Identifiability and Excitation of a Class of Rational Systems*

Alexandre S. Bazanella[1], Michel Gevers[2], Daniel Coutinho[3] and Rafael Rui[1]

***Abstract*— In a recent paper (Gevers et al., CDC 2013) [6] we have presented identifiability and excitation conditions for a class of linearly parametrized polynomial systems with a scalar state. The present paper expands the model class in two ways: we consider a class of rational systems with vector states that can be written in the so-called "generalized controller form". The identifiability for such systems can be determined using the Ritt algorithm as shown in the seminal paper (Ljung and Glad, 1994) [10]. Here we present an alternative and simple method which offers a lot of insight into the structural conditions on the model class that make it globally identifiable and into the generation of informative experiments.**

I. Introduction

The question of *identifiability* of a model structure can be stated as follows: does there exist an experiment such that the data collected from this experiment allow one to uniquely determine the parameter values? The seminal papers [**?**], [10] provided a broad answer to this question for large classes of linearly and nonlinearly parametrized systems using tools from differential algebra. It is important to note that identifiability is a property of the chosen model structure (i.e. the parametrization); it is independent of the true system and of the data. The question asked is: are there data that make the mapping from parameter space to output space injective? The question of finding such a data set is that of *informativity*: once it has been established that the model structure is identifiable, it becomes an experiment design problem.

The solution provided in [10] is based on the application of Ritt's algorithm to the model equations; see e.g. [4]. It is very powerful but very heavy on computations; for rather simple-looking model structures, it is often impossible to terminate the computations of Ritt's algorithm on a standard PC. Ritt's algorithm essentially produces a large set of differential algebraic equations that are equivalent to the initial model structure, by a long sequence of differentiations and substitutions. When the procedure terminates in the required form, it means that the model structure is equiavlent to a set of linear regression equations, where each parameter can be written as the ratio of two polynomial expressions in the input u, the output y and their derivatives. This indicates that the model class **may be** globally identifiable. Sufficient conditions for global identifiability are then obtained if one can produce an experiment that makes a number of these polynomials in u, y and their derivatives nonzero.

While the procedure of [10] is very powerful, it is very computer intensive and not very intuitive. Hence the desire to produce alternative methods for special classes of nonlinear systems that better reveal the connection between identifiability, informativity and the parametric structure of the model class. A large number of authors have produced such results, in particular in biological and biochemical systems. Let us just mention a few recent papers: [13], [12], [8], [2] as well as [11] which presents results for discrete-time systems.

Our present work follows the same objective. In [6] we have produced necessary and sufficient conditions for identifiability and informativity for a class scalar polynomial models. In the present paper, we expand the results of [6] to a model class with rational functions and vector states.

For this class of model structures, we first illustrate the approach of [10] based on the Ritt algorithm, with the aim of illustrating not just the linear regression equations that are obtained at the end of the algorithm, but also the identifiability analysis that follows. This analysis consists in finding experimental conditions on the input and/or the initial state that make some polynomial expressions in u, y and their derivatives nonzero.

We show that for the model class considered in this paper a necessary and sufficient condition for global identifiability is equivalent to a certain semi-infinite matrix being full rank. A sufficient condition for global identifiability is easily derived by requiring that a finite truncation of that matrix be of full rank. We demonstrate that this matrix can easily be factored into the product of two matrices that depend only on the structure of the model class, namely the dimension n of the state and the degree q of the highest degree polynomial. In other words, these matrices can be written by inspection from the model structure. An informative experiment is one that makes these matrices have full rank. One of the main advantages of the tools developed in this paper is that the full rank conditions on these matrices allow for a better understanding of the different experiment design scenarios that lead to global identifiability: excitation by the input and the initial condition, excitation by the initial condition only with no inputs, or excitation by the input whatever the unknown initial condition may be.

The paper is organized as follows. We first present the model class and some basic assumptions in Section II. Section III presents some background material on identifia-

*This work is supported by CNPq, by CAPES, by a joint programme of the Belgian FNRS and the Brazilian CNPq, by the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office.

[1]A. S. Bazanella and R. Rui are with the Graduate Program in Electrical Engineering, Universidade Federal do Rio Grande do Sul, Brazil bazanela@ece.ufrgs.br and rui@ece.ufrgs.br

[2]M. Gevers is with ICTEAM, Louvain University, B1348 Louvain la Neuve, Belgium Michel.Gevers@uclouvain.be

[3]D. Coutinho is with the Department of Automation and Systems, Universidade Federal Santa Catarina, Florianópolis, SC, Brazil coutinho@das.ufsc.br

bility of parametric model structures; it presents the classical algorithm of [4] and [10] and illustrates it with two examples. In Section IV we present a range of results for the analysis of identifiability and informativity of the considered model class, based on our alternative approach. In Section V the classical approach and our new approach are applied to an example, which illustrates the similarities and differences. As always, we conclude.

II. Model class and assumptions

Consider the following class of deterministic continuous-time nonlinear model structures in generalized controller canonical form:

$$
\begin{aligned}
y &= x_1 && (1)\\
\dot{x}_1 &= x_2 && (2)\\
\dot{x}_2 &= x_3 && (3)\\
&\vdots\\
\dot{x}_n &= \frac{1}{n(\mathbf{x})}[\phi^T(\mathbf{x})\theta + m(\mathbf{x}) + g(\mathbf{x},u)] && (4)\\
g(\mathbf{x},u) &= \sum_{i=1}^{l} g_i(\mathbf{x})u^i && (5)
\end{aligned}
$$

where $\mathbf{x} = [x_1 \quad \dots \quad x_n]^T$ is the state vector, u is the scalar input; $\theta \in \Re^d$ is a vector of real parameters; $\phi(\mathbf{x})$ is a d-dimensional vector of multivariate polynomial functions, $m(\mathbf{x})$, $n(\mathbf{x})$ and $g_i(\mathbf{x})$ are scalar multivariate polynomials, with the property that $n(\mathbf{x}) > 0$ for all $\mathbf{x}$. The model class (1)-(4), often called Brunovsky canonical form, has been widely studied in nonlinear control: see e.g. [5]. Observe that in this structure, the state vector can be written as $\mathbf{x} = [y, \dot{y}, \dots, y^{(n-1)}]$ where $y^{(i)}$ denotes the i-th derivative of y.

The degree of the polynomial of highest degree in $\phi(\mathbf{x})$ is called the degree of $\phi(\mathbf{x})$ and denoted q. Moreover, the term $m(\mathbf{x})$ is such that $\nexists \eta : m(\mathbf{x}) \equiv \eta^T\phi(\mathbf{x})$. This last assumption does not represent a loss of generality in the model; if $m(\mathbf{x}) = \eta_1^T\phi(\mathbf{x})$ for some η_1, then the two terms could be grouped together with a reparametrization. For each parameter value θ the model class represents a model, whose trajectory and output are denoted $\mathbf{x}(t,\theta,u(t),\mathbf{x}(0))$ and $y(t,\theta,u(t),\mathbf{x}(0))$.

In the same spirit as [3] and [10], we shall make the following assumption on the input signal $u(.)$.

Assumption 1: The signal $u(t)$ is analytic and is such that the solution $\mathbf{x}(t,\theta,u(t),\mathbf{x}(0))$ of (2)-(4) is analytic in t. The virtue of this assumption, which is standardly used in nonlinear system identification, is that knowing all derivatives of an analytic signal at some time is equivalent to knowing that signal at all times.

Ideally the choice of parametrization made in (1)-(4) should be such that the model class can describe exactly the true system $\mathcal{S}$; we shall throughout make the following assumption.

Assumption 2: There exists a parameter value θ_0 such that the true system is described by (1)-(4) with $\theta = \theta_0$.

III. Background on identifiability and informativity

We briefly review the definitions from [6]. These are the nonlinear deterministic counterpart of the classical definitions as can be found in [9], which are for linear time-invariant systems in a stochastic framework with quasi-stationary processes.

Consider two different models obtained from the model class (1)-(4) at two different parameter values θ^* and θ, both with the same initial condition $\mathbf{x}(0) = \mathbf{x}_0$ and the same input $u(t)$.

Definition 1: (Identifiability at θ^*) The model class (1)-(4) is *globally identifiable at* θ^* if there exists an experiment $z(.) \stackrel{\Delta}{=} \{u(.), \mathbf{x}_0\}$ such that, for all $\theta \in \Re^d$, the outputs $y(t,\theta^*,u(t),\mathbf{x}(0)) = y(t,\theta,u(t),\mathbf{x}(0)) \ \forall t \geq 0$ **only if** $\theta = \theta^*$. ■

This definition relies on the possible existence of an appropriate experiment $z(.)$ which allows to differentiate between different values of θ by measuring the output. Such an experiment, when it exists, is called informative. The following is the deterministic counterpart of the concept of informativity for stationary stochastic LTI systems introduced in [1].

Definition 2: (Informativity at θ^*) An experiment $z(.) \stackrel{\Delta}{=} \{u(.), \mathbf{x}_0\}$ is globally informative at θ^* for the model class (1)-(4) if for all $\theta \in \Re^d$, the outputs $y(t,\theta,u(t),\mathbf{x}(0)) = y(t,\theta^*,u(t),\mathbf{x}(0)) \ \forall t \geq 0$ **only if** $\theta = \theta^*$. ■

Definitions 1 and 2 of identifiability and informativity consider that the experiment may consist of the choice of an initial condition and the choice of an input. There are situations where there is no input to the system and where the estimation of the parameters must be secured from a transient response to a properly chosen initial condition; this situation can be handled by Definition 1 with $u \equiv 0$. It is also often the case that the initial condition is not available to the designer of the experiment and that it is in fact unknown to him/her. In this situation identifiability must be secured from the input only, assuming that the initial condition is unknown, and might actually kill the transfer of excitation from the input to the state. To handle this case, we introduce the following definition.

Definition 3: (Identifiability at θ^* from u only) The model class (1)-(4) is *globally identifiable at θ^* from the input* if there exists an input $u(.)$ such that, for all initial conditions $\mathbf{x}_0$ and for all $\theta \in \Re^d$, the outputs $y(t,\theta^*,u(t),\mathbf{x}(0)) = y(t,\theta,u(t),\mathbf{x}(0)) \ \forall t \geq 0$ **only if** $\theta = \theta^*$. ■

Otherwise said, the model class is identifiable from the input only if there exists an input $u(t)$ which yields an informative experiment for all possible initial conditions. These definitions exhibit the two ingredients that are necessary for a

meaningful identification: informativity, which is a property of the applied experiment (input signal, initial condition, or a combination of these), and identifiability, which refers to the possible existence of an informative experiment given a particular model structure. These two ingredients depend only on the model structure. We should add, however, that under our Assumption 2, the purpose of the identification is to identify the true system. This requires that one is able to produce an informative experiment at θ_0; in that sense, an experiment that is informative for the identification of the true system depends on that system.

Classical results on identifiability of deterministic systems for broad classes of nonlinear systems are given in [4] and in [10], where an algorithmic solution (Ritt's algorithm) is also given. The main result of [10] can be stated as follows. Ritt's algorithm is applied to the model class in such a way as to obtain, if possible, an equivalent (equivalence meaning that both sets of equations have the same set of solutions) set of equations in which the dependence on θ is algebraic and linear. The model structure is globally identifiable at any θ if the following two conditions hold simultaneously for the equations obtained at the end of the algorithm:
1) the equations involving θ all take the following form

$$p_i(y, \dot{y}, \ldots, u, \dot{u}, \ldots)\theta_i = q_i(y, \dot{y}, \ldots, u, \dot{u}, \ldots) \qquad (6)$$

where $p_i(y, \dot{y}, \ldots, u, \dot{u}, \ldots)$ and $q_i(y, \dot{y}, \ldots, u, \dot{u}, \ldots)$ are scalar polynomials in y, u and their derivatives;
2) the polynomials $p_i(y, \dot{y}, \ldots, u, \dot{u}, \ldots)$, called "initials", and the multipliers $S_i(y, \dot{y}, \ldots, u, \dot{u}, \ldots)$ used at each step of the algorithm, called "separants", are all nonzero.
Note that these are sufficient conditions only.

The second condition depends on the experimental conditions; establishing identifiability thus requires verifying the existence of experimental conditions such that the initials and separants are nonzero. Furthermore, the informativity issue, that is the issue of constructing such experimental conditions that make the initials and separants nonzero, also requires separate consideration. In a number of examples (see [10]) the resulting expressions are simple enough so as to allow such verification of identifiability, and sometimes also solving the informativity issue, by inspection. In more complex cases this is usually not the case. Moreover, Ritt's algorithm itself is quite costly computationally, requiring more resources than a typical personal computer has to offer even for modestly complex models. The following example illustrates the merits and difficulties of applying Ritt's algorithm for the verification of the global identifiability of a model structure.

Example 3.1:

$$\dot{x}_1 = x_2, \quad y = x_1 \qquad (7)$$
$$\dot{x}_2 = x_3 \qquad (8)$$
$$\dot{x}_3 = \frac{\theta_1 x_1^3 + \theta_2 x_1 x_2 + \theta_3 x_2 x_3^2}{1 + x_2^2 + x_3^2} + u \qquad (9)$$

Application of Ritt's algorithm to this model yields

$$q_1(y, \dot{y}, \ddot{y}, y^{(3)}, u) = -p(y, \dot{y}, \ddot{y}, y^{(3)}, u) y \theta_1 \qquad (10)$$
$$q_2(y, \dot{y}, \ddot{y}, y^{(3)}, u) = p(y, \dot{y}, \ddot{y}, y^{(3)}, u)\theta_2 \qquad (11)$$
$$q_3(y, \dot{y}, \ddot{y}, y^{(3)}, u) = p(y, \dot{y}, \ddot{y}, y^{(3)}, u)\theta_3 \qquad (12)$$

where $p(y, \dot{y}, \ddot{y}, y^{(3)}, u)$ is given by (13) and the only separant that does not contain $p(y, \dot{y}, \ddot{y}, y^{(3)}, u)$ as a factor is given by (14) below.

$$\begin{aligned} p(y, \dot{y}, \ddot{y}, y^{(3)}, u) = {} & x_1^{(3)} \dot{x}_1^2 x_1^2 x_3 - x_1^{(3)} \dot{x}_1 x_1^3 \dot{x}_3 - 3\ddot{x}_1^2 \dot{x}_1 x_1^2 x_3 \\ & + 2\ddot{x}_1^2 x_1^3 \dot{x}_3 + 6\ddot{x}_1 \dot{x}_1^3 x_1 x_3 - 4\ddot{x}_1 \dot{x}_1^2 x_1^2 \dot{x}_3 + \ddot{x}_1 \dot{x}_1 x_1^3 \ddot{x}_3 \\ & - 6\dot{x}_1^5 x_3 + 6\dot{x}_1^4 x_1 \dot{x}_3 - 2\dot{x}_1^3 x_1^2 \ddot{x}_3 \end{aligned} \qquad (13)$$

$$S(y, \dot{y}, \ddot{y}, y^{(3)}, u) = x_1^2 \ddot{x}_1 - 2x_1 \dot{x}_1^2 \qquad (14)$$

Observe that the polynomials p and S can all be expressed as functions of y, u and their derivatives by making successive substitutions using the model structure (7)-(9). The results of [10] tell us that the model structure of Example 3.1 is globally identifiable at all θ if there exists an experiment (i.e. an input signal) such that the initials p, py and the separant S are all nonzero. Recall that this is a sufficient and not necessary condition. It appears difficult to establish the existence of such an experiment. In addition, assuming that such experiment does exist, it is not clear how to find or characterize one.

IV. IDENTIFIABILITY

In this section we provide conditions on the structure of the model class (1)-(4) that guarantee identifiability for such model class. We start by presenting an equivalent and more workable characterization of global identifiability at θ^*.

Since the parameter vector appears only in the last equation, identifiability and informativity rely entirely on the following equation.

$$n(\mathbf{x})\dot{x}_n - m(\mathbf{x}) - g(\mathbf{x}, u) = \phi^T(\mathbf{x})\theta \qquad (15)$$

It follows from the model equations that $\dot{x}_n = y^{(n)}$. The left hand side of this equation is a measured quantity, and so

is the regressor vector $\phi(\mathbf{x})$. Therefore identifiability rests entirely on the properties of the vector $\phi(\mathbf{x})$. The following Lemma provides an alternative characterization of global identifiability.

Lemma 4.1: The model structure (1)-(4) is globally identifiable at some θ^* if and only if there exists an experiment $z(.) \stackrel{\Delta}{=} \{u(t)\ t \geq 0, \mathbf{x}_0\}$ such that

$$\beta^T \phi(\mathbf{x}(t, \theta^*, u(t), \mathbf{x}_0)) = 0 \quad \forall t \geq 0 \Longrightarrow \beta = 0 \qquad (16)$$

Proof: Consider the solution $\mathbf{x}(t, \theta^*)$ of the model (1)-(4) at θ^*, and the solution $\mathbf{x}(t, \theta)$ of the same model at some other θ, both solutions driven by the same input signal and with the same initial condition.Suppose that for all input signals $\{u(t), t \geq 0\}$ and for all initial conditions $\mathbf{x}_0$ the solutions $y(t, \theta)$ and $y(t, \theta^*)$ are identical; it follows that $\mathbf{x}(t, \theta^*)$ and $\mathbf{x}(t, \theta)$ are also identical. The left hand sides of (15) are then also identical. Therefore $\phi(\mathbf{x}(t, \theta)) = \phi(\mathbf{x}(t, \theta^*)) \quad \forall t \geq 0$ and

$$(\theta^* - \theta)^T \phi(\mathbf{x}(t, \theta^*)) \equiv 0 \quad \forall t \geq 0$$

It then follows that $\theta = \theta^*$ if and only if (16) holds. ■

For convenience and for clarity, let us define

$$R_\infty(\mathbf{x}) \stackrel{\Delta}{=} [\phi(\mathbf{x})\ \dot{\phi}(\mathbf{x})\ \ddot{\phi}(\mathbf{x}) \ldots] \qquad (17)$$

and let $\varrho(\cdot)$ represent the rank of a matrix. We can now state a nonscalar version of Theorem 4.1 in [7].

Theorem 4.1: The model structure (1)-(4) is globally identifiable at some θ^* if and only if there exists an experiment $z(\cdot) \stackrel{\Delta}{=} \{u(t), t \geq 0; \mathbf{x}_0\}$ such that $\varrho(R_\infty(\mathbf{x}_0)) = d$.
Proof: It follows from Lemma 4.1 that the model structure (1)-(4) is **not** identifiable at θ^* if and only if for all initial conditions $\mathbf{x}_0$ and for all inputs $u(t), t \geq 0$,

$$\exists \beta \neq \mathbf{0}: \quad \beta^T \phi(\mathbf{x}(\theta^\star, t)) \equiv 0 \quad \forall t \geq 0 \qquad (18)$$

Because the solutions of (15) are analytic by Assumption 1, (18) is equivalent to the satisfaction of

$$\beta^T \phi(\mathbf{x}) = \beta^T \dot{\phi}(\mathbf{x}) = \ldots = \beta^T \phi^{(k)}(\mathbf{x}) = 0 \quad \forall k \geq 0 \quad (19)$$

for some $\beta \neq 0$ at any particular value of $\mathbf{x}$, in particular at $\mathbf{x}(0) = \mathbf{x}_0$. ■

Theorem 4.1 provides a strong result to characterize the identifiability of the model structure (1)-(4), but it is hard to verify whether the condition is satisfied. A more tractable (but only sufficient) condition is to construct a finite matrix having the first k-columns of $R_\infty(\mathbf{x})$ with $k \geq q$ and to show that it has full row rank; thus we define the $d \times (k+1)$ matrix

$$R_k(\mathbf{x}) \stackrel{\Delta}{=} [\phi(\mathbf{x})\ \dot{\phi}(\mathbf{x})\ \ldots\ \phi^{(k)}(\mathbf{x})] \qquad (20)$$

We have the following obvious but very useful result.

Corollary 4.1: The model structure (1)-(4) is globally identifiable at some θ^* if there exists an experiment $z(\cdot) \stackrel{\Delta}{=} \{u(t), t \geq 0; \mathbf{x}_0\}$ such that $\varrho(R_k(\mathbf{x}_0)) = d$ for some k. ■

We shall illustrate this approach with an example in Section V. Alternatively, $R_\infty(\mathbf{x})$ can be decomposed into a product of two matrices, whereby the rank of $R_\infty(\mathbf{x})$ can be related to some structural properties of the model class.

Both approaches require the computation of the successive time derivatives of $\phi(\mathbf{x})$ for the model structure (2)-(4), which proceeds as follows:

$$\dot{\phi}(\mathbf{x}) = \sum_{k=1}^{n} \frac{\partial \phi(\mathbf{x})}{\partial x_k} \dot{x}_k \qquad (21)$$

$$\ddot{\phi}(\mathbf{x}) = \sum_{k=1}^{n} [\sum_{j=1}^{n} (\frac{\partial^2 \phi(\mathbf{x})}{\partial x_k x_j} \dot{x}_k \dot{x}_j) + \frac{\partial \phi(\mathbf{x})}{\partial x_k} \ddot{x}_k] \qquad (22)$$

$$\begin{aligned} \phi^{(3)}(\mathbf{x}) = \sum_{k=1}^{n} \{ \sum_{j=1}^{n} [\sum_{i=1}^{n} (\frac{\partial^3 \phi(\mathbf{x})}{\partial x_k \partial x_j \partial x_i} \dot{x}_k \dot{x}_j \dot{x}_i) \qquad (23) \\ + \quad 2 \frac{\partial^2 \phi(\mathbf{x})}{\partial x_k \partial x_j} \ddot{x}_k \dot{x}_j + \frac{\partial^2 \phi(\mathbf{x})}{\partial x_k \partial x_j} \dot{x}_k \ddot{x}_j] + \frac{\partial \phi(\mathbf{x})}{\partial x_k} x_k^{(3)} \} \end{aligned}$$

and so on. It is seen that each derivative of $\phi(\mathbf{x})$ with respect to time - say the m-th derivative - can be written as a linear combination of the derivatives, up to the m-th order, of $\phi(\mathbf{x})$ with respect to each one of the state variables $x_1, \ldots x_n$. Let us organize all the partial derivatives of $\phi(\mathbf{x})$ with respect to the state variables in a single matrix called $J_q^n(\mathbf{x})$. To this end, define $\partial_k \phi(\mathbf{x})$ as the row vector containing all the $k-th$ order partial derivatives, that is:

$$\begin{aligned} \partial \phi(\mathbf{x}) &= \left[\begin{array}{ccc} \frac{\partial \phi(\mathbf{x})}{\partial x_1} & \ldots & \frac{\partial \phi(\mathbf{x})}{\partial x_n} \end{array} \right] \\ \partial_2 \phi(\mathbf{x}) &= \left[\begin{array}{cccc} \frac{\partial^2 \phi(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 \phi(\mathbf{x})}{\partial x_1 \partial x_2} & \ldots & \frac{\partial^2 \phi(\mathbf{x})}{\partial x_n^2} \end{array} \right] \\ \partial_3 \phi(\mathbf{x}) &= \left[\begin{array}{cccc} \frac{\partial^3 \phi(\mathbf{x})}{\partial x_1^3} & \frac{\partial^3 \phi(\mathbf{x})}{\partial x_1^2 \partial x_2} & \frac{\partial^3 \phi(\mathbf{x})}{\partial x_1^2 \partial x_3} & \ldots \\ \frac{\partial^3 \phi(\mathbf{x})}{\partial x_{n-1} \partial x_n^2} & \frac{\partial^3 \phi(\mathbf{x})}{\partial x_n^3} & & \end{array} \right], \end{aligned}$$

up to $\partial_q \phi(\mathbf{x})$, noting that $\partial_k \phi(\mathbf{x}) = \mathbf{0}\ \forall k > q$, since q is the degree of $\phi(\mathbf{x})$. Then define the *identifiability matrix*:

$$J_q^n(\mathbf{x}) \stackrel{\Delta}{=} \left[\begin{array}{ccccc} \phi(\mathbf{x}) & \partial \phi(\mathbf{x}) & \partial_2 \phi(\mathbf{x}) & \ldots & \partial_q \phi(\mathbf{x}) \end{array} \right]. \qquad (24)$$

The matrix $J_q^n(\mathbf{x})$ depends only on $\mathbf{x}$ and its dimensions are $d \times m$, where

$$m = 1 + \sum_{i=1}^{q} CR_i^n \qquad (25)$$

and CR_i^n denotes the i-combination of n elements with repetition, that is:

$$CR_i^n = \frac{(n+i-1)!}{i!(n-1)!} \qquad (26)$$

Then the matrix $R_\infty(\mathbf{x})$ can be written as

$$R_\infty(\mathbf{x}) = J_q^n(\mathbf{x}) W_\infty(\mathbf{x}) \qquad (27)$$

where $W_\infty(\mathbf{x})$ is a matrix in which each entry is a combination of time derivatives of the state variables. A remarkable feature of the matrix $W_\infty(\mathbf{x})$ is that it can be written on the basis of the knowledge of the dimension n of the state

and the degree q of $\phi(\mathbf{x})$ only. Thus, all model structures that have the same n and q have the same expression for $W_\infty(\mathbf{x})$. An example in Section V will illustrate this feature.

An obvious necessary condition for identifiability arises from the decomposition (27).

Lemma 4.2: The model structure (1)-(4) is identifiable only if $\varrho(J_q^n(\mathbf{x})) = d$.

Proof: Sylvester's inequality applied to (27) yields

$$\begin{aligned}&\varrho(J_q^n(\mathbf{x})) + \varrho(W_\infty(\mathbf{x})) - m \leq \varrho(R_\infty(\mathbf{x}))\\ &\leq \min\{\varrho(J_q^n(\mathbf{x})), \varrho(W_\infty(\mathbf{x}))\} \end{aligned} \tag{28}$$

The result follows immediately from the second inequality. ■

It is worth mentioning that $\varrho(J_q^n(\mathbf{x})) = d$ requires $m \geq d$. Now, from the first inequality in (28) it follows that, provided that the necessary condition $\varrho(J_q^n(\mathbf{x})) = d$ is satisfied, $\varrho(W_\infty(\mathbf{x})) = m$ would suffice for identifiability.

The following theorem provides a sufficient condition for global identifiability of the model structure studied in this paper.

Theorem 4.2: The model structure (1)-(4) is globally identifiable at any θ if, for some experiment $z(.)$, $\varrho(J_q^n(\mathbf{x})) = d$ and $\varrho(W_\infty(\mathbf{x})) = m$, where m is a function of n and q, defined by (25)-(26).

Proof: The result follows from the left inequality in (28). ■

The following is an immediate consequence.

Corollary 4.2: The model structure (1)-(4) is globally identifiable at any θ if, for some experiment $z(.)$ and for some finite k, $\varrho(J_q^n(\mathbf{x})) = d$ and $\varrho(W_k(\mathbf{x})) = m$, where $W_k(\mathbf{x}))$ is defined from $R_k(\mathbf{x}) = J_q^n(\mathbf{x})W_k(\mathbf{x})$. ■

V. Examples

Example 5.1: Consider the following example:

$$\dot{x}_1 = x_2 \tag{29}$$

$$\dot{x}_2 = \frac{\theta_1 x_1^2 + \theta_2 x_1 x_2}{1 + x_2^2} + 1 + x_1^2 + x_1 x_2 u \tag{30}$$

that is,

$$\phi(\mathbf{x}) = \begin{bmatrix} x_1^2 \\ x_1 x_2 \end{bmatrix} \quad n(\mathbf{x}) = 1 + x_2^2$$

$$g(\mathbf{x}, u) = (1 + x_2^2) x_1 x_2 u \quad m(\mathbf{x}) = (1 + x_1^2)(1 + x_2^2)$$

A necessary condition for identifiability is $\varrho(J_q^n) = 2$ with

$$\begin{aligned} J_q^n(\mathbf{x}) &= \begin{bmatrix} \phi(\mathbf{x}) & \frac{\partial \phi(\mathbf{x})}{\partial x_1} & \frac{\partial \phi(\mathbf{x})}{\partial x_2} & \frac{\partial^2 \phi(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 \phi(\mathbf{x})}{\partial x_1 \partial x_2} & \frac{\partial^2 \phi(\mathbf{x})}{\partial x_2^2} \end{bmatrix} \\ &= \begin{bmatrix} x_1^2 & 2x_1 & 0 & 2 & 0 & 0 \\ x_1 x_2 & x_2 & x_1 & 0 & 1 & 0 \end{bmatrix} \end{aligned} \tag{31}$$

Clearly, $\varrho(J_q^n) = 2$, and the necessary condition is satisfied. Then a sufficient condition is the existence of an experiment such that the corresponding $W_\infty(\mathbf{x})$ matrix, whose first six columns are given below, has rank 6 (see Corollary 4.1).

$$W_6(\mathbf{x}) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \dot{x}_1 & \ddot{x}_1 & \overset{(3)}{x}_1 & \overset{(4)}{x}_1 & \overset{(5)}{x}_1 \\ 0 & \dot{x}_2 & \ddot{x}_2 & \overset{(3)}{x}_2 & \overset{(4)}{x}_2 & \overset{(5)}{x}_2 \\ 0 & 0 & \dot{x}_1^2 & 3\ddot{x}_1\dot{x}_1 & 4\overset{(3)}{x}_1\dot{x}_1 + 3\ddot{x}_1^2 & 5\overset{(4)}{x}_1\dot{x}_1 + 10\overset{(3)}{x}_1\ddot{x}_1 \\ 0 & 0 & 2\dot{x}_1\dot{x}_2 & 3\ddot{x}_1\dot{x}_2 + 3\dot{x}_1\ddot{x}_2 & 4\overset{(3)}{x}_1\dot{x}_2 + 6\ddot{x}_1\ddot{x}_2 + 4\dot{x}_1\overset{(3)}{x}_2 & 5\overset{(4)}{x}_1\dot{x}_2 + 10\overset{(3)}{x}_1\ddot{x}_2 + 10\ddot{x}_1\overset{(3)}{x}_2 + 5\dot{x}_1\overset{(4)}{x}_2 \\ 0 & 0 & \dot{x}_2^2 & 3\ddot{x}_2\dot{x}_2 & 4\overset{(3)}{x}_2\dot{x}_2 + 3\ddot{x}_2^2 & 5\overset{(4)}{x}_2\dot{x}_2 + 10\overset{(3)}{x}_2\ddot{x}_2 \end{bmatrix} \tag{32}$$

It is important to realize, as stated above, that all models with the same n and q have the same $W_\infty(\mathbf{x})$ matrix; they differ only in the matrix $J_q^n(\mathbf{x})$.

According to Theorem 4.1, the model is identifiable if and only if there exists an experiment such that the matrix $R_\infty(\mathbf{x})$ has rank equal to 2. From Corollary 4.1 the model is identifiable if $R_2(\mathbf{x})$ has rank equal to 2. This last matrix is given by

$$R_2(\mathbf{x}) \triangleq \begin{bmatrix} x_1^2 & 2x_1\dot{x}_1 \\ x_1 x_2 & x_1\dot{x}_2 + \dot{x}_1 x_2 \end{bmatrix} \tag{33}$$

and its rank is equal to two if and only if its determinant is not zero, that is:

$$\mid R_2(\mathbf{x}) \mid = x_1^3\ddot{x}_1 - x_1^2\dot{x}_1^2 \neq 0 \tag{34}$$

Hence, existence of an experiment for which (34) is satisfied constitutes a sufficient condition for identifiability. Such an experiment, if it exists, is then called informative. We can now consider different experimental scenarios.

(a) Consider first that the initial condition $\mathbf{x}_1(0)$ and the input $u(t)$ are available to the designer. Taking $x_1(0) \neq 0$ reduces condition (34) to $\dot{x}_2(0) \neq \frac{x_2^2(0)}{x_1(0)}$. Replacing $\dot{x}_2(0)$ by its expression given by (30) shows that for any choice of $x_2(0) \neq 0$ one can choose $u(t)$ such that this inequality hods. Other informative experiments exist as well.

(b) Now consider that $u(t) \equiv 0$ and that informativity must come from the initial condition. For any $x_1(0) \neq 0$, (34) yields

$$\frac{\theta_1 x_1(0)^2 + \theta_2 x_1(0) x_2(0)}{1 + x_2(0)^2} + 1 + x_1(0)^2 \neq \frac{x_2^2(0)}{x_1(0)}. \quad (35)$$

Since (35) is a fourth order polynomial equation in $x_2(0)$, there are at most four real roots $x_2(0)$ of this polynomial; all other values of $x_2(0)$ result in an informative experiment; the model is identifiable. Clearly, other informative experiments exist as well.

(c) Let us now analyze identifiability from the input. The model is identifiable from the input if there exists an input $u(t)$ such that (34) is satisfied for all initial conditions. The initial condition $x_2(0) = 0;\ x_1(0) = \sqrt{\frac{-1}{1+\theta_1}}$ will result in $\mathbf{x}(t) \equiv 0$ whatever $u(t)$ or θ. Hence the model is **not** identifiable from the input, at least at those θ such that $\theta_1 < -1$, because initial conditions with $x_2(0) = 0$ may "kill" the excitation coming from whatever $u(t)$.

Now, application of Ritt's algorithm to this example results in the following initials and separants, which must be different from zero:

$$p_1(x_1, \dot{x}_1, \ddot{x}_1) = -(x_1^3 \ddot{x}_1 - x_1^2 \dot{x}_1^2) \neq 0 \quad (36)$$

$$p_2(x_1, \dot{x}_1, \ddot{x}_1) = -(x_1^2 \ddot{x}_1 - x_1 \dot{x}_1^2) \neq 0 \quad (37)$$

$$S_1(x_1, \dot{x}_1) = (x_1 \dot{x}_1)^2 \neq 0 \quad (38)$$

Hence the simultaneous satisfaction of conditions (36), (37) and (38) is a sufficient condition for identifiability. Condition (36) is the same as obtained by our method, while condition (37) is equivalent to it. As for condition (38), it is an additional constraint originated from the fact that the concept of identifiability treated in [10] is that of Definition 3, i.e. identifiability from the input. Indeed, (38) excludes the initial conditions with $x_2(0) = 0$, which have been shown above to "kill" the excitation coming from the input. Thus condition (38) says that there exist inputs $u(t)$ that provide informative experiments for any initial condition except those violating (38). In the sense of Definition 3, this means that the system is not identifiable from the input only.

This example illustrates one advantage of the method developed in this paper, besides the speed of computation. Our method leads to a set of sufficient conditions that are typically a subset of those resulting from the Ritt polynomials being all nonzero. In addition, it allows us to make search for informative experiments (and hence identifiability) under different experimental scenarios.

Return to Example 3.1

The computation of the 3×3 matrix $R_3(\mathbf{x})$ for Example 3.1 takes 1 second on a PC as opposed to more than 5 minutes for the Ritt algorithm leading to the polynomials (13) and (14). The determinant of $R_3(\mathbf{x})$ is given by $p(y, \dot{y}, \ddot{y}, y^{(3)}, u)y$, where $p(y, \dot{y}, \ddot{y}, y^{(3)}, u)$ is given by (13). This leads us to confirm two statements made earlier:

(1) the computation time is significantly lower for our method than for the computation of the Ritt polynomials;

(2) the sufficient condition obtained by our method is a subset of the conditions obtained by running the algorithm developed by [10]; see the discussion on Example 5.1.

VI. Conclusions

Several results on the identifiability of rational models represented in a canonical form have been presented in this paper, extending the results of [6] to a broader class of models. A necessary and sufficient condition has been given. Then two more easily computable conditions, one necessary and one sufficient, have been derived from it. The relationships among these conditions and the classical ones arising from differentiable algebra (as in [4], [10]) have been illustrated by means of an example. As could be expected from the fact that our results are valid for a specific class of models, whereas those classical results apply for general rational models, our methods give tighter and more easily computable conditions for the considered class. Future work will focus on further exploring computational aspects, particularly in constructing informative experiments through the analysis of the rank of the matrix $W_\infty(\mathbf{x})$.

References

[1] A.S. Bazanella, X. Bombois, and M. Gevers. Necessary and sufficient conditions for uniqueness of the minimum in prediction error identification. *Automatica*, 48(8):1621–1630, August 2012.

[2] Sara Berthoumieux, Daniel Kahn, Hidde De Jong, and Eugenio Cinquemani. Structural and practical identifiability of approximate metabolic network models. In *Proceedings of the 16th IFAC symposium on System Identification*, Brussels, Belgique, 2012.

[3] S. Dasgupta, M. Gevers, G. Bastin, G. Campion, and Libei Chen. Identifiability of scalar linearly parametrized polynomial systems. In *Proc. 9th IFAC Symposium on Identification and System Parameter Estimation (SYSID 91)*, volume 1, pages 374–378, Budapest, Hungary, July 1991.

[4] S. Diop and M. Fliess. Nonlinear observability, identifiability and persistent trajectories. In *Proc. 30th IEEE Conf. on Decision and Control*, pages 714–719, Brighton, UK, December 1991.

[5] M. Fliess. Generalized controller canonical form for linear and nonlinear dynamics. *IEEE Transactions on Automatic Control*, 35(9):994–1001, September 1990.

[6] M. Gevers, A. Bazanella, D. Coutinho, and S. Dasgupta. Identifiability and excitation of polynomial systems. In *CD-ROM of 52nd IEEE Conf. on Decision and Control*, pages 4278–4283, Florence, Italy, December 2013.

[7] M. Gevers, A. Bazanella, D. Coutinho, and S. Dasgupta. Identifiability and excitation of linearly parametrized rational systems. *submitted for publication to IEEE Transactions on Automatic Control*, 2014.

[8] Johan Karlsson, Milena Anguelova, and Mats Jirstrand. An efficient method for structural identifiability analysis of large dynamic systems. In Michel Kinnaert, editor, *16th IFAC Symposium on System Identification*, volume 16, pages 941–946, Brussels, 2012. Laxemburg: IFAC.

[9] L. Ljung. *System Identification: Theory for the User, 2nd Edition*. Prentice-Hall, Englewood Cliffs, NJ, 1999.

[10] L. Ljung and T. Glad. On global identifiability for arbitrary model parametrizations. *Automatica*, 30(2):265–276, 1994.

[11] Christian Lyzell, Torkel Glad, Martin Enqvist, and Lennart Ljung. Difference algebra and system identification. *Automatica*, 47(9):1896 – 1904, 2011.

[12] Jana Nemcová. Structural identifiability of polynomial and rational systems. *Mathematical Biosciences*, 223(2):83 – 96, 2010.

[13] X. Xia and C.H. Moog. Identifiability of nonlinear systems with application to hiv/aids models. *Automatic Control, IEEE Transactions on*, 48(2):330 – 336, feb. 2003.