



HAL
open science

A graph theoretical approach to the parameters identifiability characterization.

Taha Boukhobza, Frédéric Hamelin, Christophe Simon

► **To cite this version:**

Taha Boukhobza, Frédéric Hamelin, Christophe Simon. A graph theoretical approach to the parameters identifiability characterization.. *International Journal of Control*, 2014, 87 (4), pp.751-763. 10.1080/00207179.2013.856519 . hal-00907164

HAL Id: hal-00907164

<https://hal.science/hal-00907164>

Submitted on 22 May 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

International Journal of Control
Vol. 00, No. 00, Month 2013, 1–20

RESEARCH ARTICLE

A graph theoretical approach to the parameters identifiability characterization.

T. Boukhobza^a*, F. Hamelin^a and C. Simon^a

^a*Centre de Recherche en Automatique de Nancy (CRAN), University of Lorraine, CNRS UMR 7039, BP 70239, 54506 Vandœuvre Cedex, Nancy, France*

(April 2013)

This paper is devoted to the generic parameters identifiability analysis for structured linear systems with unknown inputs. The proposed method is based on a graph-theoretic approach and assumes only the knowledge of the system's structure. Necessary and sufficient conditions are expressed in graphical terms which ensure that a given set of unknown parameters characterizing the model of the system are structurally identifiable using the available measurements, assuming that the inputs and initial conditions are quite informative. Even if we are not interested in proposing identification schemes, these given conditions are necessary and sufficient to the parameter identification whatever the used algorithms to determine numerically the parameters values. Finally, note that the given conditions are quite easy to check because they are based on searching and comparing some specific paths and cycles in a digraph. This makes our approach well suited to study large scale systems.

Keywords: Parameters identifiability; structured linear systems; graph theoretical approach

1 Introduction

This paper deals with the analysis of a priori global identifiability of general structured linear dynamical systems with unknown inputs. In such systems, the matrices have a number of fixed zero entries while the rest of the entries are not numerically known. Following the definition of local identifiability by Bellman and Åström (Bellman and Åström 1970) and later by Chapman and Godfrey (Chapman and Godfrey 1985), a model is identifiable if error-free data lead to a finite number of solutions for all the parameters in the model equations. A model is unidentifiable if at least one parameter is unidentifiable, *i.e.* has an infinite number of possible values. Global identifiability concerns the possibility of uniquely determining the model parameters from input-output data. There are at least two reasons for assessing this property. First, the model parameters can have a physical meaning, and it can be interesting to know whether it is possible to determine their values from experimental data. Second, numerical optimization approaches will have difficulties when trying to estimate the parameters of a non-identifiable model.

A summary of existing research on structural identifiability analysis shows that in 1970s some techniques were developed from control theories for linear models (Bellman and Åström 1970), especially compartmental models. A detailed review and discussion of the different ideas and terminology in the field of linear identifiability can be found in the PhD thesis of Delforge (Delforge 1984). More recently, some methods for testing the structural identifiability of non-linear models have been developed. In (Miao et al. 2011), identifiability analysis methodologies are reviewed for nonlinear ODE models developed in the past couple of decades, including structural identifiability analysis, practical identifiability analysis, and sensitivity-based identifiability analysis. The main drawbacks of these methods are related to the difficulty of assessing the observability condition and the complexity to solve linear inhomogeneous partial differential equations for general non-linear dynamic systems. The more recently developed DAISY

*Corresponding author. Email: taha.boukhobza@univ-lorraine.fr

software (Bellu et al. 2007) enables its application to non-expert users. In line with this approach, the implicit function theorem based method was proposed by Xia and Moog (Xia and Moog 2003). It is based on computing the time derivatives of the outputs to eliminate unobserved states and determining a finite number of equations consisting of known inputs/outputs and unknown parameters. These two kinds of methods can be applied to small-scale systems effectively but, when dealing with medium or large-scale systems, the calculations involved become intractable.

In this context, the aim of this paper is to propose a new graph-theoretic tool to provide necessary and sufficient conditions that ensure that some desired parameters are identifiable according to the structure of the system and the location of the measurements. Our approach is well-suited to treat large scale systems because it uses classical programming techniques and is free from numerical difficulties. Moreover, the key feature of all of the results presented in this paper is that the global identifiability of structured linear dynamical systems is scrutinizing with unknown inputs. To our knowledge, it is a challenging problem that has received very little attention, the majority of recent studies mainly focus on general non-linear dynamic models without unknown inputs.

To study the identifiability of structured systems, we consider models where the fixed zeros are conserved while the non zero entries are replaced by free parameters to be determined. There is a huge amount of interesting studies in the literature using this kind of models. The latter capture most of the available structural information from physical laws. Their study requires a low computational burden, which allows us to deal with large scale systems. Because of these characteristics, we think that structured systems are particularly well suited to study the identifiability.

The pioneering work of Reinschke (Reinschke 1985) showed that graph theoretical tools are useful techniques to analyse basic properties of structured systems. In the context of parameter identifiability, the work of Audoly and D'angiò (Audoly and D'angiò 1980) considers a directed graph-based approach to deal with the problem of the a priori identifiability of compartmental systems from input-output experiments. The problem of determining whether or not a particular parameter, or the entire compartmental system, is structural identifiable can become extremely difficult if the number of unknown parameters and the number of compartments is large. With the objective to consider large linear compartmental systems, Eisenfeld *et al.* in (Eisenfeld 1982) and (Eisenfeld and Grundy 1983) introduce new methods based on some digraph properties and a subsystem decomposition in such a way that a parameter is structural identifiable with respect to the large system if and only if it is structural identifiable with respect to the subsystem in which it is contained. In (Eisenfeld and Grundy 1983), sufficient conditions are proposed on the digraph of the system, which guarantee that the desired decomposition can be achieved. However, none of these studies can be used to analyse generic parameter identifiability of structured linear systems with unknown inputs. It is the aim of this paper.

Compared to existing works, the contribution of the presented approach lies mainly in three points : Firstly, few works deal with the identifiability of systems with unknown inputs, representing disturbances, unmodeled dynamics or faults for example. Secondly, most part of identifiability studies use algebraic or geometric approaches and so require more information than the proposed approach which needs only the knowledge of the structure of the system. Moreover, the use of graph-theoretic approach provides quite efficient and systematic analysis tool.

Finally, our approach allows us to address quite easily and immediately the problem of sensors location to ensure the identifiability of a chosen set of parameters. This problem is very interesting from a conception point.

The paper is organized with a Section 2 presenting the problem of the generic parameter identifiability for structured linear systems with unknown inputs. In Section 3, some classical definitions and notations in the graph-theory are given and a digraph for any structured linear system is defined. Section 4 presents some general results on the existence of redundancy equations allowing the identification of unknown parameters for systems with unknown inputs. These results allow to reduce the digraph of the system to its useful part. In Section 5, some lemmas are provided which allow us to establish the necessary and sufficient identifiability conditions in Section 6, which represent the main result of the paper. A brief conclusions is given in Section 7.

2 Problem statement

In this paper, we address numerically non-specified systems on the form:

$$\Sigma_{\Lambda} : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \quad (1)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^q$ and $y \in \mathbb{R}^p$ are respectively the state vector, the input vector and the output measurement vector. Without loss of generality and for the sake of homogeneity, all the inputs are assumed to be unknown. The control input signals, whose values are known, are considered to be measured *i.e.* a virtual output equation of the form $y_k = u_k$ is associated to these inputs.

A , B , C and D represent matrices which elements are either fixed to zero or assumed to be nonzero parameters noted λ_i . It is assumed that the entries of all these matrices, constituting vector $(\lambda_1, \lambda_2, \dots, \lambda_h)^T \in \mathbb{R}^h$, can take any value in \mathbb{R}^h or equivalently that parameters λ_i are free. If all parameters λ_i are numerically fixed, a so-called admissible realization of structured system (Σ_{Λ}) is obtained. A property is said to be true generically (van der Woude 1999) if it is true for almost all the realizations of structured system (Σ_{Λ}) . Here, “for almost all the realizations” is to be understood (Dion et al. 2003, van der Woude 1999) as “for all parameter values $((\lambda_1, \lambda_2, \dots, \lambda_h)^T \in \mathbb{R}^h)$ except for those in some proper algebraic variety in the parameter space”. The proper algebraic variety for which the property is not true is the zero set of some non-trivial polynomial with real coefficients in the h system parameters $\lambda_1, \lambda_2, \dots, \lambda_h$ or equivalently it is an algebraic variety which has Lebesgue measure zero.

In this paper, we are interested in the generic identifiability of a subset of parameters denoted $\Theta = \{\lambda_{i_1}, \lambda_{i_2}, \dots, \lambda_{i_{k_0}}\} \subseteq \{\lambda_1, \lambda_2, \dots, \lambda_h\} = \Lambda$, which may represent directly or not some physical parameters of the system. The parameters to be identified are some of the entries of matrices A , B , C and D . It is assumed also that all the other parameters $(\Lambda \setminus \Theta)$ are known even if the identifiability analysis does not necessitates their numerical values.

Let’s recall briefly the definition of the generic parameter identifiability:

Definition 2.1: The unknown parameters of subset Θ of structured system (Σ_{Λ}) are generically identifiable iff, for almost all realizations of (Σ_{Λ}) , for some initial state x_0 and for input function $u(t)$, $y(t, x_0, u(t), \Theta_0) = y(t, x_0, u(t), \Theta_1)$ for $t \geq 0$ implies $\Theta_0 = \Theta_1$, where $y(t, x_0, u(t), \Theta_0)$ (resp. $y(t, x_0, u(t), \Theta_1)$) represents the output related to system Σ_{Λ} for the initial conditions x_0 , the input $u(t)$ and the unknown parameter values $\Theta = \Theta_0$ (resp. $\Theta = \Theta_1$). \square

In other words, the unknown parameters of subset Θ are generically identifiable, if, when the initial state and the input are quite informative, the measured trajectory allows to uniquely determine the values of all the elements of Θ .

Therefore, to achieve the parameters identifiability, it is necessary and sufficient to have some algebraic and differential equations linking some outputs, which allow to uniquely determine the exact values of the parameters. Since we deal with linear systems, these equations, called **redundancy equations**, have obviously the form:

$$y_i^{(\ell)} = \sum_{s < \ell} \alpha_{i,s} y_i^{(s)} + \sum_{y_j \in \mathbf{Y} \setminus \{y_i\}} \sum_{s=0}^n \alpha_{i,s} y_j^{(s)} \quad (2)$$

where each $\alpha_{i,s}$ can lead to an “**identification**” equation allowing the parameters estimation. The parameters identifiability is achieved if and only if there are sufficiently independent identification equations issued from the redundancy algebraic relations, to express all the unknown parameters.

Note that the aim of the paper is not to find these identification equations neither to propose an identification algorithm but to analyse if yes or no, according to the structure of the system, it can be written sufficiently independent and informative to ensure the unknown parameters identifiability.

To establish our result, which consists on graphical necessary and sufficient conditions for the generic identifiability of the parameters belonging to Θ , we proceed with the following steps:

Step 1: a digraph is associated to the considered system. This digraph and different notations are presented in Section 3.

Step 2: the graph is reduced to its useful part for the identifiability. This is done by establishing firstly that in a certain easily computed part of the graph, it is not possible to exhibit algebraic redundancy equations allowing the identification of the parameters (Lemma 4.2), and secondly that in the remaining part there always exists an algebraic relation linking any output to the others (Lemmas 4.3 and 4.4).

Step 3: in the useful part of the graph, using the graphic characterization of the observability subspace dimension (Lemma 4.6), the notion of eligible edge is defined. Then, it is shown that a necessary condition to the parameter identifiability is that the associated edge of the considered parameter is eligible (Lemma 5.1).

Step 4: the redundancy algebraic equations are characterized by defining the pairs of paths \mathbb{Q}_A and of path-cycle \mathbb{Q}_C . Next, it is proven that all the algebraic equations use only the parameters associated to the edges covered by these pairs. According to this fact, Lemma 5.2 provides necessary and sufficient conditions to check if a parameter appears or not in a redundancy equation. Then, this result is extended to a set of k unknown parameters in Lemma 5.3.

Step 5: a new bipartite graph is defined linking the unknown parameters to the pairs of paths and cycles representing the identification equations. Using the DM-decomposition of this bipartite graph, the sufficient (Lemma 5.5) and necessary (Lemma 5.6) condition are established to the generic parameter identifiability.

Step 6: the final and main result is summarized in Proposition 6.1.

3 Graphical representation and some classical definitions

The digraph associated to (Σ_Λ) is noted $\mathcal{G}(\Sigma_\Lambda)$. It is constituted by a vertex set \mathcal{V} and an edge set \mathcal{E} i.e. $\mathcal{G}(\Sigma_\Lambda) = (\mathcal{V}, \mathcal{E})$. The vertices are associated to the continuous state, the input and the output components of (Σ_Λ) and the directed edges represent links between these variables. More precisely, $\mathcal{V} = \mathbf{X} \cup \mathbf{Y} \cup \mathbf{U}$, where $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ is the set of state vertices, $\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_p\}$ is the set of output vertices and $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_q\}$ is the set of unknown input vertices. The edge set is $\mathcal{E} = A\text{-edges} \cup B\text{-edges} \cup C\text{-edges} \cup D\text{-edges}$, with $A\text{-edges} = \{(\mathbf{x}_j, \mathbf{x}_i) \mid A(i, j) \neq 0\}$, $B\text{-edges} = \{(\mathbf{u}_j, \mathbf{x}_i) \mid B(i, j) \neq 0\}$, $C\text{-edges} = \{(\mathbf{x}_j, \mathbf{y}_i) \mid C(i, j) \neq 0\}$, $D\text{-edges} = \{(\mathbf{u}_j, \mathbf{y}_i) \mid D(i, j) \neq 0\}$. Here, $M(i, j)$ is the $(i, j)^{th}$ element of matrix M and $(\mathbf{v}_1, \mathbf{v}_2)$ denotes a directed edge from vertex $\mathbf{v}_1 \in \mathcal{V}$ to vertex $\mathbf{v}_2 \in \mathcal{V}$. Each edge is associated to a free non-zero parameter of the system model called the weight of the edge.

Notation : The edges associated to the unknown parameters are represented in dashed lines. The set of these edges is denoted \mathcal{E}_θ and their number is k_θ .

Example 3.1 To the system defined by the following matrices, it is associated the digraph depicted in Figure 1.

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \lambda_1 \\ 0 & \lambda_2 & 0 & 0 & 0 & \lambda_3 \\ \lambda_4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_5 & \lambda_6 & 0 \\ 0 & 0 & 0 & \lambda_7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \lambda_8 \end{pmatrix}, C = \begin{pmatrix} \lambda_9 & 0 & \lambda_{10} & 0 & 0 & 0 \\ 0 & \lambda_{11} & 0 & \lambda_{12} & 0 & 0 \end{pmatrix}, D = 0. \text{ For this example, the unknown parameters to be identified are } \Theta = \{\lambda_1, \lambda_2, \lambda_7, \lambda_{10}\}. \text{ We recall that all the other parameters } \Lambda \setminus \Theta \text{ are assumed to be known.}$$

□

Let's now give some useful definitions and notations which are quite common in the graph-theory.

Definition 1: a path is simple when every vertex occurs only once in this path. The weight of a path is the product of the weight of all the edges constituting this path.

Definition 2: a cycle is a path of the form $\mathbf{v}_{s_0} \rightarrow \mathbf{v}_{s_1} \rightarrow \dots \rightarrow \mathbf{v}_{s_i} \rightarrow \mathbf{v}_{s_0}$, where $\mathbf{v}_{s_0}, \mathbf{v}_{s_1}, \dots, \mathbf{v}_{s_i}$ are distinct.

Definition 3: a path is a **Y-topped** path if its end vertex belongs to \mathbf{Y} .

In the sequel, \mathcal{V}_1 and \mathcal{V}_2 represent two subsets of \mathcal{V} . It is denoted by $\text{card}(\mathcal{V}_1)$ the cardinality of \mathcal{V}_1

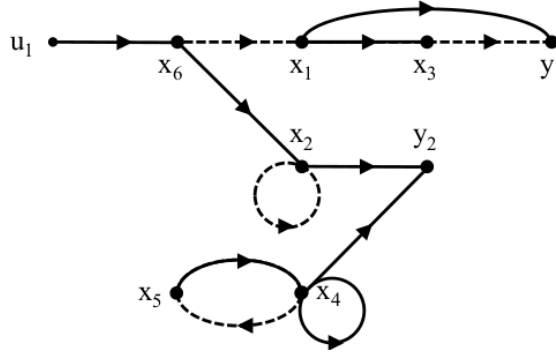


Figure 1. The digraph associated to the system of Example 3.1

and $\mathcal{V}_1 \setminus \mathcal{V}_2$ is the set of elements in \mathcal{V}_1 which are not in \mathcal{V}_2 .

Definition 4: a path is a \mathcal{V}_1 -rooted path if its begin vertex belongs to \mathcal{V}_1 .

Definition 5: denote by $pred(\mathcal{V}_1)$ the set of all the predecessors of the vertices included in \mathcal{V}_1 .

Definition 6: a path $p = \mathbf{v}_{s_0} \rightarrow \mathbf{v}_{s_1} \rightarrow \dots \rightarrow \mathbf{v}_{s_i}$ is said a \mathcal{V}_1 - \mathcal{V}_2 path if $\mathbf{v}_{s_0} \in \mathcal{V}_1$ and $\mathbf{v}_{s_i} \in \mathcal{V}_2$. Moreover, if the only vertex of p which belongs to \mathcal{V}_1 is \mathbf{v}_{s_0} and the only vertex of p which belongs to \mathcal{V}_2 is \mathbf{v}_{s_i} , p is called a direct \mathcal{V}_1 - \mathcal{V}_2 path.

Definition 7: some paths are disjoint if they have no common vertex. $\rho(\mathcal{V}_1, \mathcal{V}_2)$ is the maximal number of disjoint \mathcal{V}_1 - \mathcal{V}_2 paths. Moreover, a set of $\rho(\mathcal{V}_1, \mathcal{V}_2)$ disjoint \mathcal{V}_1 - \mathcal{V}_2 paths is called maximal \mathcal{V}_1 - \mathcal{V}_2 linking.

Definition 8: $\mu(\mathcal{V}_1, \mathcal{V}_2)$ denotes the minimal number of vertices of $\mathbf{U} \cup \mathbf{X} \cup \mathbf{Y}$ belonging to a maximal \mathcal{V}_1 - \mathcal{V}_2 linking.

Definition 9: $V_{ess}(\mathcal{V}_1, \mathcal{V}_2)$ is the vertex subset including the vertices present in all the maximal \mathcal{V}_1 - \mathcal{V}_2 linkings.

Definition 10: there exists a unique vertex subset noted $S^o(\mathcal{V}_1, \mathcal{V}_2)$, called minimum output separator, which is the set of begin vertices of all direct $V_{ess}(\mathcal{V}_1, \mathcal{V}_2)$ - \mathcal{V}_2 paths.

Example 3.1, continued: For the considered system, it can be seen that $\rho(\mathbf{U}, \mathbf{Y}) = 1$, $V_{ess}(\mathbf{U}, \mathbf{Y}) = \{\mathbf{u}_1, \mathbf{x}_6\}$, $S^o(\mathbf{U}, \mathbf{Y}) = \{\mathbf{x}_6\}$. \square

In order to express more easily the graphical necessary and sufficient conditions ensuring the parameters identifiability, a second kind of graphs called bipartite graphs will be used. Hereafter, some recalls are given on this kind of graphs and particularly on the so-called Dulmage-Mendelsohn decomposition. Consider any bipartite digraph noted \mathcal{B} defined by the triplet $(\mathbf{V}^+, \mathbf{V}^-, W)$, where \mathbf{V}^+ and \mathbf{V}^- are two vertex subsets and W the set of edges linking these two vertex subsets. Let's recall some elements related to the subdivision of such bipartite digraph into strongly connected components using the Dulmage-Mendelsohn (DM) decomposition (Dulmage and Mendelsohn 1958, Murota 1987):

A matching in bipartite graph \mathcal{B} is an edge set $M \subseteq W$ such that all the edges of M are disjoint (have no common vertex). A matching is maximal if it has a maximal cardinality. It can be associated to each maximal matching M a non bipartite digraph noted $\mathcal{B}_M = (\mathbf{V}^+, \mathbf{V}^-, W_M)$ where $(\mathbf{v}_1, \mathbf{v}_2) \in W_M \Leftrightarrow (\mathbf{v}_1, \mathbf{v}_2) \in W$ or $(\mathbf{v}_2, \mathbf{v}_1) \in M$. The set of vertices in \mathbf{V}^+ (resp. in \mathbf{V}^-) covered by the edges of M is denoted by $\partial^+ \mathbf{M}$ (resp. $\partial^- \mathbf{M}$). It is noted $\mathbf{S}_0^+ = \mathbf{V}^+ \setminus \partial^+ \mathbf{M}$ and $\mathbf{S}_0^- = \mathbf{V}^- \setminus \partial^- \mathbf{M}$.

In \mathcal{B}_M , two vertices \mathbf{v}_i and \mathbf{v}_j are said to be strongly connected if it exists in \mathcal{B}_M a path from \mathbf{v}_i to \mathbf{v}_j and a path from \mathbf{v}_j to \mathbf{v}_i . The relation "is strongly connected" is an equivalence relation and its equivalence classes can be defined. Each equivalent class is called a strongly connected component (Murota 1987).

Any bipartite digraph \mathcal{B} can be decomposed uniquely into a certain number (defined here as $\nu + 2$) of partially ordered irreducible components denoted $C_i(\mathcal{B}) = (\mathbf{V}_i^+(\mathcal{B}), \mathbf{V}_i^-(\mathcal{B}), W_i(\mathcal{B}))$, $i = 0, 1, \dots, \nu, \infty$, using DM-decomposition described below to make the paper self-contained:

\Leftrightarrow Find a maximal matching M in \mathcal{B} ,

\Leftrightarrow Let $\mathbf{V}_0^+(\mathcal{B}) = \mathbf{S}_0^+ \cup \{\mathbf{v} \in \mathbf{V}^+, \mid \exists \text{ a path in } \mathcal{B}_M \text{ from } \mathbf{S}_0^+ \text{ to } \mathbf{v}\}$

$\Leftrightarrow \mathbf{V}_0^-(\mathcal{B}) = \{\mathbf{v} \in \mathbf{V}^-, \mid \exists \text{ a path in } \mathcal{B}_M \text{ from } \mathbf{S}_0^+ \text{ to } \mathbf{v}\}$.

$\leftrightarrow W_0(\mathcal{B}) \stackrel{def}{=} \{\text{edges linking } \mathbf{V}_0^+(\mathcal{B}) \text{ to } \mathbf{V}_0^-(\mathcal{B})\}.$

$\leftrightarrow \text{Let } \mathbf{V}_\infty^+(\mathcal{B}) = \{\mathbf{v} \in \mathbf{V}^+, | \exists \text{ a path in } \mathcal{B}_M \text{ from } \mathbf{v} \text{ to } \mathbf{S}_0^-\}$

$\leftrightarrow \mathbf{V}_\infty^-(\mathcal{B}) = \mathbf{S}_0^- \cup \{\mathbf{v} \in \mathbf{V}^-, | \exists \text{ a path in } \mathcal{B}_M \text{ from } \mathbf{v} \text{ to } \mathbf{S}_0^-\}.$

$\leftrightarrow W_\infty(\mathcal{B}) \stackrel{def}{=} \{\text{edges linking } \mathbf{V}_\infty^+(\mathcal{B}) \text{ to } \mathbf{V}_\infty^-(\mathcal{B})\}.$

\leftrightarrow For $i = 1, \dots, \nu$, let $C_i(\mathcal{B})$ be the strongly connected component of the graph obtained from \mathcal{B}_M after deleting the vertices and the edges of $C_0(\mathcal{B})$ and $C_\infty(\mathcal{B})$.

\leftrightarrow A partial order relation is defined on the strongly connected components, denoted “ \preceq ” as follows: $C_i(\mathcal{B}) \preceq C_j(\mathcal{B}) \Leftrightarrow$ there exists a path starting from the vertices of $C_j(\mathcal{B})$ to the ones of $C_i(\mathcal{B})$ in \mathcal{B}_M .

It is important to note that the obtained subsets $\mathbf{V}_i^+(\mathcal{B})$ and $\mathbf{V}_i^-(\mathcal{B})$, for $i = 0, 1, \dots, \nu, \infty$, are the same whatever the choice of the maximal matching M is (Dulmage and Mendelsohn 1958, Murota 1987). $C_0(\mathcal{B})$ is called the **minimal inconsistent part** of \mathcal{B} and $C_\infty(\mathcal{B})$ is the **maximal inconsistent part** of \mathcal{B} . The union of all the strongly components $C_i(\mathcal{B})$, $i = 1, \dots, \nu$ is called the **consistent part** of \mathcal{B} .

4 Preliminaries and specific definitions

The identifiability of a set of parameters depends on the existence of some algebraic relations linking the output and their derivatives and, which are sensitive to these parameters. Therefore, it is important to characterize the existence and the number of algebraic relations linking the different output components and their derivatives. Knowing that the identifiability of the parameters depends on the existence of enough redundancy output equations sensitive to the unknown parameters, the method follows three main steps :

Step 1: A reduction of the digraph by the elimination of the part where there is no redundancy equations. This gives also a necessary identifiability condition

Step 2: A graphical characterization of the redundancy equations sensitive to the unknown parameters.

Step 3: A graphical test to know if the existing redundancy output equations allows the identifiability of all the unknown parameters.

Almost all the technical aspects of method turn around the graphical characterization of the existence and the independence of output redundancy equations. And this is strongly linked to the observability subspaces dimensions and so the existence and the length of some particular paths and cycles in the digraph associated to the system.

The existence of an algebraic equation linking some output components is equivalent to the fact that two generated observability subspaces are not disjoint, which can be characterized by some dimension calculations.

Our aim, hereafter, is to characterize the generic dimension of the observability subspace of the system or more exactly of the part of the system where it is possible to write redundancy output equations useful for the identification. To do so, the system is first subdivided into two parts from the observability point of view: the first one is the maximal over-determined part and the second one is related to the square part and the under-determined part. The parameters identifiability can be achieved only in the over-determined part. This subdivision is provided in (van der Woude 1999) and is equivalent to the DM-decomposition when representing the system by a bipartite digraph.

Definition 4.1: (Boukhobza et al. 2007) In digraph $\mathcal{G}(\Sigma_\Lambda)$, it can be defined:

- $\mathbf{X}_1 \stackrel{def}{=} \{\mathbf{x}_i | \rho(\mathbf{U} \cup \{\mathbf{x}_i\}, \mathbf{Y}) > \rho(\mathbf{U}, \mathbf{Y})\};$
- $\mathbf{Y}_0 \stackrel{def}{=} \mathbf{Y} \cap V_{ess}(\mathbf{U}, \mathbf{Y});$
- $\mathbf{Y}_1 \stackrel{def}{=} \mathbf{Y} \setminus \mathbf{Y}_0;$
- $\mathbf{U}_0 \stackrel{def}{=} \{\mathbf{u}_i | \text{there is no edge from } \mathbf{u}_i \text{ to } \mathbf{X}_1 \cup \mathbf{Y}_1\};$
- $\mathbf{U}_1 \stackrel{def}{=} \mathbf{U} \setminus \mathbf{U}_0;$
- $\mathbf{X}_s \stackrel{def}{=} S^o(\mathbf{U}_0, \mathbf{Y}) \cap \mathbf{X};$
- $\mathbf{X}_0 \stackrel{def}{=} \mathbf{X} \setminus (\mathbf{X}_1 \cup \mathbf{X}_s).$

□

The aim of this decomposition is to simplify the digraph of the system and to reduce it to its useful part for the identifiability study. This decomposition is also used in the observability context (van der Woude 1999, Boukhobza et al. 2007, Boukhobza and Hamelin 2013). Mainly, due to the presence of unknown inputs in the systems, from the observability point of view, any system can be divided into three parts: under-determined part, exactly-determined part and over-determined part. The two first parts are here regrouped to form the subsystem (Σ_0) . The over-determined part of the system constitutes the subsystem (Σ_1) . The only part containing redundancy output equations leading to the identification of unknown parameters are the over-determined part. This is why, only this part is kept and the study can be reduced to (Σ_1) the over-determined part and the rest of the digraph is removed from the study. Moreover, if an unknown parameter appears in the under-determined part or in the exactly determined part of the system, then it won't be identifiable. More precisely, the decomposition leads to two subsystems:

Subsystem (Σ_1) which has as input $U_1 \cup X_s$, state X_1 and output Y_1 . This subsystem represents, from the observability point of view, the over-determined part of the system. Indeed, it is proven in (Boukhobza et al. 2007, van der Woude 1999) that this part remains left invertible even if any one of its measurements is removed. This is due to the fact that, by definition, no output vertex is essential in an input-output linking and all the state vertices constituting this subsystem can reach an output vertex independently from the input-output paths. It has necessarily more outputs (measurements) than inputs.

Subsystem (Σ_0) which has as input U_0 , state $X_0 \cup X_s$ and output Y_0 . This subsystem represents the square and under-determined parts of the system.

In fact, if it is denoted by X_0 , X_s , X_1 , U_0 , U_1 , Y_0 and Y_1 the state, the unknown input and the output associated to vertex subsets X_0 , X_s , X_1 , U_0 , U_1 , Y_0 and Y_1 respectively, system (Σ_Λ) can be written as given in (Boukhobza et al. 2007):

$$\begin{cases} \dot{X}_0 = A_{0,0}X_0 + A_{0,s}X_s + A_{0,1}X_1 + B_{0,0}U_0 + B_{0,1}U_1 \\ \dot{X}_s = A_{s,0}X_0 + A_{s,s}X_s + A_{s,1}X_1 + B_{s,0}U_0 + B_{s,1}U_1 \\ \dot{X}_1 = A_{1,s}X_s + A_{1,1}X_1 + B_{1,1}U_1 \\ Y_0 = C_{0,0}X_0 + C_{0,s}X_s + C_{0,1}X_1 + D_{0,0}U_0 + D_{0,1}U_1 \\ Y_1 = C_{1,s}X_s + C_{1,1}X_1 + D_{1,1}U_1 \end{cases} \quad (3)$$

The proposed subdivision of the system into two subsystems is summarized in Figure 2. As it can be seen through equations (3), X_0 and U_0 does not intervene on the dynamics of X_1 and neither on the output Y_1 . Thus, system (Σ_1) can be seen as having U_1 and X_s as input vectors, X_1 as state vector and Y_1 as output vector. This system is over-determined and has more outputs than inputs (Boukhobza and Hamelin 2009). System (Σ_0) , has more inputs than outputs and is under-determined. It can be considered that this system interacts with subsystem (Σ_1) as U_1 and X_1 can be present in the dynamics of X_0 and X_s as well in the output equation Y_0 .

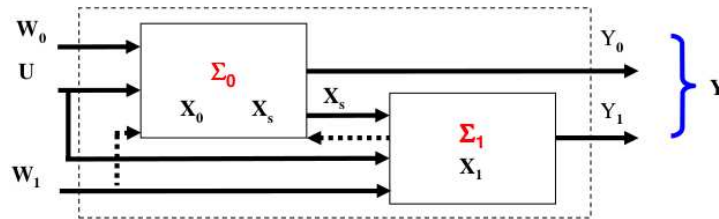


Figure 2. System subdivision for parameters identifiability

Example 3.1, continued: For this example, we have $X_1 = \{x_1, x_2, x_3, x_4, x_5\}$, $X_0 = \emptyset$, $Y_0 = \emptyset$, $Y_1 = \{y_1, y_2\}$, $U_0 = \{u_1\}$, $U_1 = \emptyset$ and $X_s = \{x_6\}$. \square

Remark: When all the inputs are known, or equivalently measured, then subsystem (Σ_0) contains only the state vertices which are not connected to the measurements.

Some important results proven in (Boukhobza and Hamelin 2009, 2011) are summarized in the following lemmas:

Lemma 4.2: *There cannot exist a relation linking only the output components of square subsystem (Σ_0) .* \square

Proof: See Appendix A.

Using this lemma result, the studied graph can be reduced to keep only its useful part *i.e.* only the vertices related to subsystem (Σ_1) : $\mathbf{U}_1 \cup \mathbf{X}_s \cup \mathbf{X}_1 \cup \mathbf{Y}_1$ and the edges linking these vertices. Note that the elimination of subsystem (Σ_0) does not correspond to the elimination of a subsystem, where there is no unknown parameter to identify. Indeed, if there are some parameters to identify related to some edges in (Σ_0) , then these parameters are not identifiable and it can be concluded that the considered subset of parameters is not identifiable. **Therefore, from now, only the over-determined part is considered *i.e.* in the sequel, our study is reduced only to the part (Σ_1) .**

Example 3.1, continued: *For the system of Example 3.1, the reduced useful digraph is depicted in Figure 3.* \square

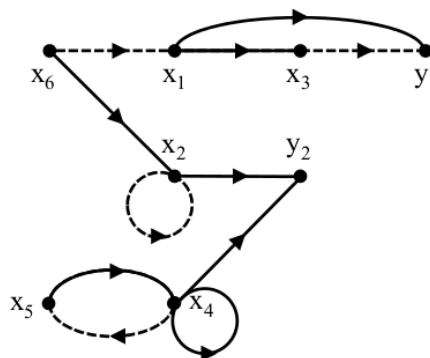


Figure 3. The reduced useful digraph for Example 1

Lemma 4.3: *For subsystem (Σ_1) , for each $\mathbf{Y}_u = \{y_{i_1}, y_{i_2}, \dots, y_{i_k}\} \subset \mathbf{Y}_1$ such that $\rho(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_u) = \text{card}(\mathbf{U}_1 \cup \mathbf{X}_s)$, there exist generically a matrix G , a function φ and an integer $v \leq n_1$ such that $(X_s^T, U_1^T)^T = \varphi(Y_u, \dot{Y}_u, \dots, Y_u^{(v)}) + GX_1$, where $Y_u^{(v)} = (y_{i_1}^{(v)}, y_{i_2}^{(v)}, \dots, y_{i_k}^{(v)})^T$.* \square

Proof: This lemma has been proved in (Boukhobza and Hamelin 2011).

As a corollary of this lemma, we have:

Lemma 4.4: *For each $y_i \in \mathbf{Y}_1$, there exists an output algebraic relation linking it to other measurements of \mathbf{Y}_1 .* \square

Proof: See Appendix B.

Definition 4.5: Consider structured linear system (Σ_Λ) associated to digraph $\mathcal{G}(\Sigma_\Lambda)$. The positive number β_1 is defined as the maximal number of vertices of $\mathbf{X}_1 \cup \mathbf{U}_1 \cup \mathbf{X}_s$ covered by a disjoint union of:

- a $(\mathbf{U}_1 \cup \mathbf{X}_s) \rightarrow \mathbf{Y}_1$ linking of maximal size;
 - \mathbf{Y}_1 -topped paths;
 - cycles covering only the elements of \mathbf{X}_1 .
- \square

Lemma 4.6: *Consider structured linear system (Σ_Λ) represented by digraph $\mathcal{G}(\Sigma_\Lambda)$, the number β_1 is equal to the generic dimension of the observable subspace of the over-determined part (Σ_1) of the system in the extended state and input space.* \square

Proof: This lemma has been proved in (Boukhobza 2010).

To achieve the parameter identifiability analysis, it is necessary to have enough independent output algebraic relations linking some outputs and their derivatives, which are sensitive to the parameters to be identified. When several output components appear in this equation, it means that for some order k_i , $y_i^{(k_i)}$ can be written as a linear combination of $y_i^{(k)}$, $k < k_i$, and of other output components with their derivatives. In this case, the intersection of observability subspace generated by y_i and the other

components is not empty. This can be characterized by the generic dimension of such subspaces.

Definition 11: for each output $\mathbf{y}_i \in \mathbf{Y}_1$,

$$\Delta_i = \{\mathbf{x}_i \in \mathbf{X}_1, \text{ such that } \mathbf{x}_i \text{ is covered by a direct } S^o(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_1 \setminus \{\mathbf{y}_i\}) \rightarrow \mathbf{Y}_1 \setminus \{\mathbf{y}_i\} \text{ path}\}$$

Definition 12: integer $\beta_{1,\bar{i}}$ is the sum of :

- $\mu(\mathbf{U}_1 \cup \mathbf{X}_s, S^o(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_1 \setminus \{\mathbf{y}_i\}))$
- the maximal number of vertices of $\Delta_i \cup S^o(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_1 \setminus \{\mathbf{y}_i\})$ covered by a disjoint union of:
 - ◊ a $S^o(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_1 \setminus \{\mathbf{y}_i\}) - \mathbf{Y}_1 \setminus \{\mathbf{y}_i\}$ linking of maximal size;
 - ◊ $\mathbf{Y}_1 \setminus \{\mathbf{y}_i\}$ -topped paths;
 - ◊ cycles covering only the elements of Δ_i .

Definition 13: $d(\mathbf{y}_i) = \beta_1 - \beta_{1,\bar{i}}$.

$\beta_{1,\bar{i}}$ represents the dimension of the observability subspace of subsystem (Σ_1) when removing the output y_i Boukhobza (2010). Obviously, for all $k > d(\mathbf{y}_i)$, $y_i^{(k)}$ can be written as a linear combination of $y_i^{(j)}$, with $j \leq d(\mathbf{y}_i)$, and other output components with their derivatives.

Definition 14: a path p is said eligible if it covers only the vertices included in $\mathbf{U}_1 \cup \mathbf{X}_s \cup \mathbf{X}_1 \cup \mathbf{Y}_1$ and if it verifies one of the following conditions:

- p is a simple $\mathbf{U}_1 \cup \mathbf{X}_s - \mathbf{Y}_1$ path.
- p is a Y -topped path and its length is strictly greater than $d(\mathbf{y}_i)$, where \mathbf{y}_i is its end vertex.
- p is a cycle covering only vertices in \mathbf{X}_1 .

Definition 15: an edge is eligible if it belongs to an eligible path.

Example 3.1, continued: For the system of Example 1 reduced to its over-determined part, $\beta_1(\{\mathbf{y}_1, \mathbf{y}_2\}) = 6$ and as $\beta_{1,\bar{i}} = \beta_{1,\bar{i}} = 2$, then $d(\{\mathbf{y}_1\}) = d(\{\mathbf{y}_2\}) = 4$. The eligible paths are then all the $\{\mathbf{u}_1\}$ - \mathbf{Y} simple paths and all the cycles covering $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ and \mathbf{x}_5 . The eligible edges of the system of Example 3.1 are depicted in Figure 4. \square

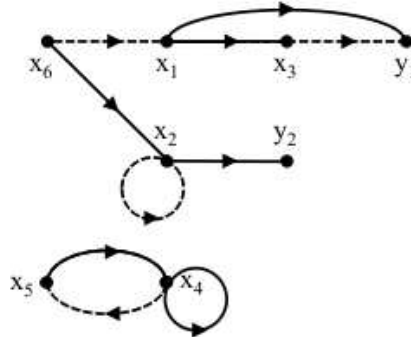


Figure 4. Eligible edges for Example 1

5 Intermediate results for the identifiability study

Using the previous definitions, the following lemma is stated:

Lemma 5.1: Let's consider structured linear system (Σ_Λ) represented by graph $\mathcal{G}(\Sigma_\Lambda)$. An unknown parameter λ_e corresponding to an edge $e \in \mathcal{E}_\theta$ is structurally identifiable only if e is eligible. \square

Proof: The proof is provided in Appendix C.

Now, the following sets are defined in $\mathcal{G}(\Sigma_\Lambda)$ and an equivalence relation \mathcal{R} between some specific paths is introduced:

Definition 15: $\Gamma_d = \{\mathbf{v}_i \text{ covered by an eligible path such that } S^o(\mathbf{v}_i, \mathbf{Y}_1) = \{\mathbf{v}_i\} \text{ when considering the subgraph constituted only by the eligible edges}\}$

Definition 16: $\Gamma_c = \{\mathbf{x}_i \in \mathbf{X}_1, | \mathbf{x}_i \text{ is covered by a cycle}\}$

Definition 17: in the set of eligible Γ_d - Γ_d direct and simple paths, relation \mathcal{R}_1 is defined as follows : $p_1 \mathcal{R}_1 p_2$ if p_1 and p_2 are two paths having the same begin and end vertices and the same length.

Definition 18: \mathbb{A} is the set of pairs of eligible Γ_d - Γ_d direct and simple paths $\{p_1, p_2\}$ such that $p_1 \mathcal{R}_1 p_2$. In \mathbb{A} , relation \mathcal{R}_p is defined as follows : $\{p_1, p_2\} \mathcal{R}_p \{p_3, p_4\}$ if one of these two conditions holds:

- $p_1 \mathcal{R}_1 p_3$ and $p_2 \mathcal{R}_1 p_4$
- $p_1 \mathcal{R}_1 p_4$ and $p_2 \mathcal{R}_1 p_3$

Definition 19: \mathbb{C} is the set of pairs of paths $\{C_1, p'_1\}$ such that:

- C_1 is a cycle covering elements of Γ_c ,
- p'_1 is a simple direct Γ_c - Γ_d path, and
- C_1 and p'_1 have the same begin vertex.

In \mathbb{C} , relation \mathcal{R}_c is defined as follows : $\{C_1, p'_1\} \mathcal{R}_c \{C_2, p'_2\}$ if C_1 and C_2 have the same length and p'_1 and p'_2 have the same end-vertex.

Example 3.1, continued: For this example, $\Gamma_d = \{\mathbf{x}_6\}$ and $\Gamma_c = \{\mathbf{x}_2, \mathbf{x}_4, \mathbf{x}_5\}$.

Furthermore, it can be seen that $\mathbf{x}_6 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{y}_1 \mathcal{R}_1 \mathbf{x}_6 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{x}_3 \rightarrow \mathbf{y}_1$, $\mathbf{x}_6 \rightarrow \mathbf{x}_2 \rightarrow \mathbf{y}_2 \mathcal{R}_1 \mathbf{x}_6 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{x}_3 \rightarrow \mathbf{y}_1$ and $\mathbf{x}_6 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{y}_1 \mathcal{R}_1 \mathbf{x}_6 \rightarrow \mathbf{x}_2 \rightarrow \mathbf{y}_2$. For relation \mathcal{R}_p , three equivalent classes have to be considered. The first class includes $\{\mathbf{x}_6 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{y}_1, \mathbf{x}_6 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{x}_3 \rightarrow \mathbf{y}_1\}$, the second one includes $\{\mathbf{x}_6 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{y}_1, \mathbf{x}_6 \rightarrow \mathbf{x}_2 \rightarrow \mathbf{y}_2\}$ and the last one includes $\{\mathbf{x}_6 \rightarrow \mathbf{x}_2 \rightarrow \mathbf{y}_2, \mathbf{x}_6 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{x}_3 \rightarrow \mathbf{y}_1\}$. Moreover, $\{\mathbf{x}_2 \rightarrow \mathbf{x}_2, \mathbf{x}_2 \rightarrow \mathbf{y}_2\} \mathcal{R}_c \{\mathbf{x}_4 \rightarrow \mathbf{x}_4, \mathbf{x}_4 \rightarrow \mathbf{y}_2\}$ and $\{\mathbf{x}_2 \rightarrow \mathbf{x}_2, \mathbf{x}_2 \rightarrow \mathbf{y}_2\} \mathcal{R}_c \{\mathbf{x}_4 \rightarrow \mathbf{x}_5 \rightarrow \mathbf{x}_4, \mathbf{x}_4 \rightarrow \mathbf{y}_2\}$. \square

Lemma 5.2: Let's consider structured linear system (Σ_Λ) represented by graph $\mathcal{G}(\Sigma_\Lambda)$. An unknown parameter corresponding to an edge $e \in \mathcal{E}_\theta$ belongs to a redundancy algebraic equation iff it is covered by a path belonging to a pair of \mathbb{A} or by a cycle belonging to a pair of \mathbb{C} . \square

Proof: The proof is provided in Appendix D.

The following identifiability necessary conditions can be stated:

Lemma 5.3: Let's consider structured linear system (Σ_Λ) represented by digraph $\mathcal{G}(\Sigma_\Lambda)$. k unknown parameters corresponding respectively to k edges belonging to \mathcal{E}_θ are structurally identifiable only if they are covered by paths belonging to pairs of \mathbb{A} or by cycles belonging to pairs of \mathbb{C} , where all these pairs belong to k distinct equivalence classes w.r.t. relations \mathcal{R}_p and \mathcal{R}_c . \square

Proof: The proof is provided in Appendix E.

It is defined now the bipartite digraph linking the parameters to the paths which allows to write the identifiability equations:

Definition 5.4: Considering edge subset \mathcal{E}_θ , which corresponds to the unknown parameters and a union of pairs $\mathcal{F} = \{\{p_{1,1}, p_{1,2}\}, \{p_{2,1}, p_{2,2}\}, \dots, \{p_{k,1}, p_{k,2}\}, \{c_1, p'_1\}, \{c_2, p'_2\}, \dots, \{c_\ell, p'_\ell\}\}$, a bipartite digraph $\mathcal{B}_I(\mathcal{F})$ called “**identifiability bipartite digraph related to \mathcal{F}** ” can be defined, such that: $\mathcal{B}_I(\mathcal{F}) = (\mathbf{V}^+, \mathbf{V}^-, W)$, where $\mathbf{V}^+ = \Theta$, $\mathbf{V}^- = \mathcal{F}$ and $W = \{(\lambda_j, \{p_{i,1}, p_{i,2}\}), \text{ if the edge associated to } \lambda_j \text{ is covered by a path } p_{j,1} \text{ or } p_{j,2} \text{ of a pair } \{p_{j,1}, p_{j,2}\} \in \{p_{i,1}, p_{i,2}\} \mathcal{R}_p\} \cup \{(e_j, \{c_i, p'_i\}), \text{ if the edge associated to } \lambda_j \text{ is covered by a cycle } c_j \text{ of a pair } \{c_j, p'_j\} \in \{c_i, p'_i\} \mathcal{R}_c\}$. \square

The two following lemmas establish the necessary and sufficient identifiability conditions using the bipartite graph defined above.

Lemma 5.5: Let's consider a union of pairs $\mathcal{F} \subseteq \mathbb{A} \cup \mathbb{C}$, where all the pairs belong to distinct equivalence classes w.r.t. relations \mathcal{R}_p and \mathcal{R}_c . If all the parameters associated to elements of \mathcal{E}_θ are covered in $\mathcal{B}_I(\mathcal{F})$ by an exactly determined DM-component of cardinality equal to one, then they are structurally identifiable, assuming that all the other parameters of the model are known. \square

Proof: The Proof is provided in Appendix F.

Lemma 5.6: All the parameters associated to \mathcal{E}_θ are structurally identifiable, assuming that all the

other parameters of the model are known, only if there exists a union of pairs $\mathcal{F} \subseteq \mathbb{A} \cup \mathbb{C}$ belonging to distinct equivalence classes w.r.t. relations \mathcal{R}_p and \mathcal{R}_c , such that each element of \mathcal{E}_θ is covered in $\mathcal{B}_I(\mathcal{F})$ by an exactly determined DM-component of cardinality equal to one. \square

Proof: The Proof is provided in Appendix G.

6 Main result : parameter identifiability characterisation

Using the previous definitions and results, it can be enounced now the main result of the paper characterizing exactly the parameters identifiability:

Proposition 6.1: *Let's consider structured linear system (Σ_Λ) represented by graph $\mathcal{G}(\Sigma_\Lambda)$. The unknown parameters corresponding to the edge subset \mathcal{E}_θ are structurally identifiable, assuming that all the other parameters of the model are known, iff there exists a family $\mathcal{F} = \{\{p_{1,1}, p_{1,2}\}, \{p_{2,1}, p_{2,2}\}, \dots, \{p_{k,1}, p_{k,2}\}, \{c_1, p'_1\}, \{c_2, p'_2\}, \dots, \{c_\ell, p'_\ell\}\}$ of pairs included in $\mathbb{A} \cup \mathbb{C}$ such that:*

- All the edges of \mathcal{E}_θ are covered by paths $p_{i,j}$, $i = 1, \dots, k$, $j = 1, 2$ or by cycles c_r , $r = 1, \dots, \ell$,
- All the pairs $\{p_{i,1}, p_{i,2}\}$, $i = 1, \dots, k$, belong to distinct equivalent classes w.r.t. \mathcal{R}_p ,
- All the pairs $\{c_r, p'_r\}$, $r = 1, \dots, \ell$, belong to distinct equivalent classes w.r.t. \mathcal{R}_c ,
- All the D-M components of the identifiability bipartite digraph associated to \mathcal{F} are in the exactly determined part and each of them contains exactly one element of Θ (the vertex set associated to the unknown parameters). \square

Proof:

Sufficiency: According to lemma 5.5, the existence of family \mathcal{F} satisfying conditions of Proposition 6.1, which are similar to the conditions of Lemma 5.5, and covering all the edges corresponding to unknown parameters, is sufficient to establish the identifiability of all the unknown parameters. \triangle

Necessity: The necessity is immediate according to Lemma 5.6. \triangle

Example 3.1, continued: *For the system of Example 3.1, four parameters have to be identified $\{\lambda_1, \lambda_2, \lambda_7, \lambda_{10}\}$. Let's take the following notations for the paths and cycles: $p_1 = \mathbf{x}_6 \rightarrow \mathbf{x}_2 \rightarrow \mathbf{y}_2$, $p_2 = \mathbf{x}_6 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{x}_3 \rightarrow \mathbf{y}_1$, $p_3 = \mathbf{x}_6 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{y}_1$, $c_1 = \mathbf{x}_2 \rightarrow \mathbf{x}_2$, $c_2 = \mathbf{x}_4 \rightarrow \mathbf{x}_5 \rightarrow \mathbf{x}_4$, $p'_1 = \mathbf{x}_2 \rightarrow \mathbf{y}_2$ and $p'_2 = \mathbf{x}_4 \rightarrow \mathbf{y}_2$.*

All the elements of family $\mathcal{F} = \{\{p_1, p_3\}, \{p_2, p_3\}, \{c_1, p'_1\}, \{c_2, p'_2\}\}$ belong to distinct equivalent classes w.r.t. \mathcal{R}_p and \mathcal{R}_c . Moreover, each of the unknown parameters is covered by at least one path p_1 or p_2 or one cycle c_1 or c_2 . Figure 5 depicts identifiability bipartite digraph associated to \mathcal{F} . It can be seen that all the DM-components contain only one element because edge $(\lambda_1, \{p_2, p_3\})$ does not belong to any maximal matching in this bipartite digraph. Therefore, the four parameters are generically identifiable. Indeed, when writing the redundancy algebraic equation linking y_1 and their derivatives to y_2 , when the equation is normalized in order to obtain the coefficient multiplying $y_2^{(6)}$ equal to one, it can be seen that the coefficient multiplying $y_1^{(6)}$ depends on known parameters and λ_1 that allows to identify λ_1 ; the coefficient multiplying $y_1^{(5)}$ depends on known parameters, λ_1 and λ_{10} that allows to identify λ_{10} ; the coefficient multiplying $y_2^{(5)}$ depends on known parameters, λ_2 and λ_{10} that allows to identify λ_2 ; finally, the coefficient multiplying $y_2^{(4)}$ depends on λ_2, λ_7 and λ_{10} that allows to identify λ_7 . Note that, if λ_9 was unknown, λ_1 and λ_{10} won't be identifiable since they would belong to the same DM-component. In this case, to recover the identifiability of these parameters, a measure has to be added on x_6, x_1 or x_3 . \square

7 Conclusion

In this paper, a graph-theoretic tool is proposed to analyse generic parameter identifiability for structured linear systems with unknown inputs. More precisely, new necessary and sufficient conditions are provided that ensure that the redundancy equations generated by the output measurements allow the

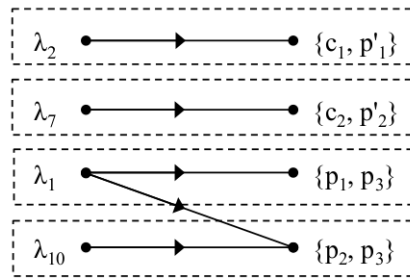


Figure 5. Identifiability bipartite digraph associated to \mathcal{F} for Example 1

parameter identifiability. Obviously, in this structural study, the different constraints on the inputs and initial conditions are not considered. It is just assumed that they are sufficiently informative and can take any desired values which allow the practical identification of the parameters. Our approach uses classical programming techniques and is free from numerical difficulties. This makes our proposed method well-suited to treat large scale systems. In future works, the graph-theoretic approach can be used when the proposed identifiability conditions are not satisfied, to determine the location of additional sensors in order to recover them.

References

- Audoly, S., and D'angiò, L. (1980), "On the identifiability of linear compartmental systems: a revisited transfer function approach based on topological properties," *Mathematical Biosciences*, 66, 201 – 228.
- Bellman, R., and Åstrom, K. (1970), "On structural identifiability," *Mathematical Biosciences*, 7, 329–339.
- Bellu, G., Saccomani, M., Audoly, S., and D'Angiò, L. (2007), "DAISY: A new software tool to test global identifiability of biological and physiological systems," *Computer Methods and Programs in Biomedicine*, 88, 52–61.
- Boukhobza, T. (2010), "Partial state and input observability recovering by additional sensor implementation: a Graph-Theoretic Approach," *International Journal of Systems Science*, 41, 1281 – 1291.
- Boukhobza, T., and Hamelin, F. (2009), "State and Input Observability Recovering by Additional Sensor Implementation: a Graph-Theoretic Approach," *Automatica*, 45, 1737–1742.
- Boukhobza, T., and Hamelin, F. (2011), "Observability of Switching Structured Linear Systems with Unknown input: a Graph-Theoretic Approach," *Automatica*, 47, 395–402.
- Boukhobza, T., and Hamelin, F. (2013), "Discrete mode observability of structured switching descriptor linear systems: A graph-theoretic approach," *Automatica*, 49, 3042?3048.
- Boukhobza, T., Hamelin, F., and Martinez-Martinez, S. (2007), "State and Input Observability for Structured Linear Systems: a Graph-Theoretic Approach," *Automatica*, 43, 1204–1210.
- Chapman, M., and Godfrey, K. (1985), "Some extensions to the exhaustive-modeling approach to structural identifiability," *Mathematical Biosciences*, 77, 305–323.
- Delforge, J. (1984), "Sur l'identifiabilité et l'identification des modèles linéaires," PhD thesis, Université d'Angers.
- Dion, J.M., Commault, C., and van der Woude, J.W. (2003), "Generic Properties and Control of Linear Structured Systems: a Survey," *Automatica*, 39, 1125–1144.
- Dulmage, A.L., and Mendelsohn, N.S. (1958), "Coverings of Bipartite Graphs," *Canadian Journal of Mathematics*, 10, 517–534.
- Eisenfeld, J. (1982), "New techniques for structural identifiability for large linear and nonlinear compartmental systems," *Mathematics and Computers in Simulation*, 24, 494 – 501.
- Eisenfeld, J., and Grundy, S. (1983), "Structural identification of large systems by reduction to subsystems: VLDL triglycerides," *Mathematical Biosciences*, 66, 129 – 149.
- Miao, H., Xia, X., Perelson, A., and Wu, H. (2011), "On Identifiability of Nonlinear ODE Models and

- Applications in Viral Dynamics,” *SIAM review*, 53, 3–39.
- Murota, K., *System Analysis by Graphs and Matroids*, New York, U.S.A.: Springer-Verlag (1987).
- Reinschke, K.J., *Multivariable Control. A Graph Theoretic Approach.*, New York, U.S.A.: Springer-Verlag (1988).
- Reinschke, K. (1985), “Structural properties of linear control systems should be investigated by means of graphs,” *Annual Review in Automatic Programming*, 12, 280–283.
- van der Woude, J.W. (1999), “The Generic Number of Invariant Zeros of a Structured Linear System,” *SIAM Journal of Control and Optimization*, 38, 1–21.
- Xia, X., and Moog, C. (2003), “Identifiability of nonlinear systems with applications to hiv/aids models,” *IEEE Transactions on Automatic Control*, 48, 330–336.

Appendix A: Proof of Lemma 4.2

Using results shown in (van der Woude 1999) (Theorem 5.1) concerning the case of square systems, the generic number of invariant zeros of the pencil matrix of system (Σ_0) is equal to $\text{card}(\mathbf{X}_0) + \text{card}(\mathbf{X}_s) + \text{card}(\mathbf{U}_0) - (\mu(\mathbf{U}_0, \mathbf{X}_s \cup \mathbf{Y}_0) - \rho(\mathbf{U}_0, \mathbf{X}_s \cup \mathbf{Y}_0) + \text{card}(\mathbf{X}_s))$. Thus, the dimension of the strongly observable subspace and so the number of possible independent observation equations is equal to $\mu(\mathbf{U}_0, \mathbf{X}_s \cup \mathbf{Y}_0) - \rho(\mathbf{U}_0, \mathbf{X}_s \cup \mathbf{Y}_0)$. Moreover, the cardinality of $\mathbf{U}_0 \cup \mathbf{X}_0$ representing the unknown variables for this system, is at least equal to $\mu(\mathbf{U}_0, \mathbf{X}_s \cup \mathbf{Y}_0) - \rho(\mathbf{U}_0, \mathbf{X}_s \cup \mathbf{Y}_0)$ because all the linkings between \mathbf{U}_0 and $\mathbf{X}_s \cup \mathbf{Y}_0$ cover the vertices of $\mathbf{U}_0 \cup \mathbf{X}_0$. Since, (Σ_0) cannot be partitioned in smaller independent rectangular and square subsystems, in the best case, the number of independent observation equations is equal to the number of unknown variables. Thus, unknown variables cannot be eliminated from any equations to obtain a relation linking only Y_0 and X_s and their derivatives. \triangle

Appendix B: Proof of Lemma 4.4

$y_i \notin V_{ess}(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_1)$ implies that there exists necessarily a subset $\mathbf{Y}_u \subset \mathbf{Y}_1 \setminus \{y_i\}$ such that $\rho(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_u) = \text{card}(\mathbf{U}_1 \cup \mathbf{X}_s)$. Using the previous lemma, there exist a function φ and an integer $\nu \geq 0$ such that, $(X_s^T, U_1^T)^T = \varphi(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) + GX_1$. Substituting this in subsystem (Σ_1) , we have:

$$\begin{cases} \dot{X}_1 = (A_{1,1} + (A_{1,s} B_{1,1})G)X_1 + \varphi_x(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) \\ \stackrel{\text{def}}{=} \tilde{A}X_1 + \varphi_x(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) \\ Y_1 = (C_{1,1} + (C_{1,s} D_{1,1})G)X_1 + \varphi_y(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) \\ \stackrel{\text{def}}{=} \tilde{C}X_1 + \varphi_y(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) \end{cases} \quad (\text{B1})$$

Consider then that the characteristic equation of matrix \tilde{A} has the form $\tilde{A}^{n_1} + \dots + a_k \tilde{A}^k + \dots + a_1 \tilde{A} + a_0 \tilde{A}^0 = 0$, where $n_1 = \text{card}(\mathbf{X}_1)$. Multiplying this equation by \tilde{C}_i where \tilde{C}_i is the line of matrix \tilde{C} related to output y_i in (B1), an equation of the following form is obtained:

$$(\tilde{C}_i \tilde{A}^{n_1} + \dots + a_k \tilde{C}_i \tilde{A}^k + \dots + a_0 \tilde{C}_i)X_1 = 0 \quad (\text{B2})$$

Since y_i is an output of (Σ_1) , at least a term $\tilde{C}_i \tilde{A}^k$ of the latter equation is not zero and depends on y_i . Then, (B2) leads to an output algebraic equation where some derivatives of y_i intervene. \triangle

Appendix C: Proof of Lemma 5.1

Every algebraic redundancy equation consists in expressing the k^{th} derivative of an output component y_i according to the other output components and/or the lower y_i derivatives. So, it is of the form $y_i^{(k)} =$

$$\sum_{s=0}^{k-1} \alpha_{i,s} y_i^{(s)} + \sum_{l|y_l \in \mathbf{Y}} \sum_{s=0}^{n_l} \alpha_{l,s} y_l^{(s)}.$$

If all the coefficients $\alpha_{l,s}$ with $l \neq i$ are equal to zero, this implies that the considered redundancy equation is necessarily issued from the polynomial characteristic of matrix A . In this case, according to the results proved in (Murota 1987), coefficients $\alpha_{i,s}$ are all issued from the cycles constituting the graph. If at least one of the coefficients $\alpha_{l,s}$, $l \neq i$ is non zero, this implies that, for all $\ell \geq k$, $y_i^{(\ell)}$ can be expressed using the $k - 1$ first derivatives of y_i and the other output derivatives. In this case, from the observability subspaces dimensions, we have equivalently $d(\mathbf{y}_i) \leq k$. This implies that the only paths which traduce the k^{th} derivative of y_i with $k \geq d(\mathbf{y}_i)$ can be associated to a redundancy equation. As $y_i^{(k)}$ is represented by y_i -topped paths of length k , $y_i^{(k)}$ is not sensitive to the coefficients representing the weight of the paths which are not $\mathbf{U} \cup \mathbf{X}_s$ -rooted paths and are shorter than $d(\mathbf{y}_i)$. Thus, when an edge is not eligible, its associated parameter cannot belong to a redundancy equation and then is not identifiable. \triangle

Appendix D: Proof of Lemma 5.2

Let's denote by $e \in \mathcal{E}_\theta$ the edge associated to the unknown parameter denoted λ_e .

Sufficiency:

First case: If e is covered by a cycle belonging to a pair of \mathcal{Q}_C then it is also covered by a cycle C in (Σ_1) . Let's denote by \mathbf{x}_1 the begin vertex of e . Since this vertex belongs to $\mathbf{X}_1(\mathbf{Y})$, there exist a $\mathbf{x}_1 \rightarrow \mathbf{y}_j$ path, where $\mathbf{y}_j \in \mathbf{Y}_1$ and a subset $\mathbf{Y}_u \subset \mathbf{Y}_1 \setminus \{\mathbf{y}_j\}$ with $\rho(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_u) = \text{card}(\mathbf{U}_1 \cup \mathbf{X}_s) < \rho(\mathbf{U}_1 \cup \mathbf{X}_s \cup \{\mathbf{x}_1\}, \mathbf{Y}_u \cup \{\mathbf{y}_j\})$. From Lemma 4.3, there exist a matrix G , a function φ and an integer $\nu \geq 0$ such that $(X_s^T, U_1^T)^T = \varphi(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) + GX_1$. Substituting this in subsystem (Σ_1) and using notations of (B1), we obtain:

$$\begin{cases} \dot{X}_1 = (A_{1,1} + [A_{1,s} \ B_{1,1}]G)X_1 + (A_{1,s} \ B_{1,1})\varphi(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) \\ \quad \stackrel{\text{def}}{=} \tilde{A}X_1 + \varphi_x(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) \\ Y_1 = (C_{1,1} + [C_{1,s} \ D_{1,1}]G)X_1 + (C_{1,s} \ D_{1,1})\varphi(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) \\ \quad \stackrel{\text{def}}{=} \tilde{C}X_1 + \varphi_y(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) \end{cases} \quad (\text{D1})$$

Since the elements of G are represented by the edges belonging to \mathbf{Y}_u -topped paths, they are independent from the elements of \tilde{A} associated to cycle C . Thus, the digraph representation of (D1) contains also C . Therefore, from (Reinschke 1988) (Theorem 21.1), the characteristic equation of matrix \tilde{A} has the form $\tilde{A}^{n_1} + \dots + a_k \tilde{A}^k + \dots + a_1 \tilde{A} + a_0 \tilde{A}^0 = 0$, where $n_1 = \text{card}(\mathbf{X}_1)$. It contains a term $a_{n_1 - \bar{k}} \tilde{A}^{n_1 - \bar{k}}$, where \bar{k} is the length of C and $a_{n_1 - \bar{k}}$ depends on the product of the weights of the edges constituting C and so, on the specific entry λ_e of matrix A according to the fact that edge e belongs to C . Thus,

$$(\tilde{C}_j \tilde{A}^{n_1} + \dots + a_k \tilde{C}_j \tilde{A}^k + \dots + a_1 \tilde{C}_j \tilde{A} + a_0 \tilde{C}_j)X_1 = 0 \quad (\text{D2})$$

where \tilde{C}_j is the line of matrix \tilde{C} related to output y_j in (D1) i.e. $y_j = \tilde{C}_j X_1 + \varphi_{y,j}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)})$. On the one hand, since \mathbf{x}_1 belongs to a cycle and as there exists a \mathbf{x}_1 - \mathbf{y}_j path, if $\ell > 0$ denotes the length of this path then, $\forall k \geq \ell - 1$, $\tilde{C}_j \tilde{A}^k \neq 0$. Furthermore, since $\ell + \bar{k} \leq n_1$, then term $a_{n_1 - \bar{k}} \tilde{C}_j \tilde{A}^{n_1 - \bar{k}} \neq 0$. On the other hand, due to expression (D1), for all $k \geq 0$ $\tilde{C}_j \tilde{A}^k X_1 = y_j^{(k)} + \varphi_y(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu+k)})$. Thus, (D2) leads to an equation of the form:

$$y_j^{(n_1+1)} + \dots + a_k y_j^{(k)} + \dots + a_0 y_j + \nu(Y_u, \dots, Y_u^{(2n_1)}) = 0 \quad (\text{D3})$$

where ν is a function depending on functions φ_y and φ_x . Therefore, as $a_{n_1 - \bar{k}}$ depends on λ_e , the latter intervenes in the redundancy algebraic equation (D2).

Second case: If e is covered by a Γ_d - Γ_d path belonging to a pair of \mathcal{Q}_A , and as all these paths are eligible, two possibilities have to be considered: 1) e is covered by a $\mathbf{U}_1 \cup \mathbf{X}_s$ - \mathbf{Y}_1 path or 2) e is covered by a

\mathbf{y}_i -topped path, which length is strictly greater than $d(\mathbf{y}_i)$. In the first case, we denote by \mathbf{v}_i the vertex of $\mathbf{U}_1 \cup \mathbf{X}_s$ from which there is a direct $\mathbf{X}_s \cup \mathbf{U}_1 - \mathbf{Y}_1$ path p of length ℓ . Since \mathbf{v}_i belongs to $\mathbf{U}_1 \cup \mathbf{X}_s$, it exists a subset $\mathbf{Y}_u = \{\mathbf{y}_{i_1}, \mathbf{y}_{i_2}, \dots, \mathbf{y}_{i_k}\} \subset \mathbf{Y}_1 \setminus \{\mathbf{y}_i\}$ such that $\rho(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_u) = \text{card}(\mathbf{U}_1 \cup \mathbf{X}_s)$ without using edges and vertices of p . From Lemma 4.3, there exist a matrix G , a function φ and an integer $\nu \leq n_1$ such that the dynamics equation of subsystem (Σ_1) is in the form of (D1). Since the elements of G are represented by edges belonging to \mathbf{Y}_u -topped paths, they are independent from the elements of \tilde{A} . From the characteristic equation of matrix \tilde{A} , an equation as (D2) can be written where also \tilde{C}_j is the line of matrix \tilde{C} related to output y_j in (D1) i.e. $y_j = \tilde{C}_j X_1 + \varphi_{y,j}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)})$. Using equations of (D1), we have $\tilde{C}_j \tilde{A}^k X_1 = y_j^{(k)} - \tilde{C}_j (\varphi_x^{(k-1)}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) + \tilde{A} \varphi_x^{(k-2)}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) + \dots + \tilde{A}^{k-2} \dot{\varphi}_x(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)})) - \varphi_{y,j}^{(k)}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)})$. Since there exists a path of length ℓ from \mathbf{v}_i to \mathbf{y}_j containing an edge e , $C_j \tilde{A}^{\ell-1}$ is not zero, for all $k \geq \ell$, and depends on the element λ_e of A . Moreover, $\varphi_x^{(k-\ell)}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)})$ is not zero for all $k \geq \ell$, because there exists a path between \mathbf{v}_i and an element of \mathbf{Y}_u , as by definition of \mathbf{Y}_u , $\rho(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_u) = \text{card}(\mathbf{U}_1 \cup \mathbf{X}_s)$. This implies that, $C_j \tilde{A}^{\ell-1} \varphi_x^{(k-\ell)}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)})$ is not zero for all $k \geq \ell$ and depends on λ_e . In particular, this implies that term $\tilde{C}_j \tilde{A}^{n_1}$ in relation (D2) is not zero and depends on λ_e . Thus, substituting term $\tilde{C}_j \tilde{A}^k X_1$ by $y_j^{(k)} - \tilde{C}_j (\varphi_x^{(k-1)}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) + \tilde{A} \varphi_x^{(k-2)}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)}) + \dots + \tilde{A}^{k-2} \dot{\varphi}_x(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)})) - \varphi_{y,j}^{(k)}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)})$ in relation (D2), an output algebraic relation is obtained knowing that term $C_j \tilde{A}^{\ell-1} \varphi_x^{(k-\ell)}(Y_u, \dot{Y}_u, \dots, Y_u^{(\nu)})$ depends on λ_e .

The second possibility is when e is covered by a \mathbf{y}_i -topped path which length is strictly greater than $d(\mathbf{y}_i)$. As e is also covered by a $\Gamma_d - \mathbf{y}_i$ path, this implies that there exists a \mathbf{y}_i -topped path p of length strictly greater than $d(\mathbf{y}_i)$ which covers e and such that \mathbf{x}_ℓ belongs to a direct $S^o(\{\mathbf{v}_p\}, \mathbf{Y}_1) - \mathbf{y}_i$ path, where $\mathbf{v}_p = \mathbf{x}_j$ is the begin vertex of the eligible path p . \mathbf{y}_i belongs to \mathbf{Y}_1 implies that there exists a vertex subset $\mathbf{Y}_u \subseteq \mathbf{Y}_1 \setminus \{\mathbf{y}_i\}$ such that $\rho(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_u) = \text{card}(\mathbf{U}_1 \cup \mathbf{X}_s)$. From Lemma 4.3, this implies that there exist a matrix G , a function φ and an integer $\nu \leq n_1$ such that the dynamics equation of subsystem (Σ_1) can be put under the form of (D1). Moreover, by definition of $d(\mathbf{y}_i)$, the dimension of the observability subspace is equal to $d(\mathbf{y}_i)$ plus the dimension of the observability subspace without y_i . This means that the $(d(\mathbf{y}_i))^{th}$ derivative of y_i can be expressed using the other output components and the lower derivatives of y_i . Thus, $\forall k \geq d(\mathbf{y}_i)$, there exists a minimal subset $\tilde{\mathbf{Y}} \subseteq \mathbf{Y}_1 \setminus (\mathbf{Y}_u \cup \{\mathbf{y}_i\})$ such that $y_i^{(k)}$ is a linear combination of the first derivatives of $y_i^{(s)}$, with $s = 0, \dots, k - 1$ and all the possible derivatives (until the dimension of the under-determined part) of components y_l of $\tilde{\mathbf{Y}}$. Thus, there exist some constant real parameters denoted $\alpha_{i,s}$ and $\alpha_{l,s}$ and a function v such that:

$$y_i^{(k)} = \sum_{s < k_i} \alpha_{i,s} y_i^{(s)} + \sum_{l | y_l \in \tilde{\mathbf{Y}}} \sum_{s=0}^{n_1} \alpha_{l,s} y_l^{(s)} + v(Y_u, \dots, Y_u^{(2n_1)}) \quad (D4)$$

where $n_1 = \text{card}(\mathbf{X}_1)$. Using the system equations described in (D1), it can be deduced that for each s and l , $y_i^{(s)} = \tilde{C}_i \tilde{A}^s X_1 + v_{i,s}(Y_u, \dots, Y_u^{(n_1+s)})$ and $y_l^{(s)} = \tilde{C}_l \tilde{A}^s X_1 + v_{l,s}(Y_u, \dots, Y_u^{(n_1+s)})$, where \tilde{C}_i (resp. \tilde{C}_l) is the row of \tilde{C} corresponding to output y_i (resp. y_l). Substituting the latter equalities in (D4), all functions of Y_u and its derivatives can be simplified as they are independent from the considered elements of X_1 and (D4) can be written as:

$$\tilde{C}_i \tilde{A}^k X_1 = \sum_{s < d(\mathbf{y}_i)} \alpha_{i,s} \tilde{C}_i \tilde{A}^s X_1 + \sum_{l | y_l \in \tilde{\mathbf{Y}}} \sum_{s=0}^{n_1} \alpha_{l,s} \tilde{C}_l \tilde{A}^s X_1 \quad (D5)$$

Since subset $\tilde{\mathbf{Y}}$ is minimal i.e. $\forall \mathbf{y}_j \in \tilde{\mathbf{Y}}$, the difference of dimensions the observability subspaces when considering respectively $\mathbf{Y}_u \cup \tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\} \setminus \{\mathbf{y}_j\}$ and $\mathbf{Y}_u \cup \tilde{\mathbf{Y}} \setminus \{\mathbf{y}_j\}$ as output components is strictly greater than k_i , then in relation (D5), all the components of $\tilde{\mathbf{Y}}$ intervene. Let's denote by \mathbf{e}_j the j^{th} Euclidean vector. As this relation is valid whatever state vector X_1 is, it can be removed and by post-multiplying

relation (D5) by e_j , it comes:

$$\tilde{C}_i \tilde{A}^k e_j = \left(\sum_{s < d(\mathbf{y}_i)} \alpha_{i,s} \tilde{C}_i \tilde{A}^s + \sum_{l | \mathbf{y}_l \in \tilde{\mathbf{Y}}} \sum_{s=0}^{n_1} \alpha_{l,s} \tilde{C}_l \tilde{A}^s \right) e_j \quad (\text{D6})$$

where each non-zero component of $\tilde{C}_l \tilde{A}^s$ is associated to the paths arriving to $\mathbf{y}_l \in \tilde{\mathbf{Y}}$ of length $s + 1$. Since all the $\{\mathbf{x}_j\} - \tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\}$ paths starting from \mathbf{x}_j cover by definition $S^o(\{\mathbf{x}_j\}, \tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\})$ denoted in the sequel \mathbf{x}_r , then there exist k_r and k' such that $k_r + k' = k$ and $\tilde{C}_i \tilde{A}^k e_j = \tilde{C}_i \tilde{A}^{k_r} \Delta_r \tilde{A}^{k'} e_j$ where Δ_r is a diagonal matrix having only one non-zero element $\Delta_r(r, r) = 1$. The same reasoning can be done for each term $\tilde{C}_l \tilde{A}^s e_j$ and so there exist s_r and s' such that $s_r + s' = s$ and $\tilde{C}_l \tilde{A}^s e_j = \tilde{C}_l \tilde{A}^{s_r} \Delta_r \tilde{A}^{s'} e_j$. The fact that e belongs to a $S^o(\mathbf{v}_p, \tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\}) - \tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\}$ path means that edge e appears in only some of $S^o(\mathbf{v}_p, \tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\}) - \tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\}$ paths. Thus, some terms $\tilde{C}_i \tilde{A}^{k_r}$ and $\tilde{C}_l \tilde{A}^{s_r}$, but not all of them, contain the non-zero parameter λ_e corresponding to edge e . Denoting by $C_r = e_r^T$, where e_r is the r^{th} Euclidean vector, we have that $\tilde{C}_i \tilde{A}^k e_j = \tilde{C}_i \tilde{A}^{k_r} \Delta_r \tilde{A}^{k'} e_j = \alpha'_{i,k} C_r \tilde{A}^{k'} e_j$ and $\tilde{C}_l \tilde{A}^s e_j = \tilde{C}_l \tilde{A}^{s_r} \Delta_r \tilde{A}^{s'} e_j = \alpha'_{l,s} C_r \tilde{A}^{s'} e_j$, where $\alpha'_{l,s} \stackrel{\text{def}}{=} (\tilde{C}_l \tilde{A}^{s_r} \Delta_r) (C_r C_r^T)^{-1} = \tilde{C}_l \tilde{A}^{s_r} \Delta_r$. Thus, after substitution of the previous terms in relation (D6), it comes:

$$\alpha'_{i,k} C_r \tilde{A}^{k'} e_j = \left(\sum_{s_r \leq s < d(\mathbf{y}_i)} \alpha'_{i,s} \alpha_{i,s} C_r \tilde{A}^{s-s_r} + \sum_{l | \mathbf{y}_l \in \tilde{\mathbf{Y}}} \sum_{s=s_r}^{n_1} \alpha'_{l,s} \alpha_{l,s} C_r \tilde{A}^{s-s_r} \right) e_j \quad (\text{D7})$$

where some, but not all coefficients $\alpha'_{i,k}$, $\alpha'_{i,s}$ and $\alpha'_{l,s}$ depend on λ_e . This weight cannot be factorized and simplified because all the coefficients do not depend on it (some $S^o(\mathbf{v}_p, \tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\}) - \tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\}$ paths do not contain edge e). Therefore, (D7) is valid only if some of the coefficients $\alpha'_{i,k}$, $\alpha_{i,s}$ and $\alpha_{l,s}$ depend also on the weight λ_e of e . Thus, by means of equation (D4) containing coefficients $\alpha_{i,s}$ and $\alpha_{l,s}$, an output algebraic relation depending on λ_e is obtained.

Necessity: Firstly, when edge e is not covered by an eligible cycle or by a $\mathbf{U}_1 \cup \mathbf{X}_s - \mathbf{Y}_1$ path, it has been shown in Lemma 5.1 that it is not possible to obtain an output algebraic relation achieving the parameter identifiability from the characteristic equation of any matrix \tilde{A} in (D1) whatever \mathbf{Y}_u is.

Moreover, when e is not covered by an eligible cycle or by a $\mathbf{U}_1 \cup \mathbf{X}_s - \mathbf{Y}_1$ path, for any possible choice of \mathbf{Y}_u , $\varphi_y(Y_u, \dot{Y}_u, \dots, Y_u^{(v)})$ and $\varphi_x(Y_u, \dot{Y}_u, \dots, Y_u^{(v)})$ do not depend on the specific parameter associated to e . Furthermore, any output relation can be put on the form of (D4) for some integer k , output component \mathbf{y}_i and a set of output components $\tilde{\mathbf{Y}}$, which does not include \mathbf{y}_i and are disjoint from \mathbf{Y}_u . Thus, the only way to have an identification of unknown parameter λ_e is that there exists at least one coefficient $\alpha_{i,s}$ or $\alpha_{l,s}$ of (D4) depending on λ_e . Without loss of generality, it can be assumed that $\tilde{\mathbf{Y}}$ is minimal and so that in the output algebraic relations all the components of $\tilde{\mathbf{Y}}$ appear. If it is not the case, they are removed from $\tilde{\mathbf{Y}}$. Relation (D4) implies that the observable subspace obtained using outputs $\{\mathbf{y}_i\} \cup \mathbf{Y}_u$ and the one generated by $\tilde{\mathbf{Y}} \cup \mathbf{Y}_u$ have non zero intersection. Defining \tilde{k}_i as the difference between the dimension of the observability subspace when considering as outputs $\tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\} \cup \mathbf{Y}_u$ and the dimension of the observability subspace considering as outputs $\tilde{\mathbf{Y}} \cup \mathbf{Y}_u$. Obviously, $\tilde{k}_i \geq d(\mathbf{y}_i)$, relation (D4) cannot be satisfied for some $k \leq \tilde{k}_i - 1$. Thus, in output algebraic relation (D4), $k \geq \tilde{k}_i$. Furthermore, as $\tilde{\mathbf{Y}}$ is minimal, it is not possible to express $y_i^{(\tilde{k}_i)}$ using only a part of $\tilde{\mathbf{Y}}$. To guarantee identifiability of λ_e , there must exist at least one \mathbf{y}_i -topped path p , of length greater or equal to $\tilde{k}_i + 1$ and so of length strictly greater than $d(\mathbf{y}_i)$, associated to this relation which covers λ_e . Let's denote by \mathbf{x}_j the begin vertex of p (it cannot be a $\mathbf{U}_1 \cup \mathbf{X}_s - \mathbf{Y}_1$ path when the Lemma is not satisfied) and e_j the j^{th} Euclidean vector. When e is not covered by an eligible $\mathbf{U}_1 \cup \mathbf{X}_s - \mathbf{Y}_1$ path, in Equation (D4), $v(Y_u, \dot{Y}_u, \dots, Y_u^{(n)})$ does not depend on λ_e . Then, (D4) can be rewritten as relation (D6) *i.e.*

$$\tilde{C}_i \tilde{A}^k e_j = \left(\sum_{s < d(\mathbf{y}_i)} \alpha_{i,s} \tilde{C}_i \tilde{A}^s + \sum_{\mathbf{y}_l \in \tilde{\mathbf{Y}}} \sum_{s=0}^{n_1} \alpha_{l,s} \tilde{C}_l \tilde{A}^s \right) e_j \quad (\text{D8})$$

where $\tilde{C}_l \tilde{A}^s$ is associated to y_l -topped paths and $\mathbf{y}_l \in \tilde{\mathbf{Y}}$ is of length $s + 1$. All the $\{\mathbf{x}_j\}$ - $\tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\}$ paths cover $\mathbf{x}_r = S^o(\mathbf{x}_j, \tilde{\mathbf{Y}} \cup \{\mathbf{y}_i\})$. So, if e is not covered by a path belonging to a pair of \mathbb{Q}_A or by a cycle belonging to a pair of \mathbb{Q}_C , then it is neither covered by a cycle, nor by a $\mathbf{U}_1 \cup \mathbf{X}_s$ - \mathbf{Y}_1 path nor by a $S^o(\{\mathbf{v}_p\}, \mathbf{Y}_1)$ - \mathbf{y}_i path. Then, the paths starting from \mathbf{x}_r to \mathbf{Y} do not contain e . Thus, all the redundancy equations can be written similarly to (D7):

$$\alpha' C_r \tilde{A}^k e_j = \left(\sum_{s_r \leq s < d(\mathbf{y}_i)} \alpha'_{i,s} \alpha_{i,s} C_r \tilde{A}^{s-s_r} + \sum_{l | \mathbf{y}_l \in \tilde{\mathbf{Y}}} \sum_{s=s_r}^{n_1} \alpha'_{l,s} \alpha_{l,s} C_r \tilde{A}^{s-s_r} \right) e_j \quad (\text{D9})$$

where α' and $\alpha'_{l,s}$ do not depend on λ_e . Equation (D9) implies that there exists an output algebraic relation with the use of only virtual output $C_r x$ in addition to Y_u . This relation can be due to the characteristic equation of matrix \tilde{A} , which includes only the terms related to cycles in the digraph. When e does not belong to an eligible cycle, also all α' and $\alpha'_{l,s}$ do not depend on λ_e and it is also the case for coefficients $\alpha_{i,s}$ and $\alpha_{l,s}$ in equation (D4). So, all the existing relations of the form (D4) do not contain λ_e . The same reasoning can be done for all $\mathbf{Y}_u \subseteq \mathbf{Y}_1$ s.t. $\rho(\mathbf{U}_1 \cup \mathbf{X}_s, \mathbf{Y}_u) = \text{card}(\mathbf{U}_1 \cup \mathbf{X}_s)$ and so λ_e cannot be identified. Δ

Appendix E: Proof of Lemma 4.6

According to Lemma 5.2, k unknown parameters corresponding respectively to k edges belonging to \mathcal{E}_θ are structurally identifiable only if they are all covered by paths belonging to pairs of \mathbb{Q}_A or by a cycle belonging to a pair of \mathbb{Q}_C . Therefore, to prove Lemma 5.3, it has to be shown that two parameters are structurally identifiable only if the associated edges are covered by two paths, parts of elements of $\mathbb{Q}_A \cup \mathbb{Q}_C$ belonging to distinct equivalence classes w.r.t. relations \mathcal{R}_p and \mathcal{R}_c . Assume that it is not the case i.e. there are two parameters λ_{r_1} and λ_{r_2} associated to two edges e_1 and e_2 , which are covered by two elements of $\mathbb{Q}_A \cup \mathbb{Q}_C$ belonging to the same equivalence classes w.r.t. relations \mathcal{R}_p or \mathcal{R}_c . There are four possibilities to this fact:

- e_1 and e_2 are covered by elements of \mathbb{Q}_A belonging to the same equivalent class w.r.t. relation \mathcal{R}_p and these two edges can be covered by the same Γ_d - Γ_d direct path or equivalently by the same Γ_d - \mathbf{Y} path. Let's denote \mathbf{x}_j the begin and \mathbf{y}_i the end vertices of this Γ_d - \mathbf{Y} path. In this case, redundancy equation (D4) where y_i intervenes can also be written on the form (D6) where the only term depending on λ_{r_1} and λ_{r_2} is $\tilde{C}_i \tilde{A}^k e_j$. However, since e_1 and e_2 are covered by the same path, then $\tilde{C}_i \tilde{A}^s$ depends on the product $\lambda_{r_1} \lambda_{r_2}$. Thus, the redundancy equation may allow the identifiability of the product $\lambda_{r_1} \lambda_{r_2}$ but not of λ_{r_1} and λ_{r_2} separately.

- e_1 and e_2 are covered by elements of \mathbb{Q}_A belonging to the same equivalent class w.r.t. relation \mathcal{R}_p and these two edges can be covered by two different Γ_d - Γ_d direct paths having the same begin and end vertices or equivalently by two different Γ_d - \mathbf{Y} paths having the same begin and end vertices. Let's denote \mathbf{x}_j the begin and \mathbf{y}_i the end vertices of this Γ_d - \mathbf{Y} path. In this case, redundancy equation (D4) where y_i intervenes can also be written on the form (D6) where the only term on the form $\tilde{C}_i \tilde{A}^s$ depending on λ_{r_1} and λ_{r_2} is $\tilde{C}_i \tilde{A}^k e_j$ because e_1 and e_2 are covered by two paths of same length. More precisely, $\tilde{C}_i \tilde{A}^s$ depends on a combination of λ_{r_1} and λ_{r_2} . Thus, the redundancy equation may allow the identifiability of this combination but not of λ_{r_1} and λ_{r_2} separately.

- e_1 and e_2 are covered by elements of \mathbb{Q}_C belonging to the same equivalent class w.r.t. relation \mathcal{R}_c and these two edges can be covered by the same cycle noted C . In this case, the redundancy algebraic equation, where λ_{r_1} and λ_{r_2} appear, is issued from the polynomial characteristic of matrix \tilde{A} (D1). From (Reinschke 1988) (Theorem 21.1), the characteristic equation of matrix \tilde{A} has the form $\tilde{A}^{n_1} + \dots + a_k \tilde{A}^k + \dots + a_1 \tilde{A} + a_0 \tilde{A}^0 = 0$, where $n_1 = \text{card}(\mathbf{X}_1)$, coefficients $a_k = \sum_{c_i | c_i \text{ is a loop covering } k \text{ vertices}} (-1)^{n_k} \text{weight}(c_i)$ and

n_k is the number of loops¹ covering k vertices. Here, since e_1 and e_2 are covered by the same cycle, in all the coefficients a_k where appear λ_{r_1} and λ_{r_2} , the weights of the loops depend on the product $\lambda_{r_1}\lambda_{r_2}$. Thus, multiplying the characteristic equation by C_i , an algebraic redundancy equation as (D2) is obtained, which makes possible the identifiability of the product $\lambda_{r_1}\lambda_{r_2}$ but not of λ_{r_1} and λ_{r_2} separately.

- e_1 and e_2 are covered by elements of \mathbb{Q}_C belonging to the same equivalent class w.r.t. relation \mathcal{R}_c and these two edges can be covered by different cycles of the same length and connected to the same vertex of Γ_d . Here also, the redundancy algebraic equation, where λ_{r_1} and λ_{r_2} appear, is issued from the polynomial characteristic of matrix \tilde{A} of system (D1). But all the coefficients of such polynomial depend on the products of the weight of different cycles or on the sum of the weights of cycles having the same length. Therefore, multiplying the characteristic equation by C_i , an algebraic redundancy equation as (D2) is obtained, where it is not possible to distinguish the value of λ_{r_1} and the value of λ_{r_2} . Thus these two values are not identifiable even if their sum and their product are known.

Consequently, if two parameters λ_{r_1} and λ_{r_2} are not identifiable when they associated to two edges e_1 and e_2 covered by two elements of $\mathbb{Q}_A \cup \mathbb{Q}_C$ belonging to the same equivalence classes w.r.t. relations \mathcal{R}_p or \mathcal{R}_c and the lemma follows. \triangle

Appendix F: Proof of Lemma 5.5

If all the DM-component cardinalities are equal to one in $\mathcal{B}_I(\mathcal{F})$ and if each element of \mathcal{F} can be associated to an equation making possible the identification of its linked parameters in $\mathcal{B}_I(\mathcal{F})$, then using the DM-decomposition properties, these equations constitute a triangular system. In this case, all the parameters are obviously structurally identifiable. Consequently, to establish the lemma, it has to be proven that each element of \mathcal{F} can be associated to an equation making possible the identification of its linked parameters in $\mathcal{B}_I(\mathcal{F})$.

Consider first that for an element $\{c_1, p'_1\} \in \mathbb{Q}_C$, c_1 covers a set of parameters $\lambda_{r_1}, \dots, \lambda_{r_k}$. Since this cycle is connected to an element of Γ_d by p'_1 and since all the vertices of Γ_d are linked to at least one output vertex, then the considered cycle is linked to at least one output vertex. Let's denote by y_i one of these outputs. The redundancy algebraic equations, where $\lambda_{r_1}, \dots, \lambda_{r_k}$ appear, are issued from the polynomial characteristic of matrix \tilde{A} (D1). From (Reinschke 1988) (Theorem 21.1), the characteristic equation of matrix \tilde{A} has the form $\tilde{A}^{n_1} + \dots + a_k\tilde{A}^k + \dots + a_1\tilde{A} + a_0\tilde{A}^0 = 0$, where $n_1 = \text{card}(\mathbf{X}_1)$, coefficients $a_k = \sum_{c_i \text{ a loop covering } k \text{ vertices}} (-1)^{n_k} \text{weight}(c_i)$ and n_k is the number of loops covering k vertices.

This equation can be multiplied by C_i to obtain redundancy algebraic equation of the form (D2). One identification equation can be written for each coefficient a_k of this redundancy algebraic equation. From this equation, for each length of cycle and so for each equivalent class w.r.t. relation \mathcal{R}_c , an identifiability equation can be deduced. Therefore, the obtained equation (noted Eq_{c_i}) related to $\{c_1, p'_1\}$ has then as unknown variables the sum of products of weights of all the cycles belonging to the pairs included in the same equivalent class than $\{c_1, p'_1\}$ w.r.t. relation \mathcal{R}_c .

Consider now the case where an element $\{p_{i,1}, p_{i,2}\} \in \mathbb{Q}_A$, c_1 covers a set of parameters $\lambda_{r_1}, \dots, \lambda_{r_k}$ in $\mathcal{B}_I(\mathcal{F})$. This element corresponds to Γ_d - Γ_d paths, and so, as all the elements of Γ_d are linked to vertices of \mathbf{Y}_1 , this element corresponds also to Γ_d - \mathbf{Y}_1 paths having distinct end vertices. Let's denote \mathbf{y}_{i_1} and \mathbf{y}_{i_2} the output end vertices of these paths and ℓ_1, ℓ_2 the length of these paths. It has been shown, in Lemma 5.2, that these paths can be associated to a redundancy algebraic equation on the form of Equation (D4) when terms $\alpha_{i_1, k_1} y_{i_1}^{(k_1)}$ and $\alpha_{i_2, k_2} y_{i_2}^{(k_2)}$ appear, with $k_1 \geq \ell_1$ and $k_2 \geq \ell_2$. From this equation, containing K parameters, it can be deduced K identification equations for each coefficient. Using the same arguments than in the proof of Lemma 5.2, coefficients α_{i_1, k_1} and α_{i_2, k_2} are respectively sensitive (inversely proportional) to the sum of the weights of all the paths belonging to the pairs included in the same equivalent class than $\{p_{i,1}, p_{i,2}\}$ w.r.t. relation \mathcal{R}_p . Therefore, the obtained equation (noted Eq_{p_i}) related to $\{p_{i,1}, p_{i,2}\}$ has then as unknown variables the sum of products of weights of all the cycles belonging to the pairs included in

¹A loop is a collection of disjoint cycles

the same equivalent class than $\{p_{i,1}, p_{i,2}\}$ w.r.t. relation \mathcal{R}_p . △

Appendix G: Proof of Lemma 5.6

From Lemma 5.2, only edges covered by eligible Γ_d -rooted paths or eligible cycles can lead to redundancy equations. Thus, a parameter which is associated to an edge not covered by such paths is not identifiable. Moreover, from Lemma 5.3, the k_θ unknown parameters are identifiable only if they are covered by k_θ paths and cycles belonging respectively to pairs of \mathbb{Q}_A and to pairs of \mathbb{Q}_C and where all these pairs belong to k_θ distinct equivalence classes w.r.t. relations \mathcal{R}_p and \mathcal{R}_c . This implies that the parameters are identifiable only if there exists a maximal matching of size k_θ in $\mathcal{B}_I(\mathcal{F})$, where \mathcal{F} denotes such union of k_θ pairs, and then there is no under-determined or over-determined DM-component in this bipartite digraph.

Assume that there exists in such bipartite digraph a DM-component of cardinality strictly greater than one. In this case, at least two redundancy algebraic equations depend at least on two unknown parameters. Let's denote these two parameters λ_{r_1} and λ_{r_2} . An equation issued from an element of \mathbb{Q}_A or \mathbb{Q}_C depending on these two parameters means that an identifiability equation depends on the product $\lambda_{r_1}\lambda_{r_2}$ or on one of ratios $\frac{\lambda_{r_1}}{\lambda_{r_2}}$ or $\frac{\lambda_{r_2}}{\lambda_{r_1}}$ (which are equivalent). Then, three cases have to be considered for these equations:

1. The two equations depend on the product $\lambda_{r_1}\lambda_{r_2}$. In this case, neither λ_{r_1} nor λ_{r_2} are identifiable.
2. The same conclusion is drawn when the two equations depend on the ratios $\frac{\lambda_{r_1}}{\lambda_{r_2}}$ or $\frac{\lambda_{r_2}}{\lambda_{r_1}}$.
3. One equation depends on the product $\lambda_{r_1}\lambda_{r_2}$ and the other on the ratio $\frac{\lambda_{r_1}}{\lambda_{r_2}}$. In this case, the parameters are identifiable. However, the first equation means that edges e_1 and e_2 associated respectively to λ_{r_1} and λ_{r_2} belong to cycle part c_1 of an element of \mathbb{Q}_C and the second implies that these two edges belong to two direct paths p_1 and p_2 constituting the elements of \mathbb{Q}_A . In this case, there exists necessarily a vertex v_1 in Γ_d covered by cycle c_1 such that a pair $(p'_1, p'_2) \in \mathbb{Q}_A$ covers only λ_{r_1} if e_1 is after e_2 in c_1 or λ_{r_2} if e_2 is after e_1 in c_1 . Thus, in \mathcal{F} , pair (p_1, p_2) can be substituted by (p'_1, p'_2) and we obtain in the bipartite digraph associated to this disjoint union that e_1 and e_2 belong to two distinct DM-components. The same reasoning can be done for all the DM-components containing more than one element. Thus, if the parameters are identifiable then there always exists a union of pairs $\mathcal{F} \subseteq \mathbb{Q}_A \cup \mathbb{Q}_C$, where all the pairs belong to distinct equivalence classes w.r.t. relations \mathcal{R}_p and \mathcal{R}_c , such that elements of \mathcal{E}_θ are each covered in $\mathcal{B}_I(\mathcal{F})$ by an exactly determined DM-component of cardinality equal to one and the lemma follows. △