Identifiability and Excitation of Linearly Parametrized Rational Systems*

Michel Gevers[1], Alexandre S. Bazanella[2], Daniel F. Coutinho[3] and Soura Dasgupta[4]

***Abstract*— This paper establishes identifiability and informativity conditions for a class of deterministic linearly parametrized rational models. The class considered is rational in the states and polynomial in the inputs. The standard definitions of identifiability and informativity for linear systems are expanded to account for the situation, common for nonlinear deterministic systems, where the identification is achieved either through the application of informative inputs or via the response to informative initial conditions. We provide necessary and sufficient conditions for identifiability from the initial state, respectively from the input, as well as necessary and sufficient conditions on the initial state to produce an informative experiment. We also provide sufficient conditions on the input to be informative when the initial condition is unknown and could therefore potentially destroy the transfer of information from the input to the regressor.**

I. INTRODUCTION

The question of identifiability of parametrized dynamical systems has occupied generations of system theorists, and the very definition of this concept has evolved over the years. For a long time, this concept embraced both the parametrization issue and the richness of the data set. Eventually, a clear separation was made between the *identifiability of the model structure*, which is a parametrization issue, and the *informativity of the data*, which is the issue of applying an excitation to the system that will produce different responses for different parameter values.

The question of identifiability of the model structure can be succinctly summarized as follows: does there exist an experiment such that the data collected from this experiment allow one to uniquely determine the parameter values? From a more analytical point of view this question translates into the following: *is the mapping from parameter vector θ to model $M(\theta)$ injective*? The seminal paper [13] 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 observe that identifiability is a property of the chosen model structure (i.e. the parametrization); it is totally independent of the true system and of the data.

The question of informativity of the data turns out to be harder to solve. For linear time invariant (LTI) systems, sufficient conditions on the input had been available for a very long time (see e.g. [12]), but necessary and sufficient conditions remained elusive until this question was solved in [9]. The relationship between identifiability, informativity and the uniqueness of the minimum in a Prediction Error Identification framework was established in [2] and required the introduction of the new concept of *local informativity*.

Informativity of the data - also known as input richness or transfer of excitation - is a topic that has attracted and continues to attract a wide attention: it is an experiment design issue. It is important not just in the context of identification, but also for the convergence of adaptive estimation and control schemes. Simply stated, the question can be summarized as follows: how to choose an experiment such that the data collected from this experiment suffices to uniquely determine the parameter values, once it has been determined that such an experiment exists (that is, once it has been determined that the model structure is identifiable)? Again, on a more analytical note, this question translates into: *What are the conditions on the excitation signals that will make the Gramian associated with a certain regression vector full rank?* When that Gramian is full rank, the regressor vector is said to be informative. The full rank condition on the Gramian is required for the estimation of the parameters. The excitation signal is typically an input signal, but in this paper we examine the situation where it can also be a properly chosen initial condition or a combination of both.

Besides the results for the identification of LTI systems mentioned above, a wide range of questions related to informativity of the data and transfer of excitation have been addressed, dealing with different classes of systems, different types of input, and different convergence requirements.

In [1] informativity conditions on the input have been obtained for parameter convergence in linear discrete-time adaptive control schemes. Similar conditions for continuous-time linear adaptive control systems have been derived in [10]. In [11] nonlinear adaptive control schemes were studied, and it was shown that nonlinearities may reduce the requirements on the richness of the external signal.

In the context of system identification, informativity conditions have been obtained for linear time-varying systems in

*This work is supported by CNPq, 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, by US NSF grants CCF-0830747, EPS-1101284, and CNS-1329657, and by a grant from the Roy J. Carver Charitable Trust.

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

[2]A. S. Bazanella is with the Department of Electrical Engineering, Universidade Federal do Rio Grande do Sul, Brazil bazanela@ece.ufrgs.br

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

[4] S. Dasgupta is with the Dept of Electrical and Computer Engineering, The University of Iowa, Iowa 52242, USA dasgupta@engineering.uiowa.edu

[14], and more specific results have been obtained in [4] for Linear Parameter Varying systems with an ARX structure. In [6] informativity conditions on the input signal have been derived for a class of discrete-time linearly parametrized systems that are linear in the output and polynomial in the input. Bilinear systems are special members of this class. In [16] the question of which type of input signals (e.g. pulses, impulses, etc) are sufficient for the identification of bilinear systems has been studied.

In [5] necessary and sufficient conditions were derived for the identifiability of continuous time scalar linearly parametrized deterministic polynomial systems. The harder question of how to generate informative inputs for such systems was raised in that paper but not solved. In addition, identification based on the response to an initial condition was not considered in [5]. A first step towards addressing these issues was presented in a recent paper [8], where results on identifiability and informativity for scalar linearly parametrized systems that are polynomial (rather than rational) in the state were presented.

In this paper we expand the results of [5] and [8] in several directions. First we expand the considered model class: we consider a class of continuous time linearly parametrized deterministic models that are rational in the state and polynomial in the input. Secondly, for such models we present necessary and sufficient conditions for identifiability from the input and the initial state; these conditions are on the structure of the model class only. Thirdly, we expand the traditional view by also presenting necessary and sufficient conditions for identifiability from the initial state when there is no input, as well as necessary and sufficient conditions on the initial state to generate an informative experiment in such situation. Indeed, it is common in some application fields that the parameters are identified from data obtained as the response to some initial condition without any external input. Finally, we present necessary and sufficient conditions for identifiability from the input for this rational model class, as well as sufficient conditions on the input to be informative, in the context where the initial state is unknown and may therefore potentially destroy the transfer of excitation from the input to the regressor.

The remainder of the paper is organized as follows. In Section II we present the nonlinear model class that is treated in this paper and the fundamental assumptions on the signals and the models. In Section III the definitions of identifiability and informativity are extended from the linear stationary stochastic case to the case of nonlinear deterministic systems. Fundamental results on identifiability for the considered class of model structures are given in Section IV. From that point on, we depart from the classical setting in which identifiability and informativity are secured from the input, to definitions that apply when these can also be secured from the response to an initial condition. The tools developed in Section IV are essential in the subsequent developments. In Section V we study the identifiability and informativity from the initial state with zero input, providing the formal definitions of these concepts and necessary and sufficient conditions for both. Then we approach, in Section VI, the problem of identifiability and informativity from the input, regardless of the initial state. Again the concepts are formalized and conditions for identifiability and informativity are presented. An example is given that illustrates the difficulty of obtaining necessary and sufficient conditions on the input to be informative regardless of the initial state. Section VII concludes.

II. MODEL CLASS AND ASSUMPTIONS

Consider the following class of deterministic continuous-time nonlinear model structures:

$$\begin{aligned} \dot{x} &= f(x,\theta) + g(x,u), \quad x(0) = x_0 \\ y &= h(x,\theta) \end{aligned} \tag{1}$$

where $x : \Re \to \Re$ is a scalar state, $\theta \in \Re^d$ is an unknown vector, $f(\cdot,\cdot)$, $g(\cdot,\cdot)$ and $h(\cdot,\cdot)$ are given (i.e. known) scalar analytical functions and $g(\cdot,0) = 0$. We shall specialize in this paper to a class of linearly parametrized models that are rational in the state and polynomial in the input, with the following form for $f(x,\theta)$, $g(x,u)$ and $h(x,\theta)$:

$$f(x,\theta) = \frac{1}{n(x)}[\theta^T \phi(x) + m(x)] \tag{2}$$

$$g(x,u) = \frac{1}{n(x)} \sum_{i=1}^{l} g_i(x) u^i = \frac{1}{n(x)} G(x) U \tag{3}$$

$$h(x,\theta) = x \tag{4}$$

where $n(x)$ is a known polynomial such that $n(x) > 0 \ \forall x \in \Re$, $\phi(x) \in \Re^d$ is a known polynomial vector in the scalar x, $m(x)$ is a known polynomial in x, $g_i(x) \in \Re$ are known polynomials in x, and where $G(x)$ and U are defined as follows:

$$G(x) = [g_1(x) \ \ g_2(x) \dots g_l(x)], \ U = [u \ \ u^2 \ \ \dots \ \ u^l]^T \tag{5}$$

We denote by q the polynomial degree of $\phi(x)$, i.e. the degree of the highest degree polynomial in $\phi(x)$. In order to avoid a degenerate situation, we make the following assumption concerning the model structure.

Assumption 1: At least one of the polynomials $\phi_1(x), \dots, \phi_d(x), m(x), g_1(x), \dots, g_l(x)$ is not the zero polynomial, and the term $m(x)$ is such that $\nexists \ \eta \neq \mathbf{0} : m(x) \equiv \eta^T \phi(x)$.

This last assumption does not represent a loss of generality in the model; if $m(x) = \eta_1^T \phi(x)$ for some η_1, then the model could be rewritten as $f(x,\rho) = \frac{1}{n(x)}[\rho^T \phi(x)]$ with the new parameter vector $\rho = \theta + \eta_1$. The family of all models (1)-(4) generated by all $\theta \in \Re^d$ is called the model class $\mathcal{M}$.

Concerning the signals, we shall make the following assumption which is standard in nonlinear system identification; see, e.g., [5] and [13].

Assumption 2: The signal $u(t)$ is analytic and the solution $x(t)$ of (1) is an analytic function of time for all $u(t)$ and $x(0)$.

The virtue of the above assumption 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 parameterization made in (1)-(4) should be such that the model class can describe exactly the true system; we shall throughout make the following assumption.

Assumption 3: There exists a parameter value θ_0 such that the true system is described by (1) with $\theta = \theta_0$.

The model structure (1)-(4) can be rewritten as

$$\dot{x} = \frac{\theta^T \phi(x) + m(x) + G(x)U}{n(x)}, \quad x(0) = x_0 \qquad (6)$$

or, alternatively, as

$$n(x)\dot{x} = \theta^T \phi(x) + m(x) + G(x)U, \quad x(0) = x_0 \qquad (7)$$

This latter expression will be most convenient for some of the derivations to follow.

III. THE CONCEPTS OF IDENTIFIABILITY AND INFORMATIVITY

We now present definitions that are the nonlinear deterministic counterpart of the classical definitions given in [12] for linear time-invariant systems in a stochastic framework of quasi-stationary processes. These definitions clearly separate the concepts of identifiability, which is a property of the model structure, and of informativity, which is a property of the experimental data. In addition, we shall depart from the Linear Time Invariant (LTI) literature on identifiability and informativity by considering that the information content in the data, that allows estimation of the unknown parameters, can come either from the external input signal $u(.)$ or from the response to an initial condition x_0. Indeed in many engineering applications of nonlinear systems (e.g. in batch chemical processes) the experiment that allows the estimation of the unknown parameters is obtained by measuring the response of the nonlinear system to some initial condition; in particular, it is often the case that there are no external inputs to the system. For the sake of generality and of possible future extensions, we present all definitions for the general model class (1), without any restriction on linearity in θ, rationality in x or polynomial structure in u.

Consider the model structure (1) at some value $\theta^\star$ with initial condition x_0:

$$\begin{aligned} \dot{x} &= f(x, \theta^\star) + g(x, u), \quad x(0) = x_0 \\ y &= h(x, \theta^\star) \end{aligned} \qquad (8)$$

and the same model structure at θ with initial condition $\hat{x}_0$:

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x}, \theta) + g(\hat{x}, u), \quad \hat{x}(0) = \hat{x}_0 \\ \hat{y} &= h(\hat{x}, \theta) \end{aligned} \qquad (9)$$

Definition 1: (Identifiability at $\theta^\star$) The model structure (1) is *globally identifiable at $\theta^\star$* if there exists an experiment $z(.) \triangleq \{u(.), x_0\}$ such that, for all $\theta \in \Re^d$, the outputs of the models (8) and (9), driven by the same $u(.)$ and with the same initial condition $x_0 = \hat{x}_0$ are identical (i.e. $\hat{y}(t, \theta) = y(t, \theta^\star) \;\; \forall t \geq 0$) **only if** $\theta = \theta^\star$.

This definition relies on the possible existence of an appropriate experiment $z(.)$ that differentiates between different values of θ by measuring the output. Such an experiment, when it exists, will be called informative. In [2], the new concept of *informativity at a particular θ^** was introduced for stationary stochastic LTI systems; the next definition is its deterministic counterpart.

Definition 2: (Informativity at $\theta^\star$) The experiment $z(.) \triangleq \{u(.), x_0\}$ is *globally informative at $\theta^\star$* for the model set (1) if, for all $\theta \in \Re^d$, the outputs of the models (8) and (9), driven by the same experimental data $z(.)$, are identical (i.e. $\hat{y}(t, \theta) = y(t, \theta^\star) \;\; \forall t \geq 0$) **only if** $\theta = \theta^\star$.

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 3, the purpose of 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.

IV. FUNDAMENTAL IDENTIFIABILITY ANALYSIS

Rewriting the model structure (7) in the form

$$n(x)\dot{x} - m(x) - G(x)U = \theta^T \phi(x), \quad x(0) = x_0 \qquad (10)$$

shows that identifiability rests entirely on the vector $\phi(x)$. Since the left hand side is a measured quantity, we observe that θ can be uniquely determined from measured data over a period $[0, t]$ if and only if $\phi(x)$ can be made to span the whole space $\Re^d$ over that same period. This amounts to finding an experiment $z(.) \triangleq \{u(\tau) \text{ for } \tau \in [0, t], x_0\}$ that makes $\phi(x)$ sufficiently rich over the interval $[0, t]$.

Definition 3: Let the trajectory $x(\tau, \theta)$ be defined as the solution of the differential equation (10) for some $u(\tau), \tau \geq 0$ and some x_0. The vector $\phi(x(\tau, \theta))$ is *sufficiently rich* over an interval $[t_0, t_0 + t]$ if there exists $\alpha > 0$ such that

$$\int_{t_0}^{t_0+t} \phi(x(\tau, \theta)) \phi^T(x(\tau, \theta)) d\tau > \alpha I \qquad (11)$$

Lemma 4.1: The model structure (10) is globally identifiable at some $\theta^\star$ if and only if there exists an experiment $z(.) \triangleq \{u(\tau) \text{ for } \tau \in [0, t], x_0\}$ such that $\phi(x(t, \theta^\star))$ is sufficiently rich over the interval $[0, t]$.

Proof: Consider the solution $x(t, \theta^\star)$ of the model (10) at $\theta^\star$, and the solution $x(t, \theta)$ 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 $\{u(\tau), \tau \in [0, t]\}$ and for all initial conditions x_0 the solutions $x(\tau, \theta^\star)$ and $x(\tau, \theta)$ of (10) are identical. The left hand sides of (10) are then also identical and $\phi(x(\tau,\theta^\star)) = \phi(x(\tau,\theta)) \ \ \forall\tau \in [0,t]$. It then follows that

$$(\theta^\star - \theta)^T\phi(x(\tau,\theta^\star)) \equiv 0 \ \ \forall\tau \geq 0$$

which is equivalent to

$$(\theta^\star-\theta)^T\phi(x(\tau,\theta^\star))\phi^T(x(\tau,\theta^\star))(\theta^\star-\theta) \equiv 0 \ \ \forall\tau \geq 0 \quad (12)$$

Integrating (12) over the interval $[0,t]$ yields

$$(\theta^\star - \theta)^T \int_0^t \phi(x(\tau,\theta^\star))\phi^T(x(\tau,\theta^\star))d\tau \ (\theta^\star - \theta) = 0$$

It follows that $\theta = \theta^\star$ if and only if $\phi(x(t,\theta^\star))$ is sufficiently rich over the interval $[0,t]$. ■

The sufficient richness of the regressor vector $\phi(x(\tau,\theta)) \in \Re^d$, as defined in Definition 3, hinges on two questions:
(i) does there exist a vector $X \triangleq [x_1, x_2, \ldots x_d] \in \Re^d$ such that the $d \times d$ matrix $L(X) \triangleq [\phi(x_1), \phi(x_2), \ldots, \phi(x_d)]$ is nonsingular?
(ii) does there exist an experiment that generates a $\phi(x(\tau,\theta))$ which spans the whole space $\Re^d$? This requires that the trajectory $x(\tau,\theta)$, which depends on θ, $u(.)$ and x_0, passes through d points $\{x_1, x_2, \ldots x_d\}$ that make $L(X)$ nonsingular.

The first is a structural issue on $\phi(x)$ only; it is independent of $\theta, u(.), x_0, m(x)$, $n(x)$ and $G(x)$. Existence of such an X is a necessary condition for identifiability [5]. To proceed in this analysis we use the fact that $\phi(x)$ is an analytic function of time, which leads to the following result.

Theorem 4.1: The model structure (10) is globally identifiable at some $\theta^\star$ if and only if there exists an experiment $z(.) \triangleq \{u(\tau) \text{ for } \tau \in [0,t], x_0\}$ such that

$$\beta^T\phi(x_0) = \beta^T\dot{\phi}(x_0) = \ldots = \beta^T\phi^{(k)}(x_0) = 0 \ \ \forall k \geq 0 \quad (13)$$

implies $\beta = 0$.

Proof: It follows from Lemma 4.1 and Definition 3 that the model structure (10) is **not** identifiable at $\theta^\star$ if and only if for all initial conditions x_0 and for all inputs $u(\tau), \tau \geq 0$,

$$\exists\beta \neq \mathbf{0}: \quad \beta^T\phi(x(\theta^\star,t)) \equiv 0 \ \ \forall t \geq 0 \quad (14)$$

Because the solutions of (10) are analytic by Assumption 2, (14) is equivalent to the satisfaction of

$$\beta^T\phi(x) = \beta^T\dot{\phi}(x) = \ldots = \beta^T\phi^{(k)}(x) = 0 \ \ \forall k \geq 0 \quad (15)$$

for some $\beta \neq 0$ at any particular value of x, in particular at $x(0) = x_0$. ■

The equivalence between (14) and (15) will form the basis for our derivations of identifiability and informativity in the following sections. To facilitate these derivations we now define the following $d \times (k+1)$ matrix

$$R_k(\theta,x) \triangleq [\phi(x) \ \ \dot{\phi}(x) \ \ldots \ \phi^{(k)}(x)], \quad k \geq 1 \quad (16)$$

In particular we shall consider

$$R_\infty(\theta,x) \triangleq [\phi(x) \ \ \dot{\phi}(x) \ \ddot{\phi}(x) \ldots] \quad (17)$$

Comment - It is important to realize that in the expressions (15)-(16)-(17) the derivatives of $\phi(x)$, evaluated at some time t_0 and hence at some state $x(t_0)$, depend also on θ and on the derivatives of x at t_0. Equivalently, by application of the model equations, one can think of the derivatives of $\phi(x)$ as depending on θ, $u(t_0)$ and the derivatives of $u(t)$ at t_0. In order not to be overwhelmed by notation, we shall only mention the θ-dependence of the matrices in these expressions.

With the notation (17), the statement (15) can be rewritten as

$$\beta^T R_\infty(\theta,x) = 0. \quad (18)$$

This leads to the following equivalent statement for Theorem 4.1.

Corollary 4.1: The model structure (10) is globally identifiable at some $\theta^\star$ if and only if

$$\varrho(R_\infty(\theta^\star,x_0)) = d \quad \text{for some} \quad z(\cdot) \quad (19)$$

where $\varrho(\cdot)$ represents the rank of a matrix. ■

Now observe that

$$\dot{\phi}(x) = \frac{\partial\phi(x)}{\partial x}\dot{x},$$
$$\ddot{\phi}(x) = \frac{\partial^2\phi(x)}{\partial x^2}\dot{x}^2 + \frac{\partial\phi(x)}{\partial x}\ddot{x},$$

and so on for all time derivatives of $\phi(x)$. We define the following $d \times (k+1)$ matrix:

$$J_k(x) \triangleq \begin{pmatrix} \phi(x) & \frac{\partial\phi(x)}{\partial x} & \frac{\partial^2\phi(x)}{\partial x^2} & \ldots & \frac{\partial^k\phi(x)}{\partial x^k} \end{pmatrix} \quad (20)$$

which contains only the derivatives of $\phi(x)$ with respect to x and thus depends only on the structure of the vector $\phi(x)$ itself (and not on θ and $u(t)$). It is important to note that all columns to the right of the $(q+1)$-th column of $J_k(x)$ are zero since $deg(\phi(x)) = q$.

Then straightforward calculations show that one can write

$$R_k(\theta,x) = J_k(x)W_k(\theta,\dot{x},\ddot{x},\ldots,x^{(k)}) \quad (21)$$

where W_k is a $(k+1) \times (k+1)$ upper-triangular matrix of the form

$$W_k(\theta,\dot{x},\ddot{x},\ldots,x^{(k)}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \dot{x} & \ddot{x} & \ldots & x^{(k)} \\ 0 & 0 & \dot{x}^2 & \ldots & \times \\ \vdots & \vdots & \vdots & \ddots & \times \\ 0 & 0 & 0 & \ldots & \dot{x}^k \end{pmatrix} \quad (22)$$

where all elements to the right of the diagonal contain derivatives of x. The factorization (22) will play a useful role in the sequel. It immediately leads to the following important *necessary condition for identifiability*.

Corollary 4.2: The model structure (10) is globally identifiable at some θ only if $\varrho(J_q) = d$. This requires in particular that $q \geq d-1$.

Proof: First note that, because of the polynomial nature of $\phi(x)$, the rank of $J_q(x)$ is the same for all x; accordingly, we henceforth refer to its rank as $\varrho(J_q)$ (removing the dependence on x from this notation). The result follows from condition (19) of Corollary 4.1, the factorization (21), and the fact that, for $k \geq q$, the matrix $J_k(x)$ can be written as $J_k(x) = \begin{bmatrix} J_q(x) & 0_{d\times(k-q)} \end{bmatrix}$. ■

We observe that this condition is on the structure of the regressor $\phi(x)$ only; in particular it imposes a constraint on the polynomial degree of the regressor vector[1].

In order to prove the main results of this Section, we shall need the following trivial technical Lemma.

Lemma 4.2: Consider the polynomial $d(x,u) \triangleq a(x) + \sum_{i=1}^{l} c_i(x)u^i$ where $a(x)$ and $\{c_i(x), i=1,\dots,l\}$ are scalar polynomials in the real variable x. Then the following two statements are equivalent

$$(i)\ \exists x_0 : d(x_0,u) = 0\ \forall u \tag{23}$$

$$(ii)\ a(x_0) = 0 \ \text{ and } c_i(x_0) = 0,\ i = 1,\dots,l \tag{24}$$

Proof: For a given x_0, the polynomial $d(x_0,u)$ is a polynomial of degree l in u with constant real coefficients. This polynomial is zero for all u (i.e. identically zero) if and only if all its coefficients are zero. ■

The first main result shows that in our analysis of identifiability we can replace the rank condition (19) on the infinite-dimensional matrix $R_\infty(\theta,x_0)$ by a rank condition on the finite matrix $R_q(\theta,x_0)$.

Theorem 4.2: For the model structure (10) with $deg(\phi(x)) = q$, for any given θ and any given x_0, the following two statements are equivalent:

$$\beta^T R_\infty(\theta,x_0) = \mathbf{0}\ \ \forall u \tag{25}$$

$$\beta^T R_q(\theta,x_0) = \mathbf{0}\ \ \forall u \tag{26}$$

Both statements imply

$$\beta^T J_q(x_0) diag(1,\dot{x},\dots,\dot{x}^q) = 0\ \ \forall u \tag{27}$$

Proof: (i) That (25) implies (26) is trivial since the equations (26) are a subset of the equations (25).
(ii) We next prove that (26) implies (25). To this end we will first show that for any positive integer k, $\beta^T R_k(\theta,x_0) = 0$ for all u implies

$$\beta^T J_k(x_0) diag\{1,\dot{x},\cdots,\dot{x}^k\} = 0\ \ \forall u, \tag{28}$$

thereby establishing (27). Then as $\phi(x)$ has degree q, and thus all its derivatives of order higher than q are zero, the result will follow from (20).

Thus, assume that (26) holds for some θ and x_0. The second equation of (26) is:

$$\beta^T \frac{\partial\phi}{\partial x}\dot{x}(0) = 0 \quad \text{for all} \quad u. \tag{29}$$

[1]It is also equivalent to the elements of $\phi(x)$ being all linearly independent in the space of polynomials over the field of real numbers.

This shows that $\beta^T R_1(\theta,x_0) = \mathbf{0}\ \ \forall u$ is equivalent to (28) with $k=1$. Now (29) is equivalent to

$$\beta^T \frac{\partial\phi}{\partial x}[\frac{1}{n(x_0)}(\theta^T\phi(x_0) + m(x_0)) + \frac{1}{n(x_0)}G(x_0)U] = 0$$

for all u. By Lemma 4.2 this is equivalent to

$$\beta^T \frac{\partial\phi}{\partial x}\frac{1}{n(x_0)}[\theta^T\phi(x_0) + m(x_0)] = 0 \tag{30}$$

$$\text{and}\quad \beta^T \frac{\partial\phi}{\partial x}\frac{1}{n(x_0)}G(x_0) = \mathbf{0} \tag{31}$$

Since $n(x_0) > 0$, the last equation is equivalent to

$$\beta^T \frac{\partial\phi}{\partial x}G(x_0) = \mathbf{0} \tag{32}$$

We have thus shown that the second equation of $\beta^T R_q(\theta,x_0) = 0\ \ \forall u$ is equivalent to (30) and (32).

Now consider the third equation of $\beta^T R_q(\theta,x_0) = 0\ \ \forall u$, i.e. $\beta^T\ddot{\phi} = 0$ for all u at x_0. Convenient expressions for the time derivatives of the state and of $\phi(x)$ are given in Lemmas 1.1 and 1.2 of the Appendix. Using (29) and (32) in (57) reduces the third equation to

$$\beta^T \frac{\partial^2\phi}{\partial x^2}\dot{x}(0)^2 = 0 \quad \text{for all} \quad u. \tag{33}$$

Thus we have shown that $\beta^T R_2(\theta,x_0) = 0$ for all u implies

$$\beta^T J_2(x_0)\begin{pmatrix} 1 & 0 & 0 \\ 0 & \dot{x} & 0 \\ 0 & 0 & \dot{x}^2 \end{pmatrix} = 0 \quad \text{for all} \quad u.$$

We now proceed by induction. Consider first that $k-1 \leq q$ and assume that $\beta^T R_{k-1}(\theta,x_0) = 0 \quad \text{for all} \quad u$ implies

$$\beta^T J_{k-1}(x_0) diag(1,\dot{x},\dots,\dot{x}^{k-1}) = 0 \quad \text{for all} \quad u. \tag{34}$$

Now consider the new equation $\beta^T\phi^{(k)} = 0$ at x_0 for all u (i.e. the last column of $\beta^T R_k(\theta,x_0) = 0$), and replace $\phi^{(k)}$ by its expression (56). It then follows that all terms of the second line of (56), premultiplied by β^T, are zero by (34), and all terms of the third line, premultiplied by β^T, are zero by (32). This shows that $\beta^T R_k(\theta,x_0) = 0$ for all u implies $\beta^T J_q(x_0) diag(1,\dot{x},\dots,\dot{x}^q) = 0$ for all u. Proceeding to $k = q+1$, we note that $\beta^T R_{q+1}(\theta,x_0) = 0$ for all u implies that the last column of $\beta^T R_{q+1}(\theta,x_0)$ is zero, i.e. $\beta^T\phi^{(q+1)} = 0$ at x_0 for all u, because $\frac{\partial^{(q+1)}\phi}{\partial x^{q+1}} = 0$ since $deg(\phi) = q$. Arguing in this way establishes (28). This implies that all columns of $\beta^T R_\infty(\theta,x_0)$ to the right of the $(q+1)$-st column are zero for all u. We have thus proved that (26) implies (25). ■

Now let $W_q(\theta,\dot{x},\dots,x^{(q)})$ be the square matrix consisting of the first $q+1$ columns of $W_\infty(\theta,\dot{x},\dots,x^{(q)})$; then

$$R_q(\theta,x) = J_q(x)W_q(\theta,\dot{x},\dots,x^{(q)}) \tag{35}$$

and we have the following final fundamental result on identifiability.

Theorem 4.3: Under Assumption 1, the model structure (10) is globally identifiable at all θ if and only if $\varrho(J_q) = d$.

Proof: Since $R_\infty(\theta, x)$ can always be factored as $R_\infty = J_q M$ for some M, necessity is obvious. We prove sufficiency by contradiction. Thus, assume that the model structure is not identifiable at some $\theta^\star$. By Corollary 4.1 this means that for all x_0 and $u(.)$ there exists $\beta \neq \mathbf{0}$ such that

$$\beta^T R_q(\theta, x_0) = \mathbf{0}. \tag{36}$$

By Theorem 4.2 this implies that there exists $\beta \neq \mathbf{0}$ such that

$$\beta^T J_q(x_0) diag(1, \dot{x}, \ldots, \dot{x}^q) = \mathbf{0} \quad \forall\, x_0,\ u. \tag{37}$$

In particular this implies, using the first equation of (37), that $\exists \beta \neq \mathbf{0}$ such that $\beta^T \phi(x_0) = 0$ for all x_0. But $\beta^T \phi(x) \equiv 0$ implies $\beta^T J_q(x) \equiv 0$ and hence $\beta^T R_q(\theta, x) \equiv \mathbf{0}$, which contradicts the assumption. ∎

Theorem 4.3 gives a necessary and sufficient condition for identifiability, and also provides the tools to further advance in the analysis of identifiability and of informativity in different scenarios. The informativity issue has not been dealt with so far in this paper. On the other hand, identifiability as defined previously in this paper means that we can chose an input **and** an initial condition such that the identification will succeed. But it is seldom the case that the designer of an identification experiment has the choice of an input signal and at the same time can decide on the initial condition, so identifiability in this sense has limited practical interest.

Thus we shall consider from now on the following two situations separately, which are the ones of practical interest:

1) the system has no external input, i.e. $u(t) \equiv 0$, and identifiability must be secured through the initial condition $x(0)$. If the model structure is identifiable through $x(0)$, then the informativity question amounts to finding a x_0 that delivers informative data via the transient response.
2) the input signal $u(.)$ must provide informative data whatever the initial condition x_0; in this second situation, identifiability must be guaranteed by the input signal regardless of the initial condition, i.e. one must assure that "adversarial" initial conditions do not kill the excitation coming from $u(.)$. If the model structure is identifiable through $u(.)$, then the informativity question amounts to finding a sufficiently rich $u(.)$ that provides informative data.

This second situation is the most commonly treated and probably the most commonly found in practice, but the first one is also found in a variety of applications, particularly in (bio-) chemical batch process [15], [3], [7]. For pedagogical reasons, we start our analysis with the simpler case where the identifiability of the parameters must be secured through the response to an initial state, assuming that there is no driving input to the system.

V. IDENTIFIABILITY AND INFORMATIVITY FROM THE INITIAL STATE

A. Definitions

When $u(t) \equiv 0$ the models (8) and (9) become:

$$\dot{x} = f(x, \theta^\star), \quad y = h(x, \theta^\star), \quad x(0) = x_0 \tag{38}$$
$$\dot{\hat{x}} = f(\hat{x}, \theta), \quad \hat{y} = h(\hat{x}, \theta), \quad \hat{x}(0) = \hat{x}_0 \tag{39}$$

Definition 4: (Identifiability at $\theta^\star$ from $x(0)$) The model structure (38) is *globally identifiable at $\theta^\star$ from the initial condition $x(0)$* if there exists an initial condition $x_0 = \hat{x}_0$ such that, for all $\theta \in \Re^d$, the outputs of the models (38) and (39) are identical (i.e. $y(t) = \hat{y}(t) \;\; \forall t \geq 0$) **only if** $\theta = \theta^\star$.

Definition 5: (Informativity of the initial condition at $\theta^\star$) The initial condition x_0 is globally informative at $\theta^\star$ for the model structure (38) if for all $\theta \in \Re^d$, the outputs of the models (38) and (39) with initial condition $x_0 = \hat{x}_0$ are identical (i.e. $y(t) = \hat{y}(t) \;\; \forall t \geq 0$) **only if** $\theta = \theta^\star$.

B. Identifiability and informativity results

We now provide necessary and sufficient conditions for identifiability and informativity at a given θ from the initial state, for the model class of rational models (6) in which the input is assumed zero:

$$\dot{x} = \frac{1}{n(x)}[\theta^T \phi(x) + m(x)], \quad x(0) = x_0 \tag{40}$$

We first show that in the case where the input is zero, Theorem 4.2 can be replaced by the following stronger statement.

Theorem 5.1: For the model class (40), with $deg(\phi(x)) = q$, the following three statements are equivalent:

$$\beta^T R_\infty(\theta, x_0) = \mathbf{0} \tag{41}$$
$$\beta^T R_q(\theta, x_0) = \mathbf{0} \tag{42}$$
$$\beta^T J_q diag(1, \dot{x}, \ldots, \dot{x}^q) = \mathbf{0} \tag{43}$$

Proof: As in the proof of Theorem 4.2, it is clear that (41) implies (42). We now prove that (42) is equivalent to (43), which implies (41). Since $\dot{\phi}(x) = \frac{\partial \phi(x)}{\partial x}\dot{x}$, it follows that $R_1(\theta, x) = J_1(x) diag(1, \dot{x})$. Now consider $\beta^T R_2(\theta, x) = \mathbf{0}$; the third equation is $\beta^T \ddot{\phi} = 0$. By equation (57) of the Appendix, this is equivalent to

$$\beta^T \frac{\partial^2 \phi}{\partial x^2}\dot{x}^2 + \beta^T \frac{\partial \phi}{\partial x} m_{2,1}(\theta, x, u)\dot{x} = 0$$

The second term is zero by the second equation of $\beta^T R_2(\theta, x) = \mathbf{0}$ and hence

$$\beta^T R_2(\theta, x) = \mathbf{0} \;\text{ if and only if }\; \beta^T J_2(x) diag(1, \dot{x}, \ddot{x}) = \mathbf{0}.$$

By an induction step, just as in the proof of Theorem 4.2, it follows that for all $k \geq 1$, $\beta^T R_k(\theta, x) = \mathbf{0}$ is equivalent to $\beta^T J_2(x) diag(1, \dot{x}, \ddot{x}) = \mathbf{0}$. That (43) implies (41) then follows from the fact that the columns of J_k become zero for all columns to the right of the $(q+1)$-st column. ∎

The main result on identifiability from the initial state follows.

Theorem 5.2: Under Assumption 1, the model class (40) is globally identifiable from the initial state at every θ if and only if $\varrho(J_q) = d$.

Proof: By Corollary 4.1, the model class is globally identifiable at some θ^* if and only if there exists x_0 such that $\varrho(R_\infty(\theta^*, x_0)) = d$. By Theorem 5.1 this is equivalent to the existence of a x_0 such that

$$\beta^T J_q(x_0) diag(1, \dot{x}(0), \ldots, \dot{x}^q(0)) = \mathbf{0} \quad \Rightarrow \quad \beta = \mathbf{0} \tag{44}$$

It follows immediately that $J_q(x_0)$ must be full rank, which proves necessity. For sufficiency, assume that J_q has full rank and suppose that for all x_0 there exists a $\beta \neq \mathbf{0}$ such that (44) holds. Then the first equation of (44) yields $\beta^T \phi(x) = 0$ for all x, which implies $\beta^T J_q(x) = \mathbf{0}$ for all x; this violates the assumption. ■

We also have an immediate characterization of an informative initial condition.

Theorem 5.3: Let the model structure (40) be globally identifiable from the initial state at a given θ^* (i.e. $\varrho(J_q) = d$) and let $d > 1$. Then an initial condition $x_0 \in \Re$ yields a globally informative experiment at θ^* if and only if x_0 is not a root of the polynomial equation

$$\phi^T(x_0)\theta^* + m(x_0) = 0. \tag{45}$$

■

That such an x_0 exists is guaranteed by Assumption 1. In fact, for any given θ almost all initial conditions are informative, since the polynomial $\theta^T \phi(x) + m(x)$ has a finite number of roots.

The condition $d > 1$ is included in Theorem 5.3 for the sake of making the statement necessary and sufficient. For $d = 1$, that is a scalar $\phi(x)$, θ may still be identifiable when $\phi(x_0)\theta + m(x_0) = 0$ provided that $\phi(x_0) \neq 0$.

C. An illustrative example

Example 1

Consider the model structure

$$\dot{x} = \frac{\theta_1 x^2 + \theta_2 x}{1 + x^2} = \frac{1}{1 + x^2} \theta^T \phi(x), \;\; x(0) = x_0 \tag{46}$$

with $\theta^T = [\theta_1 \;\; \theta_2]$ and $\phi^T(x) = [x^2 \;\; x]$. We first observe that $q = 2$ and

$$J_2(x) = \begin{pmatrix} x^2 & 2x & 2 \\ x & 1 & 0 \end{pmatrix}$$

so that $\varrho(J_2(x)) = 2$ for all x; hence the model structure is globally identifiable from the initial state at every θ. The factorization (21) with $k = q = 2$ yields the 2×3 matrix $R_2(\theta, x)$ as:

$$R_2(\theta, x) = J_2(x) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \dot{x} & \ddot{x} \\ 0 & 0 & \dot{x}^2 \end{pmatrix} = J_2(x) W(\theta, \dot{x}, \ddot{x}) \tag{47}$$

Theorem 5.3 states that an initial state x_0 is informative at some θ if and only if $\theta_1 x_0^2 + \theta_2 x_0 \neq 0$. But $\theta_1 x_0^2 + \theta_2 x_0 = 0$ implies either $x_0 = 0$ or $x_0 = -\frac{\theta_2}{\theta_1}$. Any initial condition other than these two yields $\theta_1 x_0^2 + \theta_2 x_0 \neq 0$.

Therefore any initial condition other than $x_0 = 0$ or $x_0 = -\frac{\theta_2}{\theta_1}$ is informative. One of these uninformative initial conditions is a function of the unknown θ. Thus we conclude that two experiments with two different nonzero initial conditions will always allow the estimation of the parameter vector θ.

Before concluding, we illustrate the equivalence between (42) and (43) for this example. We have

$$R_2(\theta, x) = \begin{pmatrix} x^2 & 2x\dot{x} & 2x\ddot{x} + 2\dot{x}^2 \\ x & \dot{x} & \ddot{x} \end{pmatrix}$$

$$J_2(x) \; diag(1, \dot{x}, \dot{x}^2) = \begin{pmatrix} x^2 & 2x\dot{x} & 2\dot{x}^2 \\ x & \dot{x} & 0 \end{pmatrix}$$

From the model equation, we also have

$$\ddot{x} = \frac{1}{(1 + x^2)^2} \left(2\theta_1 x + \theta_2 - \theta^2 x^2 \right) \dot{x}$$

The determinants of the minors of R_2 are $-x^2\dot{x}$, $-x^2\ddot{x} - 2x\dot{x}^2$ and $-2\dot{x}^3$, while those of $J_2(x) diag(1, \dot{x}, \dot{x}^2)$ are $-x^2\dot{x}$, $-2x\dot{x}^2$ and $-2\dot{x}^3$. Substituting for $\ddot{x}$, we conclude that, for both matrices, the determinants of all minors are zero at some x_0 if and only if $\dot{x} = 0$ at x_0.

VI. IDENTIFIABILITY AND INFORMATIVITY FROM THE INPUT

We now seek identifiability conditions and informativity from the input *regardless of the initial state*. The definitions of Section III can be adapted to this case as follows.

A. Definitions: Identifiability and informativity from $u(.)$

Consider the model structure

$$\dot{x} = \frac{\theta^{\star T} \phi(x) + m(x) + G(x)U}{n(x)}, \quad x(0) = x_0 \tag{48}$$

and the same model structure at some other parameter vector θ with the same initial condition:

$$\dot{\hat{x}} = \frac{\theta^T \phi(\hat{x}) + m(\hat{x}) + G(\hat{x})U}{n(\hat{x})}, \quad \hat{x}(0) = x_0 \tag{49}$$

Definition 6: (Identifiability at $\theta^\star$ from u) The model structure (48) is *globally identifiable at $\theta^\star$ from the input u* if there exists an input $u(.)$ such that, for all initial conditions x_0 and for all $\theta \in \Re^d$, the solutions of the models (48) and (49) are identical (i.e. $x(t) = \hat{x}(t) \;\; \forall t \geq 0$) **only if** $\theta = \theta^\star$.

This definition is consistent with the definition of global identifiability at some θ^* used in [12] and adopted in [2] for LTI systems.

Definition 7: (Informativity of the input at $\theta^\star$) The input signal $u(.)$ is globally informative at $\theta^\star$ for the model structure (48) if, for all initial conditions x_0 and for all $\theta \in \Re^d$, the solutions of the models (48) and (49) with this input $u(.)$ are identical (i.e. $x(t) = \hat{x}(t) \;\; \forall t \geq 0$) **only if** $\theta = \theta^\star$.

B. Identifiability from the input

The following theorem is the main result on identifiability from the input.

Theorem 6.1: Consider the model structure (48) with $deg(\phi(x)) = q$ and let $d > 1$. This model structure is globally identifiable at $\theta^\star$ from the input u if and only if the following two conditions hold simultaneously:
(i) $\varrho(J_q) = d$;
(ii) the polynomials $\theta^{\star T}\phi(x) + m(x)$ and $\{g_i(x), i = 1,\ldots,l\}$ have no common real root w.r.t. x.
Proof: First note that, from Definition 6 and Theorem 4.1, identifiability at θ^* from u is equivalent to the existence of an input $u(.)$ such that $\varrho(R_\infty(\theta^\star, x_0)) = d \;\; \forall x_0$.
We first prove sufficiency, by contradiction: suppose that the system is not globally identifiable at $\theta^\star$. Then there exists a x_0 and $\beta \neq 0$ such that (25) holds. By Theorem 4.2 this implies (27), which implies that either $\varrho(J_q) < d$, which violates condition (i), or $\dot{x}(0) = 0 \; \forall u$ at x_0. By Lemma 4.2, the latter is equivalent to $\theta^{\star T}\phi(x_0) + m(x_0) = 0$ and $g_i(x_0) = 0, i = 1,\ldots,l$, which violates condition (ii).
Next we prove necessity, by showing that if either (i) or (ii) is violated then there exists a x_0 such that $\varrho(R_\infty(\theta^\star, x_0)) < d$. Suppose first that $\varrho(J_q) < d$. Then there exists a x_0 and a $\beta \neq \mathbf{0}$ such that $\beta^T J_q(x_0) = \mathbf{0}$. Since $R_\infty(\theta, x_0)$ can always be written as $R_\infty(\theta, x_0) = J_q(x_0)M(\theta, x_0)$ for some $M(\theta, x_0)$, it follows that $\beta^T R_\infty(\theta, x_0) = \mathbf{0}$. Suppose now that condition (ii) is violated, i.e. there exists x_0 such that all polynomials appearing in (ii) are zero at x_0; in particular $G(x_0) = \mathbf{0}$. It then follows from the model equation (48) and from $n(x_0) > 0$ that $\dot{x}(0) = 0$ for all u. Therefore, by equation (56) of the Appendix, $\phi^{(k)} = \mathbf{0}$ at x_0 for all k, and hence the model structure is not identifiable. ■

The condition $d > 1$ is there for the sake of making the statement necessary and sufficient. For $d = 1$ we have the following

Corollary 6.1: Consider the model structure (6) with $deg(\phi(x)) = q$ and let $d = 1$. This model structure is identifiable at $\theta^\star$ from the input u if the following two conditions hold simultaneously:
(i) $\varrho(J_q) = 1$;
(ii) the polynomials $\theta^\star\phi(x)+m(x)$ and $\{g_i(x), i = 1,\ldots,l\}$ have no common real root w.r.t. x.
Moreover, condition (ii) is not necessary for identifiability from the input.
Proof: From Theorem 6.1, under the necessary condition $\varrho(J_q) = d$, identifiability from the input is equivalent to $\theta^{\star T}\phi(x_0)+m(x_0)+G(x_0)U \neq 0$ for all $x_0 \in \Re$. This, in its turn, is implied by condition (ii), but only implies condition (ii) for $d > 1$. Indeed, for $d = 1$, if $\theta^\star\phi(x_0) + m(x_0) = 0$ and $\{g_i(x_0) = 0, i = 1,\ldots,l\}$ for some x_0, then $\theta^\star$ may still be identified from the equation $\phi(x_0)\theta^\star + m(x_0) = 0$ provided that $\phi(x_0) \neq 0$. ■

C. Informativity from the input

If the identifiability conditions of Theorem 6.1 are fulfilled, then there exists an informative input signal, that is, one that allows the discrimination between two different parameter vectors θ and $\theta^\star$ regardless of the initial condition. However, Theorem 6.1 does not provide a clue as to which input signals will be informative. Obtaining necessary and sufficient conditions on the input signal u to be informative for the model class (48) regardless of the initial condition turns out to be very difficult. The objective of the present subsection is twofold: we first provide a sufficient condition for informativity of the signal u; we then illustrate with an example that this condition is not necessary, i.e. for this example we construct an informative input that does not satisfy this sufficient condition.

Theorem 6.2: Assume that the model structure (48) is globally identifiable at some $\theta^\star$. An input signal $u(.)$ is informative at $\theta^\star$ if at any given time, say $t = 0$, $\dot{x} \neq 0 \; \forall x_0$.
Proof: It follows from Definition 7 that an input signal $\{u(\tau), \tau \geq 0\}$ is informative at $\theta^\star$ if and only if the matrix $R_\infty(\theta^\star, x_0)$ has full row rank for all $x_0 \in \Re$. Now consider $R_q(\theta^\star, x_0)$ and factorize this matrix as in (21):

$$R_q(\theta^\star, x_0) = J_q(x_0)W_q(\theta^\star, \dot{x}, \ldots, x^{(q)}) \tag{50}$$

where the derivatives in W_q are evaluated at $t = 0$. It follows from the form of W_q (see (22)) that it is nonsingular for all x_0 if $\dot{x}(0) \neq 0$ for all x_0. Since $\varrho(J_q) = d$ by the global identifiability assumption, it then follows from the Sylvester inequality applied to (50):

$$\varrho(J_q) + \varrho(W_q) - (q+1) \leq \varrho(R_q),$$

that $\varrho(R_q(\theta^\star, x_0)) = d$ for all x_0 if $\dot{x}(0) \neq 0$ for all x_0. Since $R_q(\theta^\star, x_0)$ is a submatrix of $R_\infty(\theta^\star, x_0)$, the result follows. ■

The sufficient condition on u can equivalently be stated as follows.

Corollary 6.2: Assume that the model structure (48) is globally identifiable at some $\theta^\star$. An input signal $u(.)$ is informative at $\theta^\star$ if

$$\theta^{\star T}\phi(x_0) + m(x_0) + \sum_{i=1}^{l} g_i(x_0)u^i \neq 0 \;\; \forall x_0. \tag{51}$$

The following example shows that the condition of Theorem 6.2 is not necessary; but it also illustrates that there may be no input that satisfies the condition (51) even though an informative input exists. We shall construct such informative input by directly ensuring that the matrix $R_q(\theta^\star, x_0)$ has full rank for all x_0.

Example 2
Consider the model structure

$$\dot{x} = \frac{1}{1+x^2}[\theta_1 x^2 + \theta_2 x + u] = \frac{1}{1+x^2}[\theta^T\phi(x) + u], \tag{52}$$

with θ and $\phi(x)$ as in (46). The model structure is identifiable at all θ since the two conditions of Theorem 6.1 are satisfied. Now consider informativity.

We observe immediately that there is no input $u(.)$ that makes $\dot{x}(0) \neq 0$ for all x_0; thus the sufficient condition of Theorem 6.2 cannot be used to construct an informative input. However, we now show that one can construct an input $u(.)$ that makes $R_2(\theta, x_0)$ full rank for all x_0. $R_2(\theta, x)$ takes the form:

$$R_2(\theta, x_0) = \begin{pmatrix} x_0^2 & 2x_0\dot{x}(0) & 2x_0\ddot{x}(0) + 2\dot{x}^2(0) \\ x_0 & \dot{x}(0) & \ddot{x}(0) \end{pmatrix}$$

We observe that the determinants of the three minors of R_2 are, respectively, $-x_0^2\dot{x}(0)$, $-2\dot{x}^2(0)$ and $(-x_0^2\ddot{x}(0) - 2x_0\dot{x}^2(0))$. Substitute

$$\begin{aligned} \dot{x} &= \theta_1 x_0^2 + \theta_2 x_0 + u(0) \\ \ddot{x} &= 2\theta_1 \dot{x} x_0 + \theta_2 \dot{x} + \dot{u}(0) \end{aligned}$$

and consider 2 cases separately: $x_0 = 0$ and $x_0 \neq 0$.

Case 1: Suppose first that $x_0 = 0$; then $\dot{x}(0) = u(0)$, $\ddot{x}(0) = \theta_2 u(0) + \dot{u}(0)$, and

$$R_2(\theta, 0) = \begin{pmatrix} 0 & 0 & 2u^2(0) \\ 0 & u(0) & \theta_2 u(0) + \dot{u}(0) \end{pmatrix}$$

Thus, $R_2(\theta, 0)$ has full rank if and only if $u(0) \neq 0$.

Case 2: Take $u(0) \neq 0$ and consider $x_0 \neq 0$. Again two cases must be considered.
(a) Suppose first that $\dot{x}(0) \neq 0$. Then $R_2(\theta, x_0)$ has full rank because the determinant of the second minor $-2\dot{x}^2(0) \neq 0$.
(b) Suppose next that $\dot{x}(0) = 0$. Then the determinant of the third minor of $R_2(\theta, x_0)$ is given by $-x_0^2\ddot{x}(0) = -x_0^2\dot{u}(0)$ which is nonzero if we take $\dot{u}(0) \neq 0$.

These informativity conditions are independent of θ. We conclude that an input $u(.)$ such that at some time, say $t = 0$, $u(0) \neq 0$ and $\dot{u}(0) \neq 0$ is globally informative for the model structure (52).

Example 2 illustrates an interesting feature of the theory developed in this paper. For this example, we have $q + 1 > d$. In addition, since the model structure contains an input-driven term, it does not hold that $\ddot{x}$ contains $\dot{x}$ as a factor: see (53). It then follows from the factorization of R_q in (47) that $\dot{x}(0) \neq 0$ is not a necessary condition for the full rankness of R_q. Indeed we observe in our example that, even if $\dot{x}(0) = 0$, we can still find an informative input by taking $\dot{u}(0) \neq 0$. Conversely, there is no input that makes $\dot{x}(0) \neq 0$ for all x_0.

VII. CONCLUSIONS

In this paper we have provided necessary and sufficient conditions for identifiability from the initial state, as well as necessary and sufficient conditions for informativity from the initial state for a class of nonlinear rational models. Parameter estimation from a transient response is of great interest in many practical applications. In addition, for this same class of models we have provided necessary and sufficient conditions for identifiability from the input signal regardless of the unknown initial condition. Sufficient conditions for the generation of informative experiments from the input have also been derived, and we have illustrated with an example the difficulty of obtaining necessary and sufficient informativity conditions for such case.

The properties that make an initial state and/or an input globally informative at given θ are generic. But informativity, as well as identifiability, refers to a given model structure at a given parameter value; so for any given experiment there are parameter values for which it is not informative. Since the properties of identifiability and informativity should be valid at the unknown real parameter, care must still be exercised in avoiding noninformative experimental conditions.

APPENDIX

Lemma 1.1: Consider the model structure (6). The successive derivatives of x for $k \geq 1$ can all be written in the form

$$\begin{aligned} x^{(k+1)} &= f_k(\theta, x, u, \dot{u}, \ldots, u^{(k-1)})\dot{x} \\ &\quad + G(x) H_k(\theta, x, u, \dot{u}, \ldots, u^{(k)}) \end{aligned} \tag{53}$$

where H_k is a l-dimensional column vector.
Proof: The second derivative is given by

$$\begin{aligned} \ddot{x} &= \frac{1}{n(x)} \left\{ \left[\theta^T \frac{\partial \phi}{\partial x} + \frac{\partial m}{\partial x} + \frac{\partial G}{\partial x} U - \frac{\partial n}{\partial x} \dot{x} \right] \dot{x} \right. \\ &\qquad \left. + G(x) \frac{\partial U}{\partial u} \dot{u} \right\} \end{aligned}$$

By substituting for $\dot{x}$ within the square bracket, this can be rewritten in the desired form as

$$\ddot{x} = f_1(\theta, x, u)\dot{x} + G(x) H_1(x, u, \dot{u}). \tag{54}$$

Taking the time derivative of $\ddot{x}$ and substituting again yields

$$x^{(3)} = f_2(\theta, x, u)\dot{x} + G(x) H_2(\theta, x, u, \dot{u}) \tag{55}$$

where the dependence of H_2 on θ comes from the substitution of $\ddot{x}$ by (54) in a term involving $G(x)$. We now proceed by induction. Suppose we can write

$$\begin{aligned} x^{(k)} &= f_{k-1}(\theta, x, u, \dot{u}, \ldots, u^{(k-2)})\dot{x} \\ &\quad + G(x) H_{k-1}(\theta, x, u, \dot{u}, \ldots, u^{(k-1)}) \end{aligned}$$

Taking the time derivative of this expression yields

$$\begin{aligned} x^{(k+1)} &= \frac{df_{k-1}}{dt}\dot{x} + f_{k-1}\ddot{x} \\ &\quad + \frac{\partial G}{\partial x} H_{k-1}\dot{x} + G(x)\frac{dH_{k-1}}{dt} \end{aligned}$$

Substituting $\ddot{x}$ by (54) in this expression, and grouping the terms that have $\dot{x}$ as factor and those that have $G(x)$ as factor, yields the expression (53). ∎

The next Lemma gives, similarly, a general expression for the derivatives of $\phi(x)$.

Lemma 1.2: Consider the model structure (6). The successive derivatives of $\phi(x)$ for $k \geq 1$ can all be written in the form

$$\begin{aligned}\phi^{(k)} &= \frac{\partial^k \phi}{\partial x^k}\dot{x}^k \qquad (56)\\ &+\sum_{l=1}^{k-1}\frac{\partial^{k-l}\phi}{\partial x^{k-l}}m_{k,k-l}(\theta,x,u,\dot{u},\ldots,u^{(k-2)})\dot{x}\\ &+\frac{\partial \phi}{\partial x}G(x)N_k(\theta,x,u,\dot{u},\ldots,u^{(k-1)})\end{aligned}$$

where N_k is a l-dimensional column vector.
Proof: We have $\dot{\phi} = \frac{\partial \phi}{\partial x}\dot{x}$ and

$$\ddot{\phi}(x) = \frac{\partial^2 \phi}{\partial x^2}\dot{x}^2 + \frac{\partial \phi}{\partial x}\ddot{x}$$

Substituting (54) for $\ddot{x}$ we get

$$\ddot{\phi}(x) = \frac{\partial^2 \phi}{\partial x^2}\dot{x}^2 + \frac{\partial \phi}{\partial x}m_{2,1}(\theta,x,u)\dot{x} + \frac{\partial \phi}{\partial x}G(x)N_2(x,u,\dot{u}) \qquad (57)$$

which has the desired form for $k = 2$. We now proceed by induction: we assume that (56) holds for some k, and we take the time derivative. We compute separately the derivative of the three lines of (56):

- the derivative of the first line yields

$$\frac{\partial^{k+1}\phi}{\partial x^{(k+1)}}\dot{x}^{k+1} + 2\frac{\partial^k \phi}{\partial x^k}\dot{x}^{k-1}\ddot{x} = \frac{\partial^{k+1}\phi}{\partial x^{(k+1)}}\dot{x}^{k+1}$$
$$+2\frac{\partial^k \phi}{\partial x^k}\dot{x}^{k-1}[f_1(\theta,x,u)\dot{x} + G(x)H_1(x,u,\dot{u})]$$

- the derivative of the second line yields

$$\sum_{l=1}^{k-1}\frac{\partial^{k-l+1}\phi}{\partial x^{k-l+1}}m_{k,k-l}\dot{x}^2 + \sum_{l=1}^{k-1}\frac{\partial^{k-l}\phi}{\partial x^{k-l}}[\frac{dm_{k,k-l}}{dt}\dot{x} + m_{k,k-l}\ddot{x}]$$

After substituting for $\ddot{x}$ again, the second term becomes

$$\sum_{l=1}^{k-1}\frac{\partial^{k-l}\phi}{\partial x^{k-l}}[\frac{dm_{k,k-l}}{dt}\dot{x} + m_{k,k-l}f_1\dot{x} + m_{k,k-l}G(x)H_1]$$

- the derivative of the third line yields

$$\frac{\partial^2 \phi}{\partial x^2}G(x)N_k\dot{x} + \frac{\partial \phi}{\partial x}G(x)\frac{dN_k}{dt} + \frac{\partial \phi}{\partial x}\frac{\partial G(x)}{\partial x}N_k\dot{x}$$

Regrouping the coefficients of the successive derivatives $\frac{\partial^{k-l}\phi}{\partial x^{k-l}}$ yields the form (56). ■

References

[1] E. W. Bai and S. S. Sastry. Persistence of excitation, sufficient richness and parameter convergence in discrete time adaptive control. *Systems and Control Letters*, 6:153–163, 1985.

[2] 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.

[3] Dochain D. State and parameter estimation in chemical and biochemical processes: a tutorial. *Journal of Process Control*, 13:801–818, 2003.

[4] A.G. Dankers, R. Toth, P.S.C. Heuberger, X. Bombois, and P.M.J. Van den Hof. Informative data and identifiability in lpv-arx prediction-error identification. In *CD-ROM of 50th IEEE Conf. on Decision and Control and European Control Conference*, pages 799–804, Orlando, Florida, USA, December 2011. IEEE.

[5] 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.

[6] S. Dasgupta, Y. Shrivastava, and G. Krenzer. Persistent excitation in bilinear systems. *IEEE Transactions on Automatic Control*, 36(3):305–313, March 1991.

[7] Franceschini G. and Macchietto S. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63:4846–4872, 2008.

[8] 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.

[9] M. Gevers, A.S. Bazanella, X. Bombois, and L. Mišković. Identification and the information matrix: how to get just sufficiently rich? *IEEE Transactions on Automatic Control*, 54(12):2828–2840, December 2009.

[10] G. Kreisselmeier. Richness and excitation on an interval with application to continuous-time adaptive control. *IEEE Transactions on Automatic Control*, 35(2):165–171, February 1990.

[11] J.S. Lin and I. Kanellakopoulos. Nonlinearities enhance parameter convergence in output-feedback systems. *IEEE Transactions on Automatic Control*, 43(2):204–222, February 1998.

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

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

[14] I.M.Y. Mareels and M. Gevers. Persistence of excitation criteria for linear, multivariable, time-varying systems. *Mathematics of Control, Signals and Systems*, 1(3):203–226, 1988.

[15] A. Vrsalovic Presecki and D. Vasic-Racki. Modelling of the alcohol dehydrogenase production in baker's yeast. *Process Biochemistry*, 40:2781–2791, 2005.

[16] E.D. Sontag, Y. Wang, and A. Megretski. Input classes for identifiability of bilinear systems. *IEEE Transactions on Automatic Control*, 54(2):195–207, February 2009.