG there exists a particular configuration of delays such that all the excitation signals associated with a node in a function F_i have the same delay. However, this configuration is not unique, and there could be an infinite number of combinations of delays that lead to the same setting. For instance, if in the topological ordering in the proof of Proposition 1, we set for every edge $(p,q) \in V$, the delay $m_{p,q}$ as $k(p-q)$ with $k \in \mathbb{Z}^+$, we obtain a family of delays parameterized by k, that guarantee excitation signals u_j with the same delays in F_i. This particular setting, which we call the *path-independent delay case*, is important for the generalization of results to more general dynamical models where the nonlinear functions can depend on the outputs of nodes with several delays and the mixing of information is possible [15]. Furthermore, the identification of networks with this particular choice of delays is more challenging since the number of excitation variables that can be used is smaller and the next proposition highlights the relationship with the general case corresponding to any configuration of delays.

Proposition 2. *If a DAG G is identifiable in the path-independent delay case, it is also identifiable in the general case.*

Proof. Let us consider an arbitrary DAG G and an arbitrary node i. The output y_i is of the form:

$$y_i^k = u_i^{k-1} + F_i(u_1^{k-2}, \dots, u_1^{k-M_1}, \dots, u_{n_i}^{k-2}, \dots, u_{n_i}^{k-M_{n_i}}).$$

If the node i is measured, we know the function F_i, and we can restrict the analysis to the subset of the space of excitation signals u_j where for each node $j \in V \setminus \{i\}$, the excitation signals satisfy $u_j^{k-m} = u_j^{k-M_j}$ for any $m < M_j$ where M_j is the minimum delay in F_i. In this subset, all the excitation signals associated to a node j have the same delay, so that the function F_i is of the form $F_i(u_1^{k-T_{i,1}}, \dots, u_{n_i}^{k-T_{i,n_i}})$, which corresponds to the path-independent delay case. Clearly, if the set of node functions $\{\Phi\}$ is identifiable in this subset, it is

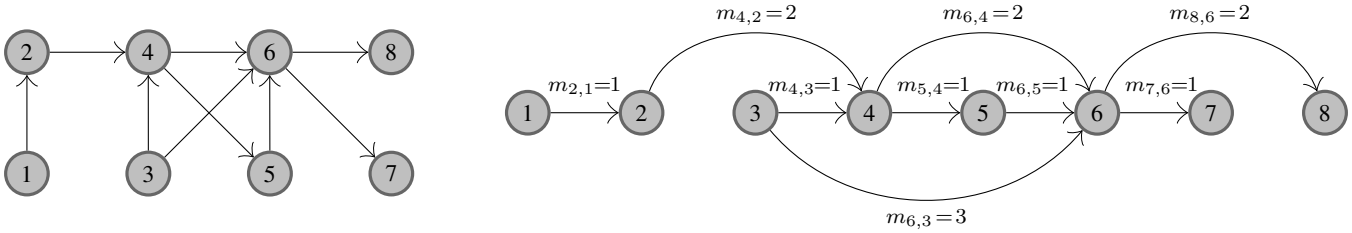

Fig. 4. DAG with a topological ordering where there exists always a particular choice of delays $m_{i,j} = i - j$ such that all the excitation signals u_j in F_i have the same delays.

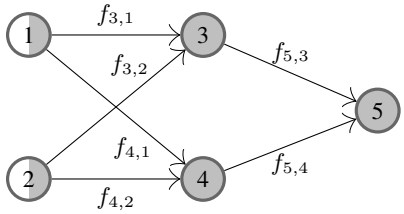

Fig. 5. A DAG where a particular choice of functions makes the network unidentifiable. If $f_{3,1} = f_{4,1}$ and $f_{3,2} = f_{4,2}$, the outputs y_3 and y_4 are the same, which implies that the functions $f_{5,3}$ and $f_{5,4}$ cannot be identified.

also identifiable in all the space, since there is no other set of node functions $\{\tilde{\Phi}\}$ that leads to F_i in the subset. Hence, if the path-independent delay case is identifiable, the general case is also identifiable. □

According to Proposition 2 the identifiability in the case of path-independent delays implies identifiability for any other configuration of delays corresponding to the general case. This implies that any sufficient condition for identifiability in the path-independent delay case is also sufficient for the general case. By contrast, a necessary condition for identifiability in the path-independent delay case could not be necessary for some particular configurations of delays. However, in our work, given a network topology G, we are interested in identifiability conditions for any configuration of delays, which implies the necessity of Proposition 2. Moreover, the path-independent delay case can be considered for any graph topology according to Proposition 1 and for this reason, in this work, we will focus on the path-independent delay case by considering that (2) is given by:

$$y_i^k = u_i^{k-1} + F_i(u_1^{k-T_{i,1}}, \ldots, u_{n_i}^{k-T_{i,n_i}}), \text{ for } 1, \ldots, n_i \in \mathcal{N}^{e \to i}. \tag{7}$$

Regarding the measurement of the node 3 in (6), the function F_3 will be of the form:

$$y_3^k = F_3(u_1^{k-T_{3,1}}),$$

where $T_{3,1} = m_{2,1} + m_{3,2} + 1$.

B. Zero measure set of functions

The notion of generic identifiability introduced in Section II-B, where a network could not be identifiable only for few particular cases of functions, can be motivated by the following example.

Example 1 (Particular choice)**.** *Let us consider the DAG in Fig. 5. If $f_{3,1} = f_{4,1} = h_1$ and $f_{3,2} = f_{4,2} = h_2$, then the measurement of the node 5 provides the output:*

$$\begin{aligned} y_5^k &= f_{5,3}(f_{3,1}(u_1^{k-T_{5,1}}) + f_{3,2}(u_2^{k-T_{5,2}})) \\ &\quad + f_{5,4}(f_{4,1}(u_1^{k-T_{5,1}}) + f_{4,2}(u_2^{k-T_{5,2}})) \\ &= f_{5,3}(h_1(u_1^{k-T_{5,1}}) + h_2(u_2^{k-T_{5,2}})) \\ &\quad + f_{5,4}(h_1(u_1^{k-T_{5,1}}) + h_2(u_2^{k-T_{5,2}})), \end{aligned} \tag{8}$$

and it is not possible to identify the functions $f_{5,3}$ and $f_{5,4}$ even if $f_{3,1}$, $f_{3,2}$, $f_{4,1}$ and $f_{4,2}$ are known, since $\tilde{f}_{5,3} = f_{5,4}$ and $\tilde{f}_{5,4} = f_{5,3}$ also satisfies (8)*. However, Theorem 1 will show that this DAG is identifiable for most functions.*

Notice that a particular choice of functions like in Example 1 can also make the DAG unidentifiable for the non-additive nonlinear model (1) since if the outputs y_3 and y_4 are the same, the functions $\Phi_5(y_3, y_4)$ and $\tilde{\Phi}_5(y_3, y_4) = \Phi_5(y_4, y_3)$ should generate the same measured function F_5 in the case of absence of symmetry with respect to the arguments (i.e., $\Phi_5(y_3, y_4) \neq \Phi_5(y_4, y_3)$).

In this subsection, we will formalize this notion of generic identifiability with respect to the functions $f_{i,j}$. First, we define the *K-analytic parametrization consistent with a given graph G* in the following way. Given a $K \in \mathbb{Z}^+$, for every edge (i, j), set variables $\beta_{i,j}^{(p)} \in \mathbb{R}$, with $p \geq K+1$, and parametrize $f_{i,j}(y_j^{k-m_{i,j}})$ by

$$f_{i,j}(y_j^{k-m_{i,j}}) = \sum_{\ell=1}^{K} \alpha_{i,j}^{(\ell)} (y_j^{k-m_{i,j}})^\ell + \sum_{\ell=K+1}^{\infty} \beta_{i,j}^{(\ell)} (y_j^{k-m_{i,j}})^\ell, \tag{9}$$

for real parameters $\alpha_{i,j}^{(q)}$, with $1 \leq q \leq K$, where we assume that the series $\sum_{\ell=K+1}^{\infty} \beta_{i,j}^{(\ell)} (y_j^{k-m_{i,j}})^\ell$ is convergent. For pairs (i, j) not connected by an edge, let $f_{i,j}(y_j^{k-m_{i,j}}) \equiv 0$. We collect all parameters $\alpha_{i,j}^{(q)}$ in a vector P, and introduce the notion of a K-generic property related to the graph G, which is related to the notion of genericity used in [19].

Definition 4 (K-generic property)**.** *We say that a property K-*generically *holds for a given graph G if, for the K-analytic parametrization consistent with the graph G, the property holds for all parameters $\alpha_{i,j}^{(\ell)}$ except possibly those lying on a zero measure set, and for every $\beta_{i,j}^{(\ell)}$.*

Notice that in Definition 4, the variables $\beta_{i,j}^{(\ell)}$ are not involved in the notion of a zero measure set. Then, based on the notion a K-generic property, we define a generic property.

Definition 5 (Generic property)**.** *We say that a property* generically *holds for a given graph G if there exists a $K \in \mathbb{Z}^+$, such that the property is K-generic.*

Now, we introduce some matrices that will be used in the proof of Theorem 1. Let us consider a static equation associated with the dynamic model (4) where we set all the functions $g_i = 0$ and the delays $m_{i,j} = 0$:

$$y_i^k = \sum_{j \in \mathcal{N}_i} f_{i,j}(y_j^k) + u_i^k, \quad \text{for all } i \in V. \tag{10}$$

We begin by defining a network matrix $J_{\mathfrak{F}}^0(\mathbf{y}^k)$ associated with the graph G composed by the derivatives of the functions $f_{i,j}$:

$$J_{\mathfrak{F}}^0(\mathbf{y}^k) = \begin{pmatrix} 0 & f'_{1,2}(y_2^k) & \cdots & f'_{1,n}(y_n^k) \\ f'_{2,1}(y_1^k) & 0 & \cdots & f'_{2,n}(y_n^k) \\ \vdots & \ddots & \ddots & \vdots \\ f'_{n,1}(y_1^k) & f'_{n,2}(y_2^k) & \cdots & 0 \end{pmatrix}, \tag{11}$$

where $\mathbf{y}^k = (y_1^k, \ldots, y_n^k)$ is a vector with the outputs of the nodes and can be considered as a free variable. By applying (10) in $J_{\mathfrak{F}}^0(\mathbf{y}^k)$, each entry of $J_{\mathfrak{F}}^0(\mathbf{y})$ can be expressed as a function of $\mathbf{u}_G^k = (u_1^k, \ldots, u_{|\mathcal{N}^e|}^k)^T$ since the graph is acyclic. Then, we define the *nonlinear network matrix* $J_G(\mathbf{u}_G^k)$ as the matrix $J_{\mathfrak{F}}^0(\mathbf{y}^k)$ evaluated in $\mathbf{y}^k$ according to (10):

$$J_G(\mathbf{u}_G^k) = \begin{pmatrix} 0 & F_{1,2}(\mathbf{u}_G^k) & \cdots & F_{1,n}(\mathbf{u}_G^k) \\ F_{2,1}(\mathbf{u}_G^k) & 0 & \cdots & F_{2,n}(\mathbf{u}_G^k) \\ \vdots & \ddots & \ddots & \vdots \\ F_{n,1}(\mathbf{u}_G^k) & F_{n,2}(\mathbf{u}_G^k) & \cdots & 0 \end{pmatrix}, \tag{12}$$

where each $F_{i,j}$ is an analytic function since it is obtained by sums and compositions of analytic functions of the form $f_{a,b}$. Notice that the coefficients of the Taylor series of each $F_{i,j}$ is a function of the parameters $\alpha_{i,j}^{(\ell)}$ and $\beta_{i,j}^{(\ell)}$ in (9).

In our work, we are particularly interested in the genericity of the rank of the square submatrices of $J_G(\mathbf{u}_G^k)$, which will be used in the proof of Theorem 1 for the derivation of a sufficient condition for identifiability. The maximal rank of a square submatrix $J_G^{A,B}(\mathbf{u}_G^k)$ is the highest possible rank. Given a square submatrix $J_G^{A,B}(\mathbf{u}_G^k)$, we say that its rank is *generic* if having the maximal rank for almost all $\mathbf{u}_G^k$ holds generically over the K-parametrizations.

Proposition 3. *The rank of any square submatrix $J_G^{A,B}(\mathbf{u}_G^k)$ is generic for almost all $\mathbf{u}_G^k$.*

Proof. The determinant of any square submatrix $J_G^{A,B}(\mathbf{u}_G^k)$, denoted by $\det(J_G^{A,B}(\mathbf{u}_G^k))$ is given by sums and products of analytic functions of the form $f'_{i,j}$, so that $\det(J_G^{A,B}(\mathbf{u}_G^k))$ is also analytic and can be expressed as:

$$\det(J_G^{A,B}(\mathbf{u}_G^k)) = \sum_{p_1,\ldots,p_{|\mathcal{N}^e|}=0} a_{p_1,\ldots,p_{|\mathcal{N}^e|}} (u_1^k)^{p_1} \cdots (u_{|\mathcal{N}^e|}^k)^{p_{|\mathcal{N}^e|}}, \tag{13}$$

If the determinant is not identically zero, there must be at least a coefficient $a_{\hat{p}_1,\ldots,\hat{p}_{|\mathcal{N}^e|}} \neq 0$. Let us denote by $a_{\min}$ one of the nonzero coefficients corresponding to the minimum degree $D_{\min}$ of the terms in (13). This coefficient is given by a polynomial $p(P_{\min})$, where $P_{\min} \in \mathbb{R}^K$ is a vector that encompasses the coefficients $\alpha_{i,j}^{(\ell)}$ and variables $\beta_{i,j}^{(\ell)}$ of the functions $f_{i,j}$ in (9) associated with terms whose degree is at most $D_{\min}$, and K is finite and nonzero. Notice that this polynomial $p(P_{\min})$ can be zero only on a subspace of dimension at most $K-1$. Then, the rank of the submatrix $J_G^{A,B}$ is maximal for all $P_{\min} \in \mathbb{R}^K$, except possibly on a subspace of dimension at most $K-1$, which has measure zero. Since one of the coefficient in (13) is nonzero, the determinant is nonzero for almost all $\mathbf{u}_G^k$. Therefore, the rank of any square submatrix $J_G^{A,B}$ is K-generic and since it holds for this particular value of K, it is also generic. □

In Section IV, we will also analyze the submatrices of $T_G(\mathbf{u}_G^k) = (I - J_G(\mathbf{u}_G^k))^{-1}$, which is well defined since $J_G(\mathbf{u}_G^k)$ is upper-triangular in a DAG [4]. Notice that Proposition 3 is also valid for any square submatrix $T_G^{A,B}(\mathbf{u}_G^k)$.

IV. Directed Acyclic Graphs

In this section, we will derive a sufficient and necessary condition for the generic identifiability of DAGs based on the notion of vertex-disjoint paths.

A. Sufficient condition

First, similarly to the linear case and the additive model, we will analyze the role of the sources and sinks in the identifiability of DAGs for the non-additive model (1) characterized by node functions that are not necessarily separable in edge functions.

Lemma 1. *The outgoing edges of sources are not identifiable if the sources are not excited. The excitation of sinks is never necessary for the identifiability of the network. The measurement of sources is never necessary.*

Proof. The measurement of an out-neighbor j of the source i provides the output:

$$y_j^k = u_j^{k-1} + \Phi_j(y_1^{k-m_{j,1}-1}, \ldots, y_i^{k-m_{j,i}-1}, \ldots, y_{M_j}^{k-m_{j,M_j}-1}). \tag{14}$$

Since the source i is not excited, (14) becomes:

$$y_j^k = u_j^{k-1} + \Phi_j(y_1^{k-m_{j,1}-1}, \ldots, 0, \ldots, y_{M_j}^{k-m_{j,M_j}-1}). \tag{15}$$

Notice that any node function $\tilde{\Phi}_j = \Phi_j + \psi_i(y_i^{k-m_{j,i}-1})$ where ψ_i is analytic and $\psi_i(0) = 0$, also satisfies (15). This implies that it is not possible to identify Φ_j and hence, the excitation of all the sources is necessary for identifiability of the network. Now, the excitation of a sink ℓ can only affect its output:

$$y_\ell^k = u_\ell^{k-1} + \Phi_\ell,$$

so that the identifiability of Φ_ℓ is independent of the value of u_ℓ^{k-1}, which implies that the excitation of the sink is never necessary for identifiability.

Finally, the measurement of a source i provides the output $y_i = u_i^{k-1}$ that does not include any unknown dynamics of the network. □

Next, we establish a link between vertex-disjoint paths and the rank of submatrices $T_G^{A,B}(\mathbf{u}_G^k)$.

Definition 6 (Vertex-Disjoint Paths [4])**.** *A group of paths are mutually vertex disjoint if no two paths of this group contain the same vertex.*

Proposition 4. *The generic rank of any square submatrix $T_G^{A,B}(\mathbf{u}_G^k)$ is at least the number of vertex-disjoint paths from A to B.*

Proof. From Proposition 3, the rank of any square submatrix $T_G^{A,B}(\mathbf{u}_G^k)$ is generic. Let us consider the particular case of linear functions of the form $f_{i,j}(y_j^k) = \varphi_{i,j} y_j^k$, which clearly belong to the class $\mathcal{F}_A$. For this type of functions, the entries of the matrix $J_{\mathfrak{F}}^0(\mathbf{y}^k)$ in (11) are constant values given by $\varphi_{i,j}$ and since they are independent of $\mathbf{y}^k$, the matrix $J_G(\mathbf{u}_G^k)$ in (12) satisfies $J_G(\mathbf{u}_G^k) = J_{\mathfrak{F}}^0(\mathbf{y}^k)$. Notice that in this case, the matrix $J_G(\mathbf{u}_G^k)$ is equivalent to the adjacency matrix of a graph whose weights are $\varphi_{i,j}$ for each edge, and the parametrization is the same in [4] if we fix $p_{i,j} = 0$ and all $\alpha_{i,j}^m = 0$ and $\beta_{i,j}^m = 0$ in [4, Equation 18]. Then, according to [4, Proposition V.1], the generic rank of any submatrix $T_G^{A,B}(\mathbf{u}_G^k)$ for the case of linear functions is the number of vertex-disjoint paths from the subset of nodes $A \subseteq V$ to the subset of nodes $B \subseteq V$, which we denote by p. Finally, since the rank is generic, it should be at least p for more general functions in the class $\mathcal{F}_A$, which completes the proof. □

In the identifiability of linear networks characterized by transfer functions in the edges, the generic rank is equal to the number of vertex-disjoint paths [4]. However, in our work, we only need to know that the generic rank is at least the number of vertex-disjoint paths, while the other direction of the inequality is out of scope. Now, we provide a sufficient condition based on vertex-disjoint paths for the generic identifiability of DAGs in the full measurement case (all nodes are measured).

Theorem 1. *In the full measurement case, a DAG is generically identifiable in the class $\mathcal{F}_A$ if there are vertex-disjoint paths from excited nodes to the in-neighbors of each node.*

Before presenting the proof of Theorem 1, we recall a technical result that will be used in the proof.

Lemma 2 (Theorem 2.35 [24])**.** *Let $H \subset \mathbb{R}^p$ and let $f : H \to \mathbb{R}^q$, where $p \geq q$. If f is continuously differentiable at the point $a \in$ int H and the linear mapping $f'(a) : \mathbb{R}^p \to \mathbb{R}^q$ is surjective, then the range of f contains a neighborhood of $f(a)$.*

Proof of Theorem 1: We proceed by induction. For a DAG and a topological ordering, we take a node i and we denote without loss of generality by $\mathcal{N}_i = \{1, \ldots, p\}$ the set of in-neighbors. We assume by induction that all the dynamics F_j at the left of the node i have been identified. In the basic case of $\mathcal{N}_i$ being empty (i is a source), this assumption is trivial. Let us set to zero all the excitation signals except by p excited nodes with vertex-disjoint paths to the in-neighbors of i. This is equivalent to having only p excited nodes in the network. The measurement of i is given by:

$$\begin{aligned} y_i^k &= u_i^{k-1} + \Phi_i(y_1^{k-m_{i,1}}, \ldots, y_p^{k-m_{i,p}}) \\ &= \sum_{j=1}^{p} f_{i,j}(\xi_{i,j}(u_1^{k-T_{i,1}}, \ldots, u_{n_i}^{k-T_{i,n_i}})) + g_i(\xi_{i,1}, \ldots, \xi_{i,p}) \\ &= F_i(\xi_{i,1}, \ldots, \xi_{i,p}), \end{aligned}$$

where $\xi_{i,j} = u_j^{k-1} + F_j(u_1^{k-T_{j,1}}, \ldots, u_{n_j}^{k-T_{j,n_j}})$ is the output of the in-neighbor j of i as a function of all the excitation signals that arrive to i. Let us assume that there exists a set $\{\Phi\} \neq \{\tilde{\Phi}\}$ such that $F(\mathcal{N}^e) = \tilde{F}(\mathcal{N}^e)$. Since $i \in \mathcal{N}^m$, the measured functions F_i and $\tilde{F}_i$ must satisfy:

$$F_i(\xi_{i,1}, \ldots, \xi_{i,p}) = \tilde{F}_i(\tilde{\xi}_{i,1}, \ldots, \tilde{\xi}_{i,p}), \tag{16}$$

which becomes

$$F_i(\xi_{i,1}, \ldots, \xi_{i,p}) = \tilde{F}_i(\xi_{i,1}, \ldots, \xi_{i,p}), \tag{17}$$

since all the dynamics at the left of i are known (i.e., $\xi_{i,j} = \tilde{\xi}_{i,j}$ for all $j \in \mathcal{N}_i$) according to the induction. Now, let us define the mapping:

$$\Xi_i : \mathbb{R}^{n_i} \to \mathbb{R}^p$$

$$\Xi_i(w_1, \ldots, w_{n_i}) = (\xi_{i,1}(w_1, \ldots, w_{n_i}), \ldots, \xi_{i,p}(w_1, \ldots, w_{n_i})),$$

which sends the n_i excitation signals with a path to the node i to the outputs of the in-neighbors of i. We denote the Jacobian matrix of Ξ_i as $J_{\Xi_i}(\mathbf{w})$ where $\mathbf{w} = (w_1, \ldots, w_{n_i})$. First, let us consider the submatrix $T_G^{\mathcal{N}^e, \mathcal{N}^i}(\mathbf{u}_G^k)$ of the nonlinear network matrix $T_G(\mathbf{u}_G^k)$ defined in Section III-B where the rows are the in-neighbors of i and the columns are the nodes with the excitation signals. According to Proposition 4, since there are vertex-disjoint paths from excited nodes to the in-neighbors of i, the generic rank of $T_G^{\mathcal{N}^e, \mathcal{N}^i}(\mathbf{u}_G^k)$ is p. This implies that the determinant $\det(T_G^{\mathcal{N}^e, \mathcal{N}^i}(\mathbf{u}_G^k))$ is not zero and there exists at least a minimum non-zero coefficient $a_{\min}(P_{\min})$ where the vector of parameters $P_{\min} \in \mathbb{R}^K$ is associated with coefficients of the functions $f'_{i,j}$ according to the analytic parametrization (9).

Now, let us analyze the rank of the Jacobian matrix $J_{\Xi_i}(\mathbf{w})$. The determinant $\det(J_{\Xi_i}(\mathbf{w}))$ is an analytic function where each coefficient is a function of a finite number of coefficients of the functions $f_{i,j}$ and g_i depending on the degree of the coefficient of $\det(J_{\Xi_i}(\mathbf{w}))$. Let us consider an coefficient $b_{\min}(Q_{\min})$ where $Q_{\min} = (P_{\min} \quad G)^T \in \mathbb{R}^L$ and G is a vector with a finite number of coefficients of the functions g_i. Then, let us consider the particular case when $G = 0$. In this case, we have $b_{\min} = a_{\min}$ which is non zero, and it could only be zero in a subspace of dimension at most $L-1$, which has zero measure.

Since the determinant $\det(J_{\Xi_i}(\mathbf{w}))$ is an analytic function of $\mathbf{w}$, it can be zero everywhere or only on a set of measure zero [25]. This implies that there exists at least a point $\mathbf{w}^* \neq 0$ such that $\det(J_{\Xi_i}(\mathbf{w}^*)) \neq 0$ and hence the generic rank of $J_{\Xi_i}(\mathbf{w}^*)$ is p. Therefore, $J_{\Xi_i}(\mathbf{w}^*)$ is surjective and by virtue

of Lemma 2, the range of Ξ_i must contain a neighborhood of $\Xi_i(\mathbf{w}^*)$, which implies that (17) holds on a set of positive measure. Then, by the Identifiy Theorem of analytic functions [25], we guarantee that

$$F_i(z_1,\dots,z_m)=\tilde{F}_i(z_1,\dots,z_m) \text{ for all } z_1,\dots,z_m\in D\subseteq\mathbb{R},$$

which implies:

$$\Phi_i(z_1,\dots,z_m)=\tilde{\Phi}_i(z_1,\dots,z_m) \text{ for all } z_1,\dots,z_m\in D\subseteq\mathbb{R},$$

and hence the node function Φ_i is generically identifiable with respect to the parameters encompassed in the vector $Q_{\min}$. Now, notice that for $i=2$ in the topological ordering, if there is an edge $f_{2,1}$, the function $\xi_{2,1}$ is the identity function and the node function Φ_2 can be clearly identified. Then, by induction, the identifiablity analysis is valid for any node in the DAG. Thus, all the DAG is generically identifiable. □

A direct consequence of Theorem 1 and Lemma 1 is a sufficient and necessary condition for the identifiability of trees.

Proposition 5. *In the full measurement case, a tree is generically identifiable in the class $\mathcal{F}_A$ if and only if all the sources are excited.*

B. Necessary condition for polynomial functions

To analyze the necessity of the identifiability condition derived in Theorem 1 we will consider the class of polynomial functions $\mathcal{F}_{Pol}$, which are essential for the approximation of continuous functions through the Weierstrass Approximation theorem [26]. Furthermore, many nonlinear systems can be accurately identified using polynomial models [27], [28].

Definition 7 (Class of functions $\mathcal{F}_{Pol}$)**.** *Let $\mathcal{F}_{Pol}$ be the class of polynomial functions $\Phi:\mathbb{R}^m\to\mathbb{R}$.*

The restriction of the identifiability conditions to polynomial functions is due to the fact that the necessity will be proved by using the theory of *Algebraic Varieties*, which is exclusively formulated for polynomials. To prove the main result of this subsection, we will use the notion of an affine algebraic variety in the field $\mathbb{R}$.

Definition 8 (Affine algebraic variety)**.** *Let $f_1,\dots,f_s$ be polynomials in $\mathbb{R}[x_1,\dots,x_n]$. Then we call $V(f_1,\dots,f_s)$ the affine algebraic variety defined as:*

$$V(f_1,\dots,f_s)=$$
$$\{(a_1,\dots,a_n)\in\mathbb{R}^n \,|f_i(a_1,\dots,a_n)=0 \textit{ for all } 1\le i\le s\}.$$

The following lemma is well-known in algebraic varieties [29] and we provide a proof for the sake of completeness.

Lemma 3. *Let $V\subseteq\mathbb{R}^n$ be given parametrically as:*

$$\begin{aligned} x_1&=f_1(t_1,\dots,t_m)\\ &\vdots\\ x_n&=f_n(t_1,\dots,t_m), \end{aligned}$$

where $f_1,\dots,f_n$ are polynomials in $\mathbb{R}[t_1,\dots,t_m]$ and $n>m$. Then V is included in an affine algebraic variety $W(F)$, for a polynomial $F\in\mathbb{R}[x_1,\dots,x_n]$ that is not identically zero.

Proof. The proof is left to Appendix A. □

To explicitly compute the polynomial F in Lemma 3, we can apply Theorem 1 (Polynomial Implicitization) in [29].

Theorem 2. *In the full measurement case, a DAG is generically identifiable in the class $\mathcal{F}_{Pol}$ if and only if there are vertex-disjoint paths from excited nodes to the in-neighbors of each node.*

Proof. a) Sufficiency: The class of polynomial functions $\mathcal{F}_{Pol}$ is a subclass of the class of analytic functions $\mathcal{F}_A$. For the class $\mathcal{F}_{Pol}$, the K-analytic parametrization (9) is formulated with all coefficients $\beta_{i,j}^{(p)}=0$ for $p\ge K+1$. Hence, the sufficient condition of Theorem 1 also encompasses the polynomial functions.

b) Necessity: Let us consider an arbitrary node i with M_i in-neighbors, such that we have the following identifiability problem:

$$\Phi_i(y_1,\dots,y_{M_i})=\tilde{\Phi}_i(y_1,\dots,y_{M_i}). \tag{18}$$

According to (7), the outputs of the nodes can be expressed as a function of the excitation signals corresponding to the nodes in $\mathcal{N}^{e\to i}$, such that (18) becomes:

$$\Phi_i(F_1(\mathbf{u}_i^k),\dots,F_{M_i}(\mathbf{u}_i^k))=\tilde{\Phi}_i(F_1(\mathbf{u}_i^k),\dots,F_{M_i}(\mathbf{u}_i^k)), \tag{19}$$

where $\mathbf{u}_i^k=(u_1^{k-T_{i,1}},\dots,u_{n_i}^{k-T_{i,n_i}})$, and we assume that all the dynamics F_j of the in-neighbors of the node i have been identified. Notice that (19) is equivalent to:

$$\Delta\Phi_i(F_1(\mathbf{u}_i^k),\dots,F_{M_i}(\mathbf{u}_i^k))=0, \tag{20}$$

where $\Delta\Phi_i=\Phi_i-\tilde{\Phi}_i$. Let us assume that the maximum number of vertex-disjoint paths from the set of excited nodes $\mathcal{N}^{e\to i}$ to the set of in-neighbors $\mathcal{N}_i$ is $D_i<M_i$. From [4, Lemma V.3], the size of the smallest $\mathcal{N}^{e\to i}-\mathcal{N}_i$ disconnecting set[2], denoted by $\mathcal{D}_{e\to i}$, is the maximum number of vertex disjoint paths given by D_i. Clearly, the output of each in-neighbor of i can be expressed as a function of $z_1,\dots,z_{D_i}\in\mathcal{D}_{e\to i}$ since each path from the set of excited nodes $\mathcal{N}^{e\to i}$ must cross $\mathcal{D}_{e\to i}$. Since the output of each node $z_j\in\mathcal{D}_{e\to i}$ can be expressed as a function of the excitation signals $\mathbf{u}_i^k$, (20) is given by:

$$\Delta\Phi_i(G_1(z_1,\dots,z_{D_i}),\dots,G_{M_i}(z_1,\dots,z_{D_i}))=0. \tag{21}$$

By applying Lemma 3, we guarantee that there exists a function $\Delta\Phi_i\not\equiv 0$ such that $\Delta\Phi_i(G_1(z_1,\dots,z_{D_i}),\dots,G_{M_i}(z_1,\dots,z_{D_i}))=0$ for all $z_1,\dots,z_{D_i}\in\mathbb{R}$. Therefore, the function $\tilde{\Phi}_i=\Phi_i+\Delta\Phi_i$ also satisfies (19), which implies that the node function Φ_i is not unique and therefore not identifiable with the information obtained through the measurement of node i. □

Example 2. *Let us consider the identifiability of the node function Φ_{10} in the DAG in Fig. 6. The node 10 has three in-neighbors 7, 8 and 9, which are reached by the excited nodes*

[2]A set of nodes $\mathcal{B}$ is an $\mathcal{A}-\mathcal{C}$ disconnecting set if every path starting in $\mathcal{A}$ and ending in $\mathcal{C}$ contains at least one node in $\mathcal{B}$, which implies that if $\mathcal{B}$ is removed, there will be no path from $\mathcal{A}$ to $\mathcal{C}$ [4].

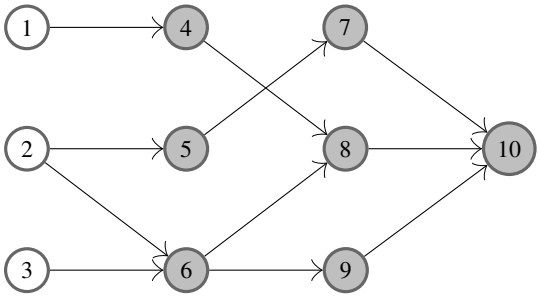

Fig. 6. The vertex-disjoint paths $\{1,4,8\}$, $\{2,5,7\}$, $\{3,6,9\}$ guarantee the identifiability of the DAG in the class $\mathcal{F}_{Pol}$. The set of nodes $\{4,5,6\}$ is a disconnecting set.

1,2 and 3 through the vertex-disjoint paths $\{1,4,8\}$, $\{2,5,7\}$, $\{3,6,9\}$. *Hence, the node function* Φ_{10} *is identifiable in the class* $\mathcal{F}_{Pol}$. *Furthermore, the set of nodes* $\mathcal{B}=\{4,5,6\}$ *is a disconnecting set of* $\mathcal{A}=\{1,2,3\}$ *and* $\mathcal{C}=\{7,8,9\}$ *since all the paths from* $\mathcal{A}$ *to* $\mathcal{C}$ *must cross* $\mathcal{B}$.

We believe that the vertex-disjoint path condition is also necessary for the more general class of analytic functions $\mathcal{F}_A$. Unfortunately, the theory of analytic varieties is not so rich as the theory of algebraic varieties and to the best of our knowledge, a result similar to Lemma 3 has not been derived and it is out of scope of this paper. For this reason, we state this potential sufficient and necessary condition as a conjecture.

Conjecture 1. *In the full measurement case, a DAG is generically identifiable in the class* $\mathcal{F}_A$ *if and only if there are vertex-disjoint paths from excited nodes to the in-neighbors of each node.*

V. Non-additive and additive model

In [20], a similar identifiability condition based on vertex-disjoint paths was obtained for the identifiability of DAGs in the additive model. However, the necessity of this condition remained as an open question. Unfortunately, the following counterexample shows that this sufficient condition is not necessary in the additive model [20].

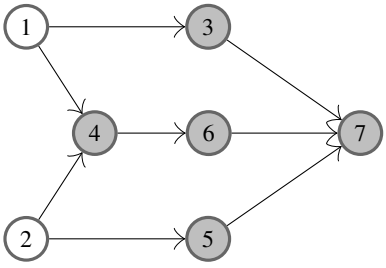

Fig. 7. A DAG where the excitation of the sources 1 and 2 is sufficient for identifiability in the additive nonlinear model but is not sufficient for identifiability in the non-additive nonlinear model since it does not satisfy the vertex-disjoint path conditions of Theorem 2.

Example 3. *Let us consider the DAG in Fig. 7. Clearly, since the node 7 has 3 in-neighbors (3,5,6) and we only have two excited nodes (1,2), there are no vertex-disjoint paths from excitations to the in-neighbors of 7 and the DAG is unidentifiable in the non-additive model. For the additive model* (5) *and pure nonlinear functions, the measurement of the node 7 is of the form:*

$$f_{7,3}(g_{7,3}(u_1))+f_{7,6}(g_{7,6}(u_1,u_2))+f_{7,5}(g_{7,5}(u_2))= \tilde{f}_{7,3}(g_{7,3}(u_1))+\tilde{f}_{7,6}(g_{7,6}(u_1,u_2))+\tilde{f}_{7,5}(g_{7,5}(u_2)), \quad (22)$$

which is equivalent to

$$\Delta f_{7,3}(g_{7,3}(u_1))+\Delta f_{7,6}(g_{7,6}(u_1,u_2))+\Delta f_{7,5}(g_{7,5}(u_2))=0, \quad (23)$$

where we omit the time dependence of the excitation signals u_1 *and* u_2. *Notice that the function* $g_{7,6}$ *must contain terms of the form* $u_1^p u_2^q$, *and if* $\Delta f_{7,6}\not\equiv 0$, *these terms cannot be canceled by* $\Delta f_{7,3}$ *and* $\Delta f_{7,5}$ *that only depend on* u_1 *and* u_2 *respectively. Similarly, the function* $\Delta f_{7,3}$ *or* $\Delta f_{7,5}$ *cannot be different from zero, because their terms cannot be canceled by the other functions. Therefore, all the functions* $\Delta f_{7,3}$, $\Delta f_{7,6}$ *and* $\Delta f_{7,5}$ *must be necessarily 0 to satisfy* (23) *and hence, the DAG is identifiable in the additive model even if there are no vertex-disjoint paths from the excitations to the in-neighbors of 7. Notice that the incoming edges of 7 are identifiable independently of the functions associated with the other edges of the network.*

This difference between the additive and the non-additive nonlinear models can be explained by the separability of the node function associated to the additive model, which limits considerably the potential class of functions that can satisfy the information obtained with the measurement of a node. However, notice that for the linear case, where the functions are necessarily separable, the vertex-disjoint path condition is also necessary. This can be justified by the fact that the nonlinearity of the functions might generate terms involving the product of two or more excitation signals as in Example 3, which will allow us to identify a node function with a reduced number of excitation signals.

VI. Conclusions and future work

In this paper, we analyzed the identifiability of a network with a non-additive nonlinear model that can be expressed as a node function in the case of full measurements. Unlike the additive model, we showed that the presence of a static component in the node functions does not affect the identifiability, allowing us to work with a more general class of functions. Then, we defined the path-independent delay case and we showed that identifiability conditions for this particular case are sufficient and necessary for identifiability in any configuration of delays of the functions. We introduced the notion of generic identifiability for nonlinear functions and characterized the set of measure zero associated with the generic notion as a subspace of finite dimension. For analytic functions in DAGs, we provided a sufficient condition in terms of vertex-disjoint paths that coincide with the identifiability conditions in the linear case. For the class of polynomial functions, we showed that the sufficient condition is also necessary, which is different from the additive nonlinear model. Finally, we formulated a conjecture about the necessity of the vertex-disjoint path

condition for the identifiability of DAGs in the class of analytic functions.

A natural continuation of this work is to derive a result similar to Lemma 3 for the case of analytic varieties that would allow us to prove the final conjecture of this paper. Also, it would be important to analyze identifiability conditions for more general DAGs where loops are present. In this case, the notion of an unfolded digraph associated with any network introduced in [15] could be an important tool.

REFERENCES

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Physics reports*, vol. 424, no. 4-5, pp. 175–308, 2006.

[2] F. Bullo, *Lectures on Network Systems*, 1.7 ed. Kindle Direct Publishing, 2022.

[3] H. H. Weerts, P. M. J. Van den Hof, and A. G. Dankers, "Identifiability of linear dynamic networks," *Automatica*, vol. 89, pp. 247–258, 2018.

[4] J. M. Hendrickx, M. Gevers, and A. S. Bazanella, "Identifiability of dynamical networks with partial node measurements," *IEEE Transactions on Automatic Control*, vol. 64, no. 6, pp. 2240–2253, 2019.

[5] H. J. van Waarde, P. Tesi, and M. K. Camlibel, "Necessary and sufficient topological conditions for identifiability of dynamical networks," *IEEE Transactions on Automatic Control*, vol. 65, no. 11, pp. 4525–4537, 2020.

[6] A. S. Bazanella, M. Gevers, and J. M. Hendrickx, "Network identification with partial excitation and measurement," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 5500–5506.

[7] X. Cheng, S. Shi, I. Lestas, and P. M. J. Van den Hof, "A necessary condition for network identifiability with partial excitation and measurement," *IEEE Transactions on Automatic Control*, vol. 68, no. 11, pp. 6820–6827, 2023.

[8] E. Mapurunga, M. Gevers, and A. S. Bazanella, "Identifiability of dynamic networks: The essential rôle of dources and dinks," *IEEE Transactions on Automatic Control*, vol. 69, no. 2, pp. 1303–1308, 2024.

[9] A. Legat and J. M. Hendrickx, "Identifiability in networked systems: Algebraic and combinatorial conditions," 2024.

[10] W. Pan, Y. Yuan, J. Gonçalves, and G.-B. Stan, "Reconstruction of arbitrary biochemical reaction networks: A compressive sensing approach," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, 2012, pp. 2334–2339.

[11] F. Dörfler and F. Bullo, "Synchronization in complex networks of phase oscillators: A survey," *Automatica*, vol. 50, no. 6, pp. 1539–1564, 2014.

[12] A. Aalto, L. Viitasaari, P. Ilmonen, L. Mombaerts, and J. Gonçalves, "Gene regulatory network inference from sparsely sampled noisy data," *Nature communications*, vol. 11, no. 1, p. 3493, 2020.

[13] A. Bizyaeva, A. Franci, and N. E. Leonard, "Nonlinear opinion dynamics with tunable sensitivity," *IEEE Transactions on Automatic Control*, vol. 68, no. 3, pp. 1415–1430, 2023.

[14] R. Vizuete and J. M. Hendrickx, "Nonlinear network identifiability: The static case," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, 2023, pp. 443–448.

[15] ——, "Nonlinear network identifiability with full excitations," *arXiv preprint arXiv:2405.07636*, 2024.

[16] C. C. Aggarwal, "Neural Networks and Deep Learning," *Springer*, 2018.

[17] F. Bonassi, C. Andersson, P. Mattsson, and T. B. Schön, "Structured state-space models are deep Wiener models," *IFAC-PapersOnLine*, vol. 58, no. 15, pp. 247–252, 2024.

[18] J. G. T. Zañudo, G. Yang, and R. Albert, "Structure-based control of complex networks with nonlinear dynamics," *Proceedings of the National Academy of Sciences*, vol. 114, no. 28, pp. 7234–7239, 2017.

[19] M. Vanelli and J. M. Hendrickx, "Local identifiability of fully-connected feed-forward networks with nonlinear node dynamics," *arXiv preprint arXiv:2412.08472*, 2025.

[20] R. Vizuete and J. M. Hendrickx, "Nonlinear identifiability of directed acyclic graphs with partial excitation and measurement," in *2024 63rd IEEE Conference on Decision and Control (CDC)*, 2024, pp. 3153–3159.

[21] J. Van der Woude, "A graph-theoretic characterization for the rank of the transfer matrix of a structured system," *Mathematics of Control, Signals and Systems*, vol. 4, pp. 33–40, 1991.

[22] L. Ljung, *System Identification: Theory for the User*. Prentice Hall PTR, 1999.

[23] J. Bang-Jensen and G. Z. Gutin, *Digraphs: Theory, Algorithms and Applications*. Springer Science & Business Media, 2008.

[24] M. Laczkovich and V. T. Sós, *Real Analysis: Series, Functions of Several Variables, and Applications*. Springer, 2017, vol. 3.

[25] S. G. Krantz and H. R. Parks, *A Primer of Real Analytic Functions*. Springer Science & Business Media, 2002.

[26] J. G. Llavona, *Approximation of Continuously Differentiable Functions*. North-Holland Mathematics Studies, 1986.

[27] A. Janczak, *Identification of Nonlinear Systems Using Neural Networks and Polynomial Models*. Springer, 2004.

[28] J. Paduart, L. Lauwers, J. Swevers, K. Smolders, J. Schoukens, and R. Pintelon, "Identification of nonlinear systems using polynomial nonlinear state space models," *Automatica*, vol. 46, no. 4, pp. 647–656, 2010.

[29] D. A. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*. Springer, 2015.

APPENDIX

A. Proof of Lemma 3

Let us denote by N the maximal degree of the polynomials $f_1, \dots, f_n$. Then, we have that each polynomial satisfies:

$$\deg([f_1]^{a_1} \cdots [f_n]^{a_n}) \leq Na_1 + Na_n \leq NM,$$

for $a_1 + \cdots + a_n \leq M$. If we have more than $\binom{NM+m}{m}$ polynomials of the form $[f_1]^{a_1} \cdots [f_n]^{a_n}$, we have a dependent set since $\binom{NM+m}{m}$ is the dimension of the vector space of NM-variate polynomials of degree m. Now, the number of polynomials of this form is $\binom{M+n}{n} = \binom{M+m+1}{m+1}$. Since $\binom{M+m+1}{m+1} = O(M^{m+1})$, for a M large enough, the polynomials $[f_1]^{a_1} \cdots [f_n]^{a_n}$ with $a_1 + \cdots + a_n \leq M$ form a linearly dependent set in $\mathbb{R}[t_1, \dots, t_m]$. This implies that there are $\alpha_{a_1,\dots,a_n} \in \mathbb{R}$ not all zero, such that:

$$\sum_{a_1+\cdots+a_n \leq M} \alpha_{a_1,\dots,a_n} [f_1]^{a_1} \cdots [f_n]^{a_n} = 0.$$

Thus V belongs to the affine variety $W(F)$ where

$$F(x_1, \dots, x_n) = \sum_{a_1+\cdots+a_n \leq M} \alpha_{a_1,\dots,a_n} x_1^{a_1} \cdots x_n^{a_n}.$$

Renato Vizuete received the M.S. degree (très bien) in Systems, Control, and Information Technologies from Université Grenoble Alpes, France (2019), and the PhD degree in Automatic Control from Université Paris-Saclay, France (2022). He is currently a FNRS postdoctoral researcher - CR at UCLouvain, Belgium, in the ICTEAM Institute. His research interests include optimization, control theory, machine learning, network theory and system identification. He was the recipient of the Networks and Communication Systems TC Outstanding Student Paper Prize of the IEEE Control Systems Society in 2022, and the Second Thesis Prize in the category Impact Science of the Fondation CentraleSupélec in 2023.

Julien M. Hendrickx is professor of mathematical engineering at UCLouvain, in the Ecole Polytechnique de Louvain since 2010, and is currently the head of the ICTEAM institute.

He obtained an engineering degree in applied mathematics (2004) and a PhD in mathematical engineering (2008) from the same university. He has been a visiting researcher at the University of Illinois at Urbana Champaign in 2003-2004, at the National ICT Australia in 2005 and 2006, and at the Massachusetts Institute of Technology in 2006 and 2008. He was a postdoctoral fellow at the Laboratory for Information and Decision Systems of the Massachusetts Institute of Technology 2009 and 2010, holding postdoctoral fellowships of the F.R.S.-FNRS (Fund for Scientific Research) and of Belgian American Education Foundation. He was also resident scholar at the Center for Information and Systems Engineering (Boston University) in 2018-2019, holding a WBI.World excellence fellowship.

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