

On a Geometric Approach to the Structural Identifiability Problem and its Application in a Water Quality Case Study

J.D. Stigter and R.L.M. Peeters

Abstract—We present and apply an alternative method for the investigation of the well-known parameter identifiability question for non-linear system models. The method is based on a geometric analysis of the parametric output sensitivities and is, in fact, an application of the tools that are available in non-linear control theory to an augmented system, including the parametric output sensitivities. Accessibility Lie algebras are calculated that yield insight (through a simple rank test) in the controllability of this augmented system. The method is demonstrated in an example that is due to Dochain *et al* [4]. Results are confirmed by the method that has certain advantages in comparison to, for example, the Taylor series approach that seeks for identifiable combinations of parameters through inspection of the individual terms in a Taylor series expansion of the output signal, i.e. application of the well-known method of Pohjanpalo [15]. Parametric output sensitivities (as already noted by Dötsch and Van den Hof [5] and Peeters and Hanzon [13]) play a crucial role in identifiability analysis and we further elaborate on this insight in the current paper. Our goals are (i) to present an interesting method for addressing the (local) identifiability question for non-linear systems and (ii) to gain better understanding in the role of parametric state- and output sensitivities in the identifiability question that stems from an alternative perspective, and that has not been presented in the identification literature. Of course, we are aware of other algorithms and software that establishes an answer to the identifiability question, albeit from a different perspective, e.g. [19], but seek in the current paper mainly for another interpretation and computational framework to address the question of local identifiability, shedding some new light on the problem.

I. INTRODUCTION

It has been argued by many that the identifiability question plays a crucial role in any practical experimental setup that is developed for determination of the values of the model parameters from input-output data, e.g. [1, 24]. Indeed, identification is a central and fundamental issue that should always be addressed. The identification literature shows many papers that deal with the identifiability problem, e.g. [2, 3, 5, 11, 14, 16–18, 21, 23–25], and several solutions exist. In an interesting paper by Dochain *et al* two of these methods (the Taylor series approximation and a transformation of the model to one that is linear-in-the-parameters) are considered for data retrieved from a Rapid Oxygen Demand TOXicity (RODToX) device in an experiment that involves injection of a substrate into a bioreactor [4]. Of the two methods that were utilized to enforce a decision on the identifiability question for certain bio-kinetic parameters (or combinations

of these parameters) in the biokinetic model, the most general approach of these is the Taylor series expansion. Its most important limitation is that it is not clear, ‘a priori’, how many terms in the series must be included to solve the identifiability question unambiguously. The approach taken here has similarities with the Taylor series expansion method and in case of the existence of an input signal in the experiment (meaning not just an initial condition but also a manipulation variable $\mathbf{u}(t)$) the result can be calculated faster and a decision on the number of terms needed in the expansion can be reached quickly with the method introduced here. But let us first introduce some notation and definitions that allow us to introduce the algorithm for an identifiability test of a given experimental setup in more detail.

A. Problem definition

Global identifiability is concerned with the question whether it is possible to uniquely identify a model from a given parameterized model class on the basis of its associated input-output behavior. To be more precise, let \mathcal{M} be a model class and $\mathcal{M}(\theta)$ an actual model from this class, where $\theta \in \Theta \subset \mathbb{R}^p$ denotes the vector of model parameters. A model $\mathcal{M}(\theta)$ addressed in this paper is represented by a set of model equations and accompanying initial conditions. From an input-output point of view, such a model is regarded to define an associated *input-output mapping* $\Sigma(\theta) : U \rightarrow Y$, which takes admissible input signals $\mathbf{u}(\cdot)$ from an input space U into output signals $\mathbf{y}(\cdot)$ in an output space Y . (In terms of the behavioral approach to systems theory, the *graph* of this input-output map corresponds to the model *behavior*, which is the set of all pairs $(\mathbf{u}(\cdot), \mathbf{y}(\cdot))$ that are consistent with the model specifications.) The question of global identifiability of a model class \mathcal{M} can now be phrased as one of (global) *injectivity* of the map $\Psi : \theta \mapsto \Sigma(\theta)$. Likewise, local identifiability can be defined in terms of *local injectivity* of Ψ . A specific model $\mathcal{M}(\theta)$ is called identifiable if Ψ is injective at θ , i.e., if it holds true for all $\tilde{\theta} \in \Theta$ that

$$\Sigma(\tilde{\theta}) = \Sigma(\theta) \Rightarrow \tilde{\theta} = \theta.$$

Note that identifiability depends not only on the structure of the model equations in $\mathcal{M}(\theta)$, but that crucial roles are played by the definitions of Θ and U too. In this paper we shall require that Θ constitutes an open and connected non-empty subset of \mathbb{R}^p and that the input space U contains arbitrary piecewise constant functions.

The space $\Psi(\Theta)$ will often exhibit a manifold structure (of course depending on the smoothness of the model equations,

Hans Stigter is with the Systems and Control Group, Wageningen University, P.O.Box 17, 6700 AA Wageningen, The Netherlands

Ralf Peeters is with the Department of Mathematics, Maastricht University, P.O.Box 616, 6200 MD Maastricht, The Netherlands

especially with respect to θ). Note that the structure of the manifold $\Psi(\Theta)$ can be manipulated through a choice of sensors that allow observation of the system's input-output behavior and, in the identification context, *may* allow reconstruction of the values of the parameters. Of course, the manifold $\Psi(\Theta)$ only has substance if the system has been properly excited, meaning that we can only speak of an interesting identification problem if we observe non-trivial input-output behavior. Excitation of the system is achieved through (i) the input signal $\mathbf{u}(\cdot) \in U$ and (ii) the initial condition $\mathbf{x}(0)$ at which the system starts. For the current study we put no emphasis on the well-known fact that some of the input signals may in practice not be admissible, i.e. substrate and biomass concentrations cannot be negative, for example, or flow-rates are usually bounded by upper and lower bounds.

For convenience our model class \mathcal{M} has been limited to the class of smooth-affine control systems (which is still a very general class):

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \theta) + \mathbf{g}(\mathbf{x}(t), \theta) \mathbf{u}(t) \quad (1)$$

$$\mathbf{x}(0) = \mathbf{x}_0(\theta) \quad (2)$$

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \theta) \quad (3)$$

where $\mathbf{f}(\mathbf{x}(t), \theta)$, $\mathbf{g}(\mathbf{x}(t), \theta)$, $\mathbf{x}_0(\theta)$ and $\mathbf{h}(\mathbf{x}(t), \theta)$ all depend smoothly on their respective arguments. (The dependence on the time t has been suppressed for the sake of readability.) This model class is very often used in mechanical, biological and chemical engineering and it therefore relates to most of the practical case studies. Two remarks can be made at this point: First, in the above model definition it may also be assumed that the initial condition $\mathbf{x}(0)$ of the system is known *with complete certainty*. Such an assumption can be relaxed by noting that (part of) the initial condition vector may be parameterized through some additional parameters that then become part of the identification problem. This may be the case, for example, in biochemical case studies where the initial substrate concentration is sometimes completely unknown so that its value must be determined from the input-output data. For clarity, we have chosen to incorporate the possible dependence of $\mathbf{x}(0)$ on θ explicitly in (2). Second, the parameters in the model are assumed to be truly constants, i.e. they do not depend on time. Of course, time varying parameters might as well be incorporated as additional states, describing the dynamics of these parameters over time, and these particular equations then contain new time-invariant coefficients that become part of the identification problem.

II. PARAMETRIC STATE AND OUTPUT SENSITIVITIES

In order to find an answer to the question whether the mapping $\Psi : \theta \mapsto \Sigma(\theta)$ is locally injective, an additional set of $n \times p$ equations is defined, known as the system of local parametric *state sensitivities* associated with the model

$\mathcal{M}(\theta)$, i.e.

$$\begin{aligned} \dot{\mathbf{x}}_\theta(t) &= \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \mathbf{x}_\theta(t) + \frac{\partial \mathbf{f}}{\partial \theta} + \left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}} \mathbf{x}_\theta(t) + \frac{\partial \mathbf{g}}{\partial \theta} \right) \mathbf{u}(t) \quad (4) \\ \mathbf{x}_\theta(0) &= \frac{\partial \mathbf{x}_0(\theta)}{\partial \theta} \end{aligned} \quad (5)$$

Here, for ease of notation, the arguments of the various functions have been dropped, and $\mathbf{x}_\theta(t) = \begin{pmatrix} \mathbf{x}_{\theta_1}(t) & \mathbf{x}_{\theta_2}(t) & \dots & \mathbf{x}_{\theta_p}(t) \end{pmatrix} = \frac{\partial \mathbf{x}}{\partial \theta}$. This set of equations is obtained from the model equation (1) through differentiation of both sides with respect to θ , and interchanging the order of differentiation with respect to t and θ on the left-hand side. Since θ itself does not depend on t , the order of differentiation can be interchanged easily. Also, the parametric sensitivity dynamics derived from the model are not approximated at some nominal value $\bar{\theta}$ and so the structural relations obtained for the sensitivity dynamics are exact. Finally, when observing the initial condition (5) for the system of local parametric sensitivities, it holds trivially true that if the initial condition (2) does not depend on any parameters, then the initial condition (5) is zero.

To further address the identifiability property of the manifold $\Psi(\Theta)$ of input-output mappings, it is useful to also consider the parametric *output sensitivity* vector $\mathbf{y}_\theta(t)$ that relates the output sensitivities to the internal states, inputs, and parametric state sensitivities. It can easily be derived from (3) as:

$$\mathbf{y}_\theta(t) = \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \mathbf{x}_\theta(t) + \frac{\partial \mathbf{h}}{\partial \theta} \quad (6)$$

where the arguments of the functions have again been omitted for ease of notation. The system of parametric sensitivities (4)–(6) can be viewed as a set of p *decoupled* systems that do not interfere with one another as depicted in figure 1. Thus, associated with each individual parameter θ_i one has a parametric sensitivity system, consisting of an output $\mathbf{y}_{\theta_i}(t)$ and a parametric state sensitivity $\mathbf{x}_{\theta_i}(t)$ which is driven by the input-state behavior of the original system (1)–(2). This will be made more explicit in the next section.

At this point it is interesting to note that the study of parametric (output) sensitivities has received ample attention in the literature, e.g. [5, 7, 10, 13, 20], albeit from many different perspectives and never from a differential geometric perspective as presented here. Having introduced the parametric output and state sensitivities, the remaining part of the paper is now concerned with the following objective: To gain insight into local injectivity of the map $\Psi : \theta \mapsto \Sigma(\theta)$ by studying the joint input-output behavior of the systems (1)–(6). This map Ψ is studied in the following sections with the tools that are available in differential geometry and this will show to be quite rewarding.

III. THE USE OF PARAMETRIC SENSITIVITIES FOR IDENTIFIABILITY ANALYSIS

A. The accessibility Lie algebra and the observation space

For a non-linear system $\mathcal{M}(\theta)$ of the form studied in this paper, the *accessibility Lie algebra* \mathcal{A} (also called the *control Lie algebra* in the literature) is defined as the smallest Lie

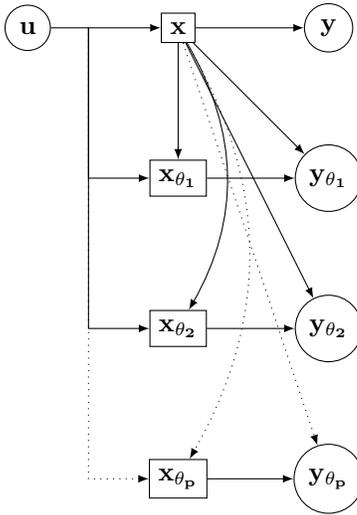


Fig. 1. Flow diagram summarizing the original system with I/O behavior $[u, y]$ and p sensitivity systems with I/O behavior $[u, y_{\theta}]$. Each of the sensitivity systems, together with the original system $[u, y]$, defines a so-called augmented system $\mathcal{M}_a^i(\theta)$ as explained in section III-B.

subalgebra of the space of smooth vector fields on the state manifold, which (i) contains the vector fields f and g , and (ii) is closed under the Lie bracket operation. The associated *controllability distribution* $\Delta_{\mathcal{A}}$ is defined as the mapping which assigns to every point x on the state manifold the linear subspace spanned by the vectors $\{v(x)|v \in \mathcal{A}\}$. The controllability rank condition (CRC) is said to hold at a point x if $\dim(\Delta_{\mathcal{A}}(x)) = n$, which means that an open subset of the state manifold can be reached from x under piecewise constant inputs. The observation space \mathcal{O} of the system is defined as the smallest subspace of smooth functions on the system manifold, which (i) contains the output function h , and (ii) is closed under Lie differentiation with respect to the vector fields in the accessibility Lie algebra \mathcal{A} .

We now note that when the functions in the observation space \mathcal{O} are evaluated at the initial state x_0 , they constitute quantities that show up as coefficients in Taylor series expansions of the output function $y(t)$ for certain particular choices of piecewise constant input functions $u(t)$. Therefore, these quantities are *invariants* of the system $\mathcal{M}(\theta)$ that may subsequently serve to characterize the associated input-output mapping $\Sigma(\theta)$. As a consequence, when a set of p such invariants turns out to have a nonsingular Jacobian (upon partial differentiation with respect to θ) at a given value θ_0 , then local injectivity of the mapping Ψ at θ_0 holds true, hence local identifiability.

In such an approach to identifiability analysis it is required to first compute \mathcal{A} and \mathcal{O} , and then to evaluate the functions that span \mathcal{O} at x_0 yielding invariants which are functions of θ . Finally, these invariants are differentiated with respect to θ , to evaluate local injectivity of the mapping Ψ . However, the computation of \mathcal{A} and \mathcal{O} also involves differentiation (as it involves Lie brackets and Lie differentiation with respect to the vector fields f and g). In the present paper it is proposed to *interchange the order of differentiation*

with respect to f and g on the one hand and with respect to θ on the other hand. Thus, we propose to proceed by constructing the sensitivity system first, and only then to compute the accessibility Lie algebra and the observation space of the sensitivity system. Finally, the functions spanning the observation space are evaluated at the augmented initial state. Depending on the values of p and n and on the exact nature and complexity of the model equations that represent $\mathcal{M}(\theta)$, such an approach may pay off in terms of computational effort involved.

B. Some additional definitions

Let us now define for each $i = 1, 2, \dots, p$ an augmented system $\mathcal{M}_a^i(\theta)$, which describes the parametric output sensitivity y_{θ_i} in terms of the state $x(t)$ of the model $\mathcal{M}(\theta)$ and the parametric state sensitivity $x_{\theta_i}(t)$ that appears in the i -th column of the equations (4)–(5). This augmented system $\mathcal{M}_a^i(\theta)$ has the state vector $x_a^i(t) \triangleq \begin{pmatrix} x(t) \\ x_{\theta_i}(t) \end{pmatrix}$ and is given by the following structure:

$$\begin{pmatrix} \dot{x}(t) \\ \dot{x}_{\theta_i}(t) \end{pmatrix} = \begin{pmatrix} f(x(t), \theta) + g(x(t), \theta) u(t) \\ \frac{\partial f}{\partial x} x_{\theta_i}(t) + \frac{\partial f}{\partial \theta_i} + \\ (\frac{\partial g}{\partial x} x_{\theta_i}(t) + \frac{\partial g}{\partial \theta_i}) u(t) \end{pmatrix} \quad (7)$$

$$x_a^i(0) = \begin{pmatrix} x_0(\theta) \\ \frac{\partial x_0(\theta)}{\partial \theta_i} \end{pmatrix} \quad (8)$$

$$y_{\theta_i}(t) = \frac{\partial h}{\partial x} x_{\theta_i}(t) + \frac{\partial h}{\partial \theta_i} \quad (9)$$

In addition, the following notation is introduced for the augmented *drift* vector field $f_a^i(x_a^i(t), \theta)$ and for the augmented *control* vector field $g_a^i(x_a^i(t), \theta)$, for the coupled dynamics of the state vector $x(t)$ and the parametric state sensitivity $x_{\theta_i}(t)$ on the righthand side of (7):

$$f_a^i(x_a^i(t), \theta) \triangleq \begin{pmatrix} f(x(t), \theta) \\ \frac{\partial f}{\partial x} x_{\theta_i}(t) + \frac{\partial f}{\partial \theta_i} \end{pmatrix} \quad (10)$$

$$g_a^i(x_a^i(t), \theta) \triangleq \begin{pmatrix} g(x(t), \theta) \\ \frac{\partial g}{\partial x} x_{\theta_i}(t) + \frac{\partial g}{\partial \theta_i} \end{pmatrix} \quad (11)$$

Likewise, the output function for the augmented system $\mathcal{M}_a^i(\theta)$ is conveniently introduced as:

$$h_a^i(x_a^i(t), \theta) \triangleq \frac{\partial h}{\partial x} x_{\theta_i}(t) + \frac{\partial h}{\partial \theta_i}. \quad (12)$$

C. Construction of the observation space of the system of parametric output sensitivities

To analyze local injectivity of the mapping Ψ , we proceed by computing the observation space of the complete system of parametric output sensitivities. This space is jointly generated by the observation spaces of the decoupled augmented systems $\mathcal{M}_a^i(\theta)$, $i = 1, 2, \dots, p$.

For each of the augmented systems $\mathcal{M}_a^i(\theta)$ we start by constructing its accessibility Lie algebra. This is the smallest Lie subalgebra of the space of smooth vector fields on the augmented state space which contains both the drift vector field f_a^i and the control vector field g_a^i and which is closed under the binary Lie bracket operator $[\cdot, \cdot]$. Recall that the

Lie bracket operator tests whether two vector fields \mathbf{f} and \mathbf{g} commute and is defined as

$$[\mathbf{f}, \mathbf{g}] = \mathcal{L}_{\mathbf{f}}\mathbf{g} - \mathcal{L}_{\mathbf{g}}\mathbf{f}$$

where $\mathcal{L}_{\mathbf{f}}\mathbf{g}$ denotes the Lie derivative of vector field \mathbf{g} along the vector field \mathbf{f} , [8, 9, 12]. It can be quite a laborious task to calculate the Lie derivatives, Lie brackets, and the accessible Lie algebra by hand. Hence, we have used a symbolic algebra package, i.e. Kwatny's ProPac add-on for Mathematica [9], to perform these computations with great ease.

Once the accessibility Lie algebra has been calculated, we proceed to calculate the observation space of each $\mathcal{M}_a^i(\theta)$. This is the smallest subspace of smooth functions on the augmented state space which contains the function \mathbf{h}_a^i and which is closed under Lie differentiation along the vector fields \mathbf{f}_a^i and \mathbf{g}_a^i (or equivalently, which is closed under Lie differentiation along the vector fields contained in the accessibility Lie algebra). Note that the (repeated) Lie derivatives of the functions \mathbf{h}_a^i considered here should be structured in the *same* way for each $i = 1, 2, \dots, p$, as it is our purpose to compute partial derivatives with respect to each of the parameters θ_i of the same input-output invariants of the system $\mathcal{M}(\theta)$.

We then construct a 'Jacobian matrix' Ξ of which the i -th column (corresponding to partial differentiation with respect to θ_i) is constructed as

$$\Xi_i(\mathbf{x}_a^{\theta_i}(0)) \triangleq (\mathbf{h}_a^i \quad \mathcal{L}_{\mathbf{v}_1}\mathbf{h}_a^i \quad \dots \quad \mathcal{L}_{\mathbf{v}_q}\mathbf{h}_a^i)^T \quad (13)$$

where the vector fields \mathbf{v}_k , $k = 1, 2, \dots, q$, are chosen from the accessibility Lie algebra of the system $\mathcal{M}_a^i(\theta)$ (structured in the same way for each $i = 1, 2, \dots, p$) that was calculated in the first step. The vector fields \mathbf{v}_k are chosen such that the observation space of the corresponding sensitivity system $\mathcal{M}_a^i(\theta)$ is spanned by them and we perform this construction for *each* of the parameters $\{\theta_i, i = 1, \dots, p\}$ in the parameter vector θ as to build up the complete matrix Ξ . The matrix Ξ then allows us to check for linear dependencies between columns or, in other words, linear correlations between parametric output sensitivities.

A test for local identifiability follows immediately from the injectivity requirement of the original mapping Ψ :

Corollary 1: A sufficient condition for the parameters in the system model $\mathcal{M}(\theta)$ to be locally identifiable is that the matrix Ξ has full rank p .

Note that it is our construction of the matrix Ξ that is 'special' here. Important is that this construction allows for an *alternative viewpoint* with respect to identifiability, much in line with the approach of [5] but now extended to a non-linear parameterized class of state-space models. Note also that if the matrix Ξ is *not* of full rank, then the system may still be identifiable; then further investigations are necessary to establish this. In the case of unidentifiability, the nature of the linear dependence between the columns of Ξ may often provide one with more detailed information on the actual source of unidentifiability and for instance allow one to compute a parameter transformation which pinpoints the parameter combinations that can or cannot be identified.

1) *A simple example:* We will now introduce a simple example that allows the reader to better understand the construction of the matrix Ξ as proposed in the previous section. Consider therefore the single-input-single-output linear system

$$\dot{x}(t) = (\theta_1 + \theta_2)x(t) + u(t) \quad (14)$$

$$x(0) = \theta_2 \quad (15)$$

$$y(t) = x(t) \quad (16)$$

The associated parameter sensitivity system is

$$\begin{pmatrix} \dot{x}_{\theta_1} & \dot{x}_{\theta_2} \end{pmatrix} = \begin{pmatrix} \theta_1 + \theta_2 & 0 \\ 0 & \theta_1 + \theta_2 \end{pmatrix} \begin{pmatrix} x_{\theta_1} & x_{\theta_2} \end{pmatrix} + \begin{pmatrix} x & x \end{pmatrix} \quad (17)$$

$$\begin{pmatrix} y_{\theta_1} & y_{\theta_2} \end{pmatrix} = \begin{pmatrix} x_{\theta_1} & x_{\theta_2} \end{pmatrix} \quad (18)$$

so that the augmented models $\mathcal{M}_a^1(\theta)$ and $\mathcal{M}_a^2(\theta)$ are identical, except(!) for their initial condition, which is $\begin{pmatrix} \theta_2 \\ 0 \end{pmatrix}$

for $\mathcal{M}_a^1(\theta)$ and $\begin{pmatrix} \theta_2 \\ 1 \end{pmatrix}$ for $\mathcal{M}_a^2(\theta)$. The augmented vector fields are easily determined from (17)–(18) as

$$\mathbf{f}_a^1 = \begin{pmatrix} (\theta_1 + \theta_2)x \\ (\theta_1 + \theta_2)x_{\theta_1} + x \end{pmatrix} \quad (19)$$

$$\mathbf{f}_a^2 = \begin{pmatrix} (\theta_1 + \theta_2)x \\ (\theta_1 + \theta_2)x_{\theta_2} + x \end{pmatrix} \quad (20)$$

$$\mathbf{g}_a^1 = \mathbf{g}_a^2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (21)$$

$$\mathbf{h}_a^1 = x_{\theta_1} \quad (22)$$

$$\mathbf{h}_a^2 = x_{\theta_2} \quad (23)$$

The accessibility Lie algebra for the augmented model $\mathcal{M}_a^1(\theta)$ is easily found from \mathbf{f}_a^1 , \mathbf{g}_a^1 to be spanned by

$$\{\mathbf{v}_1, \mathbf{v}_2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \theta_1 + \theta_2 \\ 1 \end{pmatrix} \right\} \quad (24)$$

Involutivity clearly holds because, for example,

$$[\mathbf{f}_a^1, [\mathbf{f}_a^1, \mathbf{g}_a^1]] = -(\theta_1 + \theta_2)^2 \mathbf{g}_a^1 + 2(\theta_1 + \theta_2) [\mathbf{f}_a^1, \mathbf{g}_a^1] \quad (25)$$

For the augmented model $\mathcal{M}_a^2(\theta)$ it is not difficult to observe that in this particular example the accessibility Lie algebra coincides with the one of $\mathcal{M}_a^1(\theta)$.

We then continue to construct the matrix $\Xi(x_0)$ column-by-column, by taking Lie derivatives in the directions \mathbf{v}_1 and \mathbf{v}_2 for each parameter. The result, which includes the 'zero-order' term \mathbf{h}_a^i , see (13), is given by

$$\Xi(\theta_2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \quad (26)$$

Since this matrix has rank 2, we conclude that the parameters can be identified, which should not surprise the reader since θ_2 is included in the initial condition and is assumed known whilst θ_1 can be easily identified from the output once θ_2 has been identified.

The method presented above is very well suited for implementation in a symbolic computer algebra package which makes the calculation of the matrix Ξ much easier.

The examples in section IV were also calculated using such a symbolic program. We will now turn our attention to the Rapid Oxygen Demand TOXicity (RODToX) case study that has been studied by Dochain and co-workers and was already mentioned in the introduction. We will further demonstrate the applicability of the method introduced here to this interesting case study in the field of water quality monitoring.

IV. THE RODTOX CASE STUDY

A. Introduction

Consider a biomass (with concentration X) in an aerated tank feeding on an injected substrate of wastewater ($S(0)\delta(t)$, with $\delta(t)$ the familiar Dirac operator). Assume that the concentration X is constant over the course of one experiment (in case of the RODTOX device this is a realistic assumption since the duration of one experiment, i.e. recovery to steady state after injection of a unit of wastewater, is usually some 30 minutes and this is a too short time interval for the biomass to be able to grow substantially – see also [22]). A dynamic description of the behaviour of the substrate concentration is the Michaelis-Menten or Monod model

$$\frac{dS}{dt} = -\frac{\mu_{max}X}{Y} \frac{S(t)}{K_m + S(t)} \quad (27)$$

$$S(0) = S_0 \quad (28)$$

where it should be noted that the initial substrate concentration S_0 is *unknown* and the other parameters to be determined from an identification experiment are $\{\mu_{max}, Y, K_m\}$. As a consequence of the injection of the substrate in the bioreactor the biomass will start consumption and this requires oxygen. The oxygen uptake rate (OUR_{ex}) of the biomass is defined as

$$OUR_{ex} = -(1 - Y) \frac{dS}{dt} \quad (29)$$

and is measured indirectly through a dissolved oxygen (or DO) sensor. In case of several substrate concentrations $\{S_i(t), i = 1, \dots, k\}$, the total oxygen uptake rate (OUR_{ex}) is simply defined by the sum of the individual contributions, i.e. $OUR_{ex} = \sum_{i=1}^k (1 - Y_i) \frac{dS_i}{dt}$. For the case of one substrate injection $S(0)\delta(t)$ only, Dochain *et al.* have applied the Taylor series expansion method of Pohjanpalo [15] to the sensor output equation $OUR_{ex}(t)$ to study the identifiability of the one-substrate model. It was demonstrated that the OUR_{ex} observations allow identification of the following parameter combinations [4]

$$\theta_1 = \frac{(1 - Y)\mu_{max}X}{Y} \quad (30)$$

$$\theta_2 = (1 - Y)S_0 \quad (31)$$

$$\theta_3 = (1 - Y)K_m \quad (32)$$

This follows from the observation that all coefficients in the Taylor series expansion for $OUR_{ex}(t)$ can be written as combinations of the set $\{\theta_i, i = 1, 2, 3\}$.

The question to be answered in the current paper is ‘What does our differential geometric method (based on parametric sensitivities) as presented in the above tell us about this case study and does it reach the same conclusions?’ We will answer this question in the next section.

B. Results of the Geometric Identification Method

We will now apply the geometric identification method to the RODTOX case study. Since there is no input signal but only an unknown initial condition we repeatedly calculate the Lie derivative of the output signal, i.e. the OUR equation, for each of the four parametric sensitivity systems associated with each of the parameters $\{S_0, \mu_{max}, K_m, Y\}$, yielding one column of the matrix $\Xi(S_0)$ for each of the parameters. If one calculates these columns with a depth of 4, thereby obtaining a 4×4 matrix $\Xi(S_0)$, one finds the following result where, again for clarity, each column in the matrix corresponds to a parameter in the set $\{S_0, \mu_{max}, K_m, Y\}$ respectively:

$$\begin{pmatrix} -\frac{K_m X(Y-1)\mu_{max}}{(K_m+S_0)^2 Y} & \dots \\ \frac{K_m(K_m-2S_0)X^2(Y-1)\mu_{max}^2}{(K_m+S_0)^4 Y^2} & \dots \\ -\frac{K_m(K_m^2-8S_0K_m+6S_0^2)X^3(Y-1)\mu_{max}^3}{(K_m+S_0)^6 Y^3} & \dots \\ \frac{K_m(K_m^3-22S_0K_m^2+58S_0^2K_m-24S_0^3)X^4(Y-1)\mu_{max}^4}{(K_m+S_0)^8 Y^4} & \dots \\ -\frac{S_0 X(Y-1)}{(K_m+S_0)Y} & \dots \\ \frac{2K_m S_0 X^2(Y-1)\mu_{max}}{(K_m+S_0)^3 Y^2} & \dots \\ -\frac{3K_m(K_m-2S_0)S_0 X^3(Y-1)\mu_{max}^2}{(K_m+S_0)^5 Y^3} & \dots \\ \frac{4K_m S_0(K_m^2-8S_0K_m+6S_0^2)X^4(Y-1)\mu_{max}^3}{(K_m+S_0)^7 Y^4} & \dots \\ \frac{S_0 X(Y-1)\mu_{max}}{(K_m+S_0)^2 Y} & \dots \\ -\frac{(2K_m-S_0)S_0 X^2(Y-1)\mu_{max}^2}{(K_m+S_0)^4 Y^2} & \dots \\ \frac{S_0(3K_m^2-10S_0K_m+2S_0^2)X^3(Y-1)\mu_{max}^3}{(K_m+S_0)^6 Y^3} & \dots \\ \frac{S_0(-4K_m^3+43S_0K_m^2-52S_0^2K_m+6S_0^3)X^4(Y-1)\mu_{max}^4}{(K_m+S_0)^8 Y^4} & \dots \\ -\frac{S_0 X \mu_{max}}{(K_m+S_0)Y^2} & \dots \\ -\frac{K_m S_0 X^2(Y-2)\mu_{max}^2}{(K_m+S_0)^3 Y^3} & \dots \\ \frac{K_m(K_m-2S_0)S_0 X^3(2Y-3)\mu_{max}^3}{(K_m+S_0)^5 Y^4} & \dots \\ -\frac{K_m S_0(K_m^2-8S_0K_m+6S_0^2)X^4(3Y-4)\mu_{max}^4}{(K_m+S_0)^7 Y^5} & \dots \end{pmatrix} \quad (33)$$

The above matrix has a determinant of zero which indicates that a full rank is not achieved so that the sufficient condition for identifiability is not satisfied. This result is completely in tune with earlier findings that resulted in the conclusion that only three out of the four parameters could be identified using the RODTOX device [4].

Further differentiation of the output, generating more rows in the matrix $\Xi(S_0)$, does not change the column dependencies in the matrix, i.e. all minors of size 4 are zero – even after generating six rows. To further investigate the dependencies of the columns of $\Xi(S_0)$ row reduction of the

4×4 matrix was performed and this yielded the following reduced matrix for $\Xi(S_0)$:

$$\begin{pmatrix} 1 & 0 & 0 & \frac{S_0}{-1+Y} \\ 0 & 1 & 0 & \frac{\mu_{max}}{Y(-1+Y)} \\ 0 & 0 & 1 & \frac{K_m}{-1+Y} \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (34)$$

Since the above matrix has rank 3, this result indicates again that only three parameter can be determined from the given experimental setup, assuming that additional rows will yield no further information and reduce to zero after row-reduction has been achieved. The reader should note the similarity of the parameter combinations appearing in the fourth column when compared to the original results (30)–(32) in [4]! If we calculate the nullspace of the matrix $\Xi(S_0)$ we find the vector $(\frac{S}{(Y-1)}, \frac{\mu_{max}}{Y(Y-1)}, \frac{K_m}{Y-1}, 1)$ and this vector can be used to find an explicit re-parametrization, as explained in [6]. Without going into the details of this calculation we suggest the following re-parametrization:

$$\tilde{S}(t) = (1 - Y)S(t) \quad (35)$$

$$\theta_1 = \frac{(1 - Y)\mu_{max}X}{Y} \quad (36)$$

$$\theta_2 = (1 - Y)K_m \quad (37)$$

the transformed model reads

$$\frac{d\tilde{S}}{dt} = \frac{\theta_1 \tilde{S}(t)}{\theta_2 + \tilde{S}(t)} \quad (38)$$

$$\tilde{S}(0) = \theta_3 \quad (39)$$

$$OUR_{ex}(t) = \frac{d\tilde{S}}{dt} \quad (40)$$

where, now, the parameter combinations $\{\theta_1, \theta_2, \theta_3\}$ are to be determined from the same input-output data. Construction of the matrix $\Xi(\theta_3)$ for the transformed model indeed showed that it has full rank for almost all values of θ_i and $Y \neq 1$ and so the identifiability condition is satisfied.

C. Adding an input to the RODTOX setup

One could further elaborate on the results obtained sofar and include an input signal in the experimental setup by continuously adding substrate to the bioreactor, i.e. a fedbatch experiment. In case the substrate concentration is high so that volume increase in the bioreactor can be neglected (meaning essentially that the biomass concentration X remains constant!) the modified model becomes

$$\frac{dS}{dt} = -\frac{\mu_{max}X}{Y} \frac{S(t)}{K_m + S(t)} + \dots + \frac{u(t)}{V} (S_{in} - S(t)) \quad (41)$$

$$S(0) = S_0 \quad (42)$$

where the input $u(t)$ is a (small) flow rate and the volume is assumed constant. An important final assumption, which may not be true in practice, is knowledge of the inlet concentration S_{in} of the substrate that is fed into the reactor. To circumvent this practical problem arising in the above

fedbatch setup one could think of, for example, applying an initial pulse of wastewater first (setting the *unknown* initial condition S_0), after which a known substrate with known concentration S_{in} is fed into the reactor with a flow rate $u(t)$.

If we apply are going to construct the matrix Ξ for this problem we immediately find that the introduction of the input $u(t)$ in the state equation introduces a feedthrough term in the output equation. Hence, we modify the system and include a *second state* $v(t)$ in the model with dynamics $\dot{v}(t) = u(t)$ and replace the input signal $u(t)$ in the first equation with $v(t)$ so that the model becomes affine in the input $v(t)$ and with no input in the output equation.

Turning back to the construction of the matrix $\Xi(S_0)$, for which now an input vector field $\mathbf{g}(\mathbf{x}, \theta) = (0, 1)^T$ is available, an important difference in comparison with the zero-input case in the previous section can be made. Indeed, having a control vector field available, the accessibility algebras constructed for each parameter in the set $\{S_0, \mu_{max}, K_m, Y\}$ can now become *involutive* meaning that *all* terms in the series expansion of the output parametric sensitivities can be generated by a finite set of vector fields. This important fact yields a stopping criterion for the method utilized here, i.e. if an involutive algebra has been found after calculation of, say, n rows of the matrix $\Xi(x_0)$, then the $(n + 1)$ th row can be generated through Lie-bracketing using the first n rows that span the distribution. In other words, after n rows have been generated there *may* be no need to search for new directions in the output parametric sensitivity space \mathcal{O} .

Calculation of the matrix $\Xi(S_0)$ for this input-driven setup (not presented here) indeed generated an involutive distribution after four rows were generated. Moreover, the determinant of the first four rows of the matrix $\Xi(S_0)$ was found to be

$$\frac{2S_{in}K_m^3(S_{in} - S_0)^2 S_0 X^6 (Y - 1)^3 \mu_{max}^5}{(K_m + S_0)^{11} V^4 Y^6} \quad (43)$$

and this indicates that the rank condition of corollary 1 is satisfied for most practical cases. Note also that certain parameter values (for example $Y = 1$) yield a vanishing determinant, and this may indicate identifiability problems. Finally, note also that if $S_{in} = S_0$ then, locally there is no input available anymore and we are back in the first situation where identifiability was shown not to hold true.

V. CONCLUSIONS

We have revisited the question of local identifiability of the parameters in non-linear system models from experimental data; a question that has already been highlighted from different angles in the identification literature. Our goal in the current study was to demonstrate the importance of the dynamic behavior of parametric output sensitivities in relation to the identifiability question and the applicability of non-linear control theoretical concepts to an augmented system that includes the dynamics of the output parametric

sensitivities as output channels. This has led to an alternative angle of attack towards the identifiability problem that more clearly sheds light on the role of parametric output sensitivities in the identification problem.

The recipe for construction of the matrix $\Xi(\mathbf{x}_0)$ as pointed out in section III has shown to be successful in reproducing earlier identification results for the RODTOX case study as summarized in section IV – note further that this is a case study that is especially interesting because of its unknown initial condition and its nonlinear readout or observation equation. A case like this has unfortunately not been discussed frequently in the literature on this subject (indeed, very often the observation equation is assumed linear) while the suggested approach taken here seems to handle the nonlinearities in the model and observation equation quite naturally. This example has also demonstrated, at least formally, that inclusion of an additional input signal in the model can change the identifiability properties of a given system – a result that confirms the intuitive fact that controlled excitation of a system yields in general more information on the values of the parameters in the system model than ‘just’ an initial condition.

Finally, it should be mentioned that the method developed and applied here addresses the question of theoretical identifiability rather than practical identifiability as discussed in, for example, [7]. Some more reflection on the exact construction of the accessible Lie algebras learns that, indeed, there may be directions in the parametric output sensitivity space \mathcal{O} (spanned by the column of $\Xi(\mathbf{x}(0))$) that are very hard to reach since complicated Lie brackets have to be calculated for these directions. The complexity of these repeated Lie brackets can perhaps not be realized in an experimental setup through manipulation of the input signal. Put in other words: Complex manipulations of the input signals to obtain directions in the parametric output sensitivity space that involve high order Lie brackets may be so convoluted that their consequences for the measurement of the output signals may be lost due to measurement noise. Further research could focus on finding case studies where this is demonstrated for a practical setup. This would yield more insight in the practical identifiability problem which, it should be stated repeatedly, is an important subject for each practitioner who collects real data through experimentation for calibration of his/her models.

REFERENCES

- [1] R. J. Bellman and K. J. Åström, “On structural identifiability,” *Mathematical Biosciences*, vol. 7, pp. 329–339, 1970.
- [2] L. Denis-Vidal and G. Joly-Blanchard, “Equivalence and identifiability analysis of uncontrolled nonlinear dynamical systems,” *Automatica*, vol. 40, pp. 287–292, 2004.
- [3] L. Denis-Vidal, G. Joly-Blanchard, and C. Noiret, “Some effective approaches to check the identifiability of uncontrolled nonlinear systems,” *Mathematics and Computers in Simulation*, vol. 57, pp. 35–44, 2001.
- [4] D. Dochain, P. A. Vanrolleghem, and M. van Daele, “Structural Identifiability of Biokinetic Models of Activated Sludge Respiration,” *Water Research*, vol. 29(11), pp. 2571–2578, 1995.
- [5] H. G. M. Dötsch and P. M. J. van den Hof, “Test for local structural identifiability of high-order non-linearly parametrized state space models,” *Automatica*, vol. 32, no. 6, pp. 875–883, 1996.
- [6] N. D. Evans and M. J. Chappell, “Extensions to a procedure for generating locally identifiable reparametrisations of unidentifiable systems,” *Mathematical Biosciences*, vol. 168, pp. 137–159, 2000.
- [7] A. Holmberg and J. Ranta, “Procedures for Parameter and State Estimation of Microbial Growth Process Models,” *Automatica*, vol. 18(2), pp. 181–193, 1982.
- [8] A. Isidori, *Nonlinear Control Systems*. Heidelberg: Springer-Verlag, 1989.
- [9] H. G. Kwatny and G. L. Blankenship, *Nonlinear Control and Analytical Mechanics: A Computational Approach*. Boston: Birkhäuser, 2000.
- [10] L. Ljung, *System Identification – Theory for the User*. Prentice Hall, 1987.
- [11] L. Ljung and T. Glad, “On global identifiability for arbitrary model parametrizations,” *Automatica*, vol. 30, no. 2, pp. 265–276, 1994.
- [12] H. Nijmeijer and A. Van der Schaft, *Nonlinear Dynamic Control Systems*. New York: Springer-Verlag, 1990.
- [13] R. L. M. Peeters and B. Hanzon, “Symbolic computation of fisher information matrices for parametrized state-space systems,” *Automatica*, vol. 35, pp. 1059–1071, 1999.
- [14] —, “Identifiability of homogeneous systems using the state isomorphism approach,” *Automatica*, vol. 41, pp. 513–529, 2005.
- [15] H. Pohjanpalo, “System Identifiability Based on the Power Series Expansion of the Solution,” *Mathematical Biosciences*, vol. 41, pp. 21–33, 1978.
- [16] M. P. Saccomani, “Some results on parameter identification of nonlinear systems,” *Cardiovascular Engineering: An International Journal*, vol. 4(1), pp. 95–102, 2004.
- [17] M. P. Saccomani, S. Audoly, and L. D’Angiò, “Parameter identifiability of nonlinear systems: The role of initial conditions,” *Automatica*, vol. 39, pp. 619–632, 2003.
- [18] J. Schoukens, R. Pintelon, and Y. Rolain, “Identification of nonlinear and linear systems, similarities, differences and challenges,” in *14th IFAC Symposium on System Identification (SYSID 2006)*, 2006, pp. 122–124.
- [19] A. Sedoglavic, “A probabilistic algorithm to test local algebraic observability in polynomial time,” *Journal of Symbolic Computation*, vol. 33, pp. 735–755, 2002.
- [20] J. D. Stigter and M. B. Beck, “On the development and application of a continuous-discrete recursive prediction-error algorithm,” *Mathematical Biosciences*, vol. 191, no. 2, pp. 143–158, 2004.
- [21] E. T. Tunali and T. Tarn, “New results for identifiability of nonlinear systems,” *IEEE Transactions on Automatic Control*, vol. AC-32, no. 2, pp. 146–154, 1987.
- [22] P. A. Vanrolleghem, M. V. Daele, and D. Dochain, “Practical identifiability of a biokinetic model of activated sludge respiration,” *Water Research*, vol. 29(11), pp. 2561–2570, 1995.
- [23] E. Walter and L. Pronzato, “On the identifiability and distinguishability of nonlinear parametric models,” *Mathematics and Computers in Simulation*, vol. 42, pp. 125–134, 1996.
- [24] —, *Identification of Parametric Models*. Springer, 1997.
- [25] X. Xia and C. H. Moog, “Identifiability of nonlinear systems with application to HIV/AIDS models,” *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 330–336, 2003.