UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner Bases

Nicolette Meshkat, Marisa Eisenberg, and Joseph J. DiStefano, III
November 2008
CAM Report 08-80

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90095-1555

An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner Bases

Nicolette Meshkat[1]*, Marisa Eisenberg[2], and Joseph J. DiStefano, III[3]

[1]UCLA Department of Mathematics, [2]UCLA Department of Biomedical Engineering, [3]UCLA Department of Computer Science

---

Abstract

The parameter identifiability problem for dynamic system ODE models has been extensively studied. Nevertheless, except for linear ODE models, the question of establishing *identifiable combinations* of parameters when the model is unidentifiable has not received as much attention and the problem is not fully resolved for nonlinear ODEs. We extend an existing algorithm for finding globally identifiable parameters of nonlinear ODE models to generate the 'simplest' globally identifiable parameter combinations using Gröbner bases. We also provide sufficient conditions for the method to work, demonstrate our algorithm on some models, and find the associated identifiable reparameterizations.

Key words: Identifiability, Differential Algebra, Gröbner Basis, Reparameterization

---

1. Introduction

Parameter identifiability analysis addresses the problem of which unknown parameters of an ODE model can be quantified from given input/output data. If all the parameters of the model have a unique or finitely many solutions, the model and its parameter vector **p** are said to be identifiable. Many models, however, yield infinitely many solutions for some parameters, and the model and its parameter vector **p** are then said to be unidentifiable. This raises the question, given an unidentifiable model, can we find combinations of the elements of **p** that are identifiable, e.g. so the model can be solved? Finding these parameter combinations is the main focus of this paper.

For linear ODE models, the problem of finding identifiable combinations in **p** when **p** is not identifiable is solved globally using transfer function and other linear algebra methods [1,4,10]. For nonlinear ODE models, the problem has been more challenging, with little resolution beyond application to simple models [19], most providing computationally-intensive local solutions. Evans and Chappell [9] and Gunn et al [11] adapt the Taylor series approach of Pohjanpalo [16] to find locally identifiable combinations. Chappell and Gunn [3] use the similarity transformation approach to generate identifiable reparameterizations, but again only locally. Denis-Vidal and Joly-Blanchard [6] find reparameterizations using equivalence of systems based on the straightening out theorem to get global identifiability.

*Corresponding author: Email address: nmeshkat@math.ucla.edu

Preprint submitted to Mathematical Biosciences *November 25, 2008*

However, for systems of dimension greater than one, this method does not find a necessary condition for identifiability and is not implemented as easily as other methods [6]. Denis-Vidal et al [5,7] and Verdiere et al [22] find globally identifiable combinations of parameters in a differential algebraic approach similar to Saccomani et al [17] , via an "inspection" method as discussed later.

In this work, we first establish a 'simplest' set (defined below) of globally identifiable parameter combinations for a practical class of nonlinear ODE models. To accomplish this, we extend the method of Bellu and coworkers [2] using a variation on the Gröbner basis approach and exemplify our algorithm and its application to reparameterization.

2. Nonlinear ODE Model

Our model is of the form:

$$\begin{aligned} \dot{\boldsymbol{x}}(t,\boldsymbol{p}) &= \boldsymbol{f}(\boldsymbol{x}(t,\boldsymbol{p}),\boldsymbol{u}(t),t;\ \boldsymbol{p}), t \in [t_0, T] \\ \boldsymbol{y}(t,\boldsymbol{p}) &= \boldsymbol{g}(\boldsymbol{x}(t,\boldsymbol{p});\boldsymbol{p}) \\ \boldsymbol{x_0} &= \boldsymbol{x}(t_0,\boldsymbol{p}) \end{aligned} \tag{2.1}$$

Here $\boldsymbol{x}$ is a n-dimensional state variable, $\boldsymbol{x_0}$ is the initial state at time t_0, $\boldsymbol{p}$ is a P-dimensional parameter vector, $\boldsymbol{u}$ is the r-dimensional input vector, and $\boldsymbol{y}$ is the m-dimensional output vector. As in [2], we assume $\boldsymbol{f}$ and $\boldsymbol{g}$ are rational polynomial functions of their arguments. Also, constraints reflecting known relationships among parameters, states, and/or inputs are assumed to be already included in (2.1), because they generally affect identifiability properties [8].

3. Identifiability

The question of *a priori structural identifiability* concerns finding one or more sets of solutions for the unknown parameters of a model from noise-free experimental data. Structural identifiability is a necessary condition for finding solutions to the real "noisy" data problem, often called the *numerical identifiability* problem.

Mathematically, it is sometimes convenient to express structural identifiability as an injectivity condition, as in [17]. Let $\boldsymbol{y} = \boldsymbol{\Phi}(\boldsymbol{p},\boldsymbol{u})$ be the input-output map determined from (2.1), by eliminating the state variable $\boldsymbol{x}$. Consider the equation $\boldsymbol{\Phi}(\boldsymbol{p},\boldsymbol{u}) = \boldsymbol{\Phi}(\boldsymbol{p}^*,\boldsymbol{u})$, where $\boldsymbol{p}^*$ is an arbitrary point in parameter space and $\boldsymbol{u}$ is the input function. Then one solution $\boldsymbol{p} = \boldsymbol{p}^*$ corresponds to global identifiability, finitely many distinct solutions for $\boldsymbol{p}$ to local identifiability, and infinitely many solutions for $\boldsymbol{p}$ to unidentifiability.

4. Differential Algebra Approach

A particularly productive approach to the identifiability problem for nonlinear ODE models is the differential algebra approach of Saccomani et al [17], following methods developed by Ljung and Glad [12] and Ollivier [14,15]. Their most recent contribution is the DAISY (Differential Algebra for

Identifiability of SYstems) program [2]. We summarize this approach here, as our work is an extension of their algorithm.

The first step is finding the input-output map in implicit form by reducing the model (2.1) via Ritt's pseudodivision algorithm [2]. The result is called the *characteristic set* [2]. The characteristic set is in general non-unique, but can be made unique by a suitable normalization [2].

Ritt's pseudodivision is summarized as follows. By combining the model equations and derivatives of the equations, we transform them symbolically into an equivalent system where one or more equations have the state variables eliminated. Essentially it amounts to finding a Gröbner basis for the model equations plus their derivatives, or in other words, performing successive substitutions to eliminate the state variables. An example of the algorithm can be found in [2].

The first m equations of the characteristic set, i.e. those independent of the state variables, are the *input-output relations*:

$$\boldsymbol{\Psi}(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{p}) = \boldsymbol{0} \tag{4.1}$$

For example, the ODE model

$$\dot{x} = kx + u$$

$$y = x/V$$

with the chosen ranking $\dot{x} > x > \dot{y} > y > u$ yields an input-output equation, $\Psi(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{p}) = 0$, of the form:

$$\Psi(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{p}) = V\dot{y} - kVy - u = 0$$

$\boldsymbol{\Psi}(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{p}) = 0$ are polynomial equations in the variables, $\boldsymbol{u}$, $\dot{\boldsymbol{u}}$, $\ddot{\boldsymbol{u}}$, ... , $\boldsymbol{y}$, $\dot{\boldsymbol{y}}$, $\ddot{\boldsymbol{y}}$, ... with rational coefficients in the parameter $\boldsymbol{p}$. That is, we can write $\Psi_{\mathrm{j}}(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{p}) = \sum_i c_i\,(\boldsymbol{p})\psi_i(\boldsymbol{u}, \boldsymbol{y})$, where $c_i(\boldsymbol{p})$ is a rational function in the parameter $\boldsymbol{p}$ and $\psi_i(\boldsymbol{u}, \boldsymbol{y})$ is a polynomial function in the variables $\boldsymbol{u}$, $\dot{\boldsymbol{u}}$, $\ddot{\boldsymbol{u}}$, ... , $\boldsymbol{y}$, $\dot{\boldsymbol{y}}$, $\ddot{\boldsymbol{y}}$, ..., etc. We assume the terms $\boldsymbol{u}$, $\dot{\boldsymbol{u}}$, $\ddot{\boldsymbol{u}}$, ... , $\boldsymbol{y}$, $\dot{\boldsymbol{y}}$, $\ddot{\boldsymbol{y}}$, ... are non-vanishing. Since $\psi_i(\boldsymbol{u}, \boldsymbol{y})$ are linearly independent, global identifiability becomes injectivity of the map $\boldsymbol{c}(\boldsymbol{p})$. That is, the model (2.1) is a priori globally identifiable if and only if $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ implies $\boldsymbol{p} = \boldsymbol{p}^*$ for arbitrary $\boldsymbol{p}^*$ [2]. The equations $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ are called the *exhaustive summary* [2]. These equations are then solved for the parameter vector $\boldsymbol{p}$ via the Buchberger Algorithm and elimination. The resulting M equations are displayed in the DAISY output as one of the three possible cases:

A.) unique solution, e.g. $p_i - p_i^* = 0$
B.) finite number of solutions, e.g. $p_i - p_i^* = 0 \; or \; p_i - p_j^* = 0$
C.) infinite number of solutions , e.g. $p_i = F(\boldsymbol{p}, \boldsymbol{p}^*)$

This is where DAISY terminates. Thus, in the case of unidentifiability (case (C.)), nothing more is explicitly stated about finding identifiable combinations. However, there are ways of finding identifiable combinations from the DAISY output.

One way is by the "inspection" method, which involves simple rearrangements of the coefficients of the input-output equations. There are at least M coefficients of the input-output relations, which are always identifiable. Thus, in our example, V and kV are identifiable, thus k is also identifiable. This process of "inspection" to find identifiable combinations can be done using the input-output relations in the DAISY output. However, one can imagine examples with sufficiently complicated input-output equations where the effectiveness of inspection breaks down. Furthermore, although we may find identifiable combinations directly from the input-output relations, we may not be able to find the 'simplest' identifiable combinations. Another way to find identifiable combinations is through the DAISY parameter solution. Let s be the number of free parameters, defined as the number of total parameters P minus the number of equations in the solution, M. In the case of unidentifiability, the DAISY parameter solution contains s free parameters, and thus the solution can sometimes be algebraically manipulated to find M=P-s identifiable combinations. However, this is not always possible, as in the two, three, and four compartment model examples below. Thus, the method employed in the DAISY program provides a test for identifiability of parameters, but it doesn't directly provide the simplest globally identifiable parameter combinations in unidentifiable ODE models, or their associated reparameterizations. Our algorithm extends the DAISY approach by finding such combinations.

5. Algorithm

The DAISY parameter solution is found by obtaining Gröbner Bases of the exhaustive summary via the Buchberger Algorithm and then using the properties of elimination to solve explicitly for the parameter $\boldsymbol{p}$. Our algorithm begins one step back and examines the Gröbner Bases themselves, before applying them to solve for the parameters.

From the exhaustive summary, $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$, we construct a Gröbner Basis in the form $\boldsymbol{G} = \{G_1(\boldsymbol{p},\boldsymbol{p}^*), \ldots, G_k(\boldsymbol{p},\boldsymbol{p}^*)\}$, where G_i is a polynomial function (here k ≥ P-s, depending on the ranking of parameters). In this process, we observe that additional information can be obtained from the Gröbner Basis. In particular, if in the case of unidentifiability we can obtain simplified elements of the Gröbner Basis of the form

$$q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*) \tag{5.1}$$

where $q_i(\boldsymbol{p})$ is a polynomial function of $\boldsymbol{p}$, then $q_i(\boldsymbol{p})$ is uniquely identifiable. In other words, $G_i(\boldsymbol{p},\boldsymbol{p}^*)$ is "decoupled" into a polynomial in $\boldsymbol{p}$ minus the same polynomial in $\boldsymbol{p}^*$.

Note that we may instead have elements scaled by an arbitrary polynomial function $\tilde{f}(\boldsymbol{p}^*)$,

$$\tilde{f}(\boldsymbol{p}^*)q_i(\boldsymbol{p}) - \tilde{f}(\boldsymbol{p}^*)q_i(\boldsymbol{p}^*)$$

whose solution reduces to the simplified form (5.1). For example, $p_1^* p_2 p_3 - p_1^* p_2^* p_3^*$ reduces to $p_2 p_3 - p_2^* p_3^*$.

There is no guarantee of finding elements of this form. However, even if the elements in the Gröbner Basis are not "decoupled" in this form, sometimes the element can be solved for the parameters in order to get an identifiable expression. For example, an element $p_2^* p_1 - p_1^* p_2$ implies that p_1/p_2 is identifiable. This is demonstrated in the rational coefficients example below.

As we will see shortly, when the degree of the polynomials in the Gröbner Bases is greater than the degree of the coefficients of the input-output equations, we examine *factors* of elements of the Gröbner Bases. In this case, $q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*)$ may occur as a factor of a Gröbner Basis element. For example, we may have $(p_2 p_4 - p_2^* p_4^*)(p_2 p_4 + p_2^* p_4^*)$ as an element.

Another observation is that the determination of additional expressions of the type (5.1) depend upon the choice of ranking of parameters when constructing the Gröbner basis.

Our algorithm, outlined as follows, combines the results of these observations.

Step 1: Search through all relevant rankings and determine identifiable combinations, i.e. elements of the Gröbner Bases that *can be simplified* to the decoupled form $q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*)$ when set to zero. For P parameters, we need P! rankings of the parameters. However, in most cases we can choose up to P cyclic permutations of some order of the parameters to generate enough Gröbner Bases. Group these identifiable elements in their decoupled form $q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*)$ together and call this set the *identifiable set*.

From the examples above, $p_2 p_3 - p_2^* p_3^*$, $\frac{p_1}{p_2} - \frac{p_1^*}{p_2^*}$, $p_2 p_4 - p_2^* p_4^*$, and $p_2 p_4 + p_2^* p_4^*$ could all be elements in the identifiable set. Note that the term $\frac{p_1}{p_2} - \frac{p_1^*}{p_2^*}$ is the decoupled form of the Gröbner Basis element $p_2^* p_1 - p_1^* p_2$. Since a Gröbner Basis is computed by eliminating parameters with the highest ranking first, we want each parameter to have a chance at the highest ranking, hence the need to try several rankings of parameters. This permits construction of simpler basis polynomials, involving as few parameters as possible, using the elimination properties of Gröbner Bases. These combinations may not all appear in a single Gröbner Basis, hence the need for several rankings of parameters.

Step 2: Select the M 'simplest' combinations from the identifiable set. By 'simplest', we mean the elements that have the lowest degree and the fewest terms (in $\boldsymbol{p}$). In practice, this is done by ranking the identifiable parameter combinations in the order of their degree multiplied by the number of terms.

This set may not be unique. Thus, it may be necessary to try several different sets of combinations before choosing an optimal one. Also, this set should contain at least one function of each parameter appearing in the coefficients of the input-output equations, or else Step 4 will fail.

Definition: We call the set of elements chosen the *canonical set*, i.e. the canonical set contains the simplest elements of the (decoupled) form $q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*)$, which arise either directly from elements of the Gröbner Bases, or from factors of elements of Gröbner Bases.

For example, $p_2 p_3 - p_2^* p_3^*$, $\frac{p_1}{p_2} - \frac{p_1^*}{p_2^*}$, $p_2 p_4 - p_2^* p_4^*$, and $p_2 p_4 + p_2^* p_4^*$ all have rank 1 (since the term $1/p_2$ is treated as a variable, thus of degree 1). If M=3, we pick the first three terms to be in our canonical set. Note that the first, second, and fourth terms could also be our canonical set. However, any other choice would leave out some parameter.

Step 3: Extract only the function of parameters $\boldsymbol{p}$.

Definition: This set of simplified elements, i.e. $q_i(\boldsymbol{p})$, is called the *simplified canonical set*.

In our example, p_2p_3, $\frac{p_1}{p_2}$, and p_2p_4 are in the simplified canonical set.

The parameter combinations in the simplified canonical set become our new parameters.

One should note that Steps 2 and 3 can be reversed, and in practice this is computationally faster.

Step 4: Attempt to reparameterize the input-output equations (4.1) in terms of the simplified canonical set.

Step 5: Test if the new parameters satisfy the injectivity condition of the model. If the chosen canonical set contains the simplest identifiable combinations, then identifiability results.

Step 6: If the new parameter combinations are found to be identifiable, we attempt to reparameterize our original system.

We now examine how and when this algorithm works. First we examine why Steps 1 and 2 work, i.e. why we can pick (at least) M identifiable combinations from the Gröbner Bases.

Proposition: Let **G** be the set of all P! Gröbner Bases, for all rankings. Then **G** contains at least M identifiable parameter combinations. In other words, we can always decouple at least M combinations from the Gröbner Bases, thus rendering these combinations identifiable.

Proof: We know there exists at least M=P-s identifiable combinations because there are at least M coefficients of the input-output equations, which are known to be identifiable. So there exists M identifiable combinations of the form $q_i(\boldsymbol{p})$, where $q_i(\boldsymbol{p})$ is a rational function in $\boldsymbol{p}$. Since the Gröbner Basis contains polynomial equations in $\boldsymbol{p}, \boldsymbol{p}^*$, we claim it must contain $q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*)$ as a factor of one of our terms. Assume, for a contradiction, that it does not. Since a Gröbner Basis is a solution to our exhaustive summary, this means that we do not have $q_i(\boldsymbol{p}) = q_i(\boldsymbol{p}^*)$ as a solution, which means that $q_i(\boldsymbol{p})$ is not identifiable, contradiction. Thus, each identifiable combination must appear as a factor in some Gröbner Basis, for some ranking. Since we always have the prescribed solution, $\boldsymbol{p} = \boldsymbol{p}^*$, then a decoupled element $q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*) = 0$ implies identifiability, and vice versa. Thus we can decouple at least M combinations from the Gröbner Bases, thus rendering these combinations identifiable. ■

This proposition shows we can always find at least M identifiable combinations from the Gröbner Bases. For our canonical set, we choose the simplest M identifiable combinations. We conjecture that we can always find M *algebraically independent* identifiable combinations from the Gröbner Bases.

Assumption: We further assume the canonical set is a set of algebraically independent elements (over R).

Algebraic independence means that no element can be written as an algebraic combination of the other elements over R. In practice, we can establish algebraic independence using polynomial division, for example using the PolynomialReduce function in Mathematica, or more generally using a Gröbner Basis.

Now we examine when Step 4 works. Let $c_i(p_1, \ldots, p_P)$, $i = 1, \ldots, l$, $l \geq M$, be the parameter-dependent coefficients of the input-output relations. The exhaustive summary (and therefore, the input-output coefficients) can always be written in terms its Gröbner Basis $\{G_1, \ldots, G_k\}$, in any rank ordering, by definition. In other words, we can always rewrite the coefficients $c_i(p_1, \ldots, p_P)$, $i = 1, \ldots, l$, in terms of $\{G_1, \ldots, G_k\}$. The problem is that the coefficients may not be combinations in the variables $\{G_1, \ldots, G_k\}$ alone (specifically, our ring $R(\boldsymbol{p})$ includes the parameters and real numbers, so we may have a reparameterization in terms of $\{G_1, \ldots, G_k\}$ over $R(\boldsymbol{p})$ but not over R). Thus part of the difficulty lies in choosing combinations so that we can reparameterize the coefficients only over the real numbers.

Let the canonical set have the form $\{q_1(p_1, \ldots, p_P) - q_1(p_1^*, \ldots, p_P^*), \ldots, q_M(p_1, \ldots, p_P) - q_M(p_1^*, \ldots, p_P^*)\}$. Let the combinations chosen, the simplified canonical set, be denoted $\{q_1(p_1, \ldots, p_P), \ldots, q_M(p_1, \ldots, p_P)\}$. Now we examine when one can reparameterize the coefficients in terms of the simplified canonical set. We use a variation of the method from Shannon and Sweedler [21]. Take the Gröbner Basis of the following set:

$$\{\hat{c}_1 - c_1(p_1, \ldots, p_P), \ldots, \hat{c}_l - c_l(p_1, \ldots, p_P), \hat{q}_1 - q_1(p_1, \ldots, p_P), \ldots, \hat{q}_M - q_M(p_1, \ldots, p_P)\} \qquad (5.2)$$

with the ranking $\{p_1, \ldots, p_P, \hat{q}_1, \ldots, \hat{q}_M, \hat{c}_1, \ldots, \hat{c}_l\}$. Here $\hat{c}_1, \ldots, \hat{c}_l$ and $\hat{q}_1, \ldots, \hat{q}_M$ are tag variables, i.e. variables introduced in order to eliminate other variables [21]. We denote this Gröbner Basis $\widehat{\boldsymbol{G}}$. Then we take the elements of $\widehat{\boldsymbol{G}}$ involving only $\hat{c}_1, \ldots, \hat{c}_l$ and $\hat{q}_1, \ldots, \hat{q}_M$, set them to zero and solve for $\hat{c}_1, \ldots, \hat{c}_l$. This gives a solution for the coefficients in terms of the new parameters. To construct a predicate that determines whether a given coefficient can be reparameterized, we do this one step at a time, i.e. include only one $\hat{c}_i - c_i(p_1, \ldots, p_P)$ expression in (5.2).

$$\{\hat{c}_i - c_i(p_1, \ldots, p_P), \hat{q}_1 - q_1(p_1, \ldots, p_P), \ldots, \hat{q}_M - q_M(p_1, \ldots, p_P)\} \qquad (5.3)$$

We find the Gröbner Basis of (5.3) for each $c_i(p_1, \ldots, p_P), 1 \leq i \leq l$, in order the get the entire solution for $\hat{c}_1, \ldots, \hat{c}_l$ as described above. We find the following necessary and sufficient conditions for a unique rational reparameterization.

Lemma 1: A unique rational reparameterization for a coefficient $\hat{c}_i = c_i(p_1, \ldots, p_P)$ in terms of the simplified canonical set exists if and only if the Gröbner Basis $\widehat{\boldsymbol{G}}$ contains a linear polynomial $f(\hat{q}_1, \ldots, \hat{q}_M) - g(\hat{q}_1, \ldots, \hat{q}_M)\hat{c}_i$ with no dependency on $p_1, \ldots, p_P$, possibly raised to a higher power.

Proof: → Assume that there exists a unique rational reparameterization for the coefficient $\hat{c}_i$. Then without loss of generality, $\hat{c}_i$ is of the form $\hat{c}_i = f(\hat{q}_1, \ldots, \hat{q}_M)/g(\hat{q}_1, \ldots, \hat{q}_M)$, where f and g are polynomials. Then $f(\hat{q}_1, \ldots, \hat{q}_M) - g(\hat{q}_1, \ldots, \hat{q}_M)\hat{c}_i = 0$. Now a Gröbner Basis of the set (5.3) would have the following order of terms: terms only in $\hat{c}_i$, followed by terms in $\hat{c}_i$ and $\hat{q}_j$, followed by terms in $\hat{c}_i$, $\hat{q}_j$, and p_k, where $1 \leq j \leq M, 1 \leq k \leq P$. Since $\hat{c}_i = c_i(p_1, \ldots, p_P)$, we will not have terms only in $\hat{c}_i$.

Thus our first elements of the Gröbner Basis will be terms in $\hat{c}_i$ and $\hat{q}_j$. Assume, by contradiction, that no such term exists in the Gröbner Basis. However, we know that $f(\hat{q}_1, \dots, \hat{q}_M) - g(\hat{q}_1, \dots, \hat{q}_M)\hat{c}_i = 0$. Thus, a Gröbner Basis would have to include a function only in $\hat{c}_i$ and $\hat{q}_j$ because this means that the p_k can be eliminated. Thus a term involving only $\hat{c}_i$ and $\hat{q}_j$, $1 \le j \le M$, does exist in our Gröbner Basis. The question is whether the term is precisely $f(\hat{q}_1, \dots, \hat{q}_M) - g(\hat{q}_1, \dots, \hat{q}_M)\hat{c}_i$. Let $h(\hat{c}_i, \widehat{\boldsymbol{q}})$ be the polynomial in the Gröbner Basis, where $\widehat{\boldsymbol{q}} = (\hat{q}_1, \dots, \hat{q}_M)$. If $h(\hat{c}_i, \widehat{\boldsymbol{q}})$ is linear in $\hat{c}_i$, then we are done. Otherwise, $h(\hat{c}_i, \widehat{\boldsymbol{q}})$ is of higher order in $\hat{c}_i$. Then there will be possibly multiple roots in $\hat{c}_i$. However, there is a unique rational reparameterization for the coefficient $\hat{c}_i$, so there cannot be multiple distinct roots. Likewise there cannot be an infinite number of solutions in $\hat{c}_i$, or else a $\hat{c}_j$ would have to appear as a free parameter. If there were no solutions in $\hat{c}_i$, then $\hat{c}_i$ would not appear in $h(\hat{c}_i, \widehat{\boldsymbol{q}})$, but we have already established that h is a function of both $\hat{c}_i$ and $\widehat{\boldsymbol{q}}$. Thus the only other possibility is that there are repeated roots, i.e. that $h(\hat{c}_i, \widehat{\boldsymbol{q}})$ is of the form $(f(\hat{q}_1, \dots, \hat{q}_M) - g(\hat{q}_1, \dots, \hat{q}_M)\hat{c}_i)^\alpha$, where α is a positive integer. Thus the Gröbner Basis contains a linear polynomial $f(\hat{q}_1, \dots, \hat{q}_M) - g(\hat{q}_1, \dots, \hat{q}_M)\hat{c}_i$ with no dependency on $p_1, \dots, p_P$, possibly raised to a higher power.

← Now assume that the Gröbner Basis contains a linear polynomial $f(\hat{q}_1, \dots, \hat{q}_M) - g(\hat{q}_1, \dots, \hat{q}_M)\hat{c}_i$, possibly raised to a higher power, with no dependency on $p_1, \dots, p_p$. Solving for $\hat{c}_i$, we get that $\hat{c}_i = f(\hat{q}_1, \dots, \hat{q}_M)/g(\hat{q}_1, \dots, \hat{q}_M)$. This is a rational reparameterization of $\hat{c}_i$. Assume this reparameterization is not unique, i.e. there exists another such polynomial $\tilde{h}(\hat{c}_i, \widehat{\boldsymbol{q}})$ in the Gröbner Basis. If $\tilde{h}(\hat{c}_i, \widehat{\boldsymbol{q}})$ is not a power of $f(\hat{q}_1, \dots, \hat{q}_M) - g(\hat{q}_1, \dots, \hat{q}_M)\hat{c}_i$, then there are other solutions for $\hat{c}_i$ appearing in other Gröbner Basis elements. However, this violates the form of a Gröbner Basis, for if there were another solution for $\hat{c}_i$, it must appear as a product with $f(\hat{q}_1, \dots, \hat{q}_M) - g(\hat{q}_1, \dots, \hat{q}_M)\hat{c}_i$. Thus the reparameterization is unique. ■

The lemma just proven is more of a mathematical description of what it means to reparameterize, and can be thought of as a test, not an a priori condition on whether a simplified canonical set allows reparameterization of the coefficients of the input-output equations.

Once the coefficients have been reparameterized, we can examine identifiability as in Step 5. Since we have decreased the number of parameters from P to P-s, local or global identifiability will result.

Lemma 2: Let s be the number of free parameters. Let $p_1, \dots, p_P$ be the parameters. If the coefficients can be rationally reparameterized in $P - s$ variables, then global or local identifiability results (one or finitely many solutions).

Proof: We have that the canonical set contains algebraically independent elements. Since we have mapped the coefficients from P algebraically independent parameters to P-s algebraically independent parameters, the free parameters have disappeared. We rule out the case of no solution because we always have the prescribed solution $(p_1^*, \dots, p_P^*)$. ■

These two lemmas lead us to the following theorem:

Theorem 1: Suppose we have a model described by (2.1), for which we determine the canonical set $\boldsymbol{Q}$ and the associated simplified canonical set $\boldsymbol{q}$ as described, such that $|\boldsymbol{Q}|$=P-s. If $\widehat{\boldsymbol{G}}$ contains a linear polynomial $f(\hat{q}_1, \dots, \hat{q}_M) - g(\hat{q}_1, \dots, \hat{q}_M)\hat{c}_i$ for each $c_i(p_1, \dots, p_P) \in \boldsymbol{c}(\boldsymbol{p})$, possibly raised to a higher power, then there exists a unique rational reparameterization of $\boldsymbol{c}(\boldsymbol{p})$ in terms of $\boldsymbol{q}$ and the simplified canonical set $\boldsymbol{q}$ is identifiable. Moreover, if the canonical set $\boldsymbol{Q}$ corresponds to entire elements from the Gröbner Bases, then global identifiability results. If elements of the canonical set came from factors of elements of Gröbner Bases, then local identifiability results.

Proof: Lemmas 1 and 2 give identifiability. The task is to show that we get a unique solution when only entire elements of the Gröbner Bases are used. The canonical set provides a solution set for $q_1(p_1, \dots, p_P), \dots, q_M(p_1, \dots, p_P)$. Moreover, each element in the canonical set is a linear expression in $q_1(p_1, \dots, p_P), \dots, q_M(p_1, \dots, p_P)$, by construction. Thus there is only one solution for $q_1(p_1, \dots, p_P), \dots, q_M(p_1, \dots, p_P)$, since if any other solution existed, it would have to appear as a Gröbner Basis element. In other words, our solution would have to appear as a factor in another Gröbner Basis term, which violates our assumption. Thus the reparameterized coefficients in the exhaustive summary must have only one solution for the $q_1(p_1, \dots, p_P), \dots, q_M(p_1, \dots, p_P)$. If we take a factor of a Gröbner Basis element in the canonical set, then we see that the $q_1(p_1, \dots, p_P), \dots, q_M(p_1, \dots, p_P)$ may have multiple roots and thus the reparameterized exhaustive summary should also contain multiple roots. ■

We now focus again on Step 4 of our algorithm, the reparameterization of the coefficients of the input-output equations by the simplified canonical set. We would like to examine the mathematical properties of a canonical set that permits this reparameterization. Since the canonical set was formed from the Gröbner Bases of the ideal generated by the exhaustive summary, it is natural to then examine the ideal generated by the canonical set. When the canonical set contains simplified decoupled elements (not factors) of the Gröbner Bases, then the ideal generated by the canonical set is the *same* as the ideal generated by the exhaustive summary. To simplify the notation, let $\boldsymbol{p} = (p_1, \dots, p_P)$, $\left(d_1(\boldsymbol{p}, \boldsymbol{p}^*), \dots, d_l(\boldsymbol{p}, \boldsymbol{p}^*)\right)$ be the exhaustive summary, and $\boldsymbol{Q} = \{Q_1(\boldsymbol{p}, \boldsymbol{p}^*), \dots, Q_M(\boldsymbol{p}, \boldsymbol{p}^*)\}$ be the canonical set.

Theorem 2: Let $\boldsymbol{Q} = \{Q_1(\boldsymbol{p}, \boldsymbol{p}^*), \dots, Q_M(\boldsymbol{p}, \boldsymbol{p}^*)\}$ be a canonical set that can rationally reparameterize coefficients $c_j(p_1, \dots, p_P), 1 \leq j \leq l$. Further assume that each element of $\boldsymbol{Q}$ is the decoupled form of an element from a Gröbner Basis G_k of the exhaustive summary.

Then, the Gröbner Basis of $\boldsymbol{Q}$ is the same as some Gröbner Basis G_k of the exhaustive summary for a given ranking. That is, their ideals are congruent:

$$\left(d_1(\boldsymbol{p}, \boldsymbol{p}^*), \dots, d_l(\boldsymbol{p}, \boldsymbol{p}^*)\right) = \left(Q_1(\boldsymbol{p}, \boldsymbol{p}^*), \dots, Q_M(\boldsymbol{p}, \boldsymbol{p}^*)\right)$$

Proof: Let C and B be the algebraic set of zeros of the exhaustive summary and the canonical set, respectively:

$$C = \{\boldsymbol{p} \mid d_j(\boldsymbol{p}, \boldsymbol{p}^*) = 0, 1 \leq j \leq l\},$$

$$B = \{\boldsymbol{p} \mid Q_i(\boldsymbol{p}, \boldsymbol{p}^*) = 0, 1 \leq i \leq M\},$$

Label each element of a Gröbner basis G_k with two subscripts as G_{ks}. It is clear that the algebraic set of each basis element G_{ks} contains C, since C is the intersection of algebraic sets of all basis vectors in a

Gröbner basis G_k. On the other hand, the canonical set $\boldsymbol{Q}$ contains a subset of elements from the Gröbner Bases of the exhaustive summary, so its algebraic set B is an intersection of sets containing C. Thus B contains C.

Now assume there is a root from B that C does not contain, call it $\tilde{\boldsymbol{p}} = (\tilde{p}_1, \dots, \tilde{p}_P)$. Then $Q_i(\tilde{\boldsymbol{p}}, \boldsymbol{p}^*) = 0$ for all $Q_i \in \boldsymbol{Q}$. Let $d_j(\boldsymbol{p}, \boldsymbol{p}^*)$ be the exhaustive summary. Since the coefficients can be reparameterized, we have $d_j(\tilde{\boldsymbol{p}}, \boldsymbol{p}^*) = d_j\left(Q_1(\tilde{\boldsymbol{p}}, \boldsymbol{p}^*), \dots, Q_i(\tilde{\boldsymbol{p}}, \boldsymbol{p}^*), \dots, Q_M(\tilde{\boldsymbol{p}}, \boldsymbol{p}^*)\right) = 0$ for $1 \le j \le l$, thus $\tilde{\boldsymbol{p}} = (\tilde{p}_1, \dots, \tilde{p}_P)$ is a root of the exhaustive summary, so $\tilde{\boldsymbol{p}}$ is in C. Thus C contains B.

Therefore, C=B and the Gröbner Basis of $\boldsymbol{Q}$ is the same as the Gröbner basis of the exhaustive summary. It then follows that the ideal spanned by the canonical set is equal to the ideal spanned by the exhaustive summary. ■

Note that reparameterization of the coefficients $c_j(p_1, \dots, p_P), 1 \le j \le l$, of the input-output equations by the simplified canonical set implies reparameterization by the canonical set, hence the use of the canonical set in Theorem 2. If instead, factors of Gröbner Basis elements are used, we have the following corollary.

Corollary: Let $\boldsymbol{Q} = \{Q_1(\boldsymbol{p}, \boldsymbol{p}^*), \dots, Q_M(\boldsymbol{p}, \boldsymbol{p}^*)\}$ be a canonical set that can rationally reparameterize coefficients $c_j(p_1, \dots, p_P), 1 \le j \le l$ such that some $Q_i \in \boldsymbol{Q}$ is a factor of an element in a Gröbner Basis G_k of the exhaustive summary.

Then the ideal generated by the exhaustive summary is contained in the ideal generated by the canonical set:

$$\left(d_1(\boldsymbol{p}, \boldsymbol{p}^*), \dots, d_l(\boldsymbol{p}, \boldsymbol{p}^*)\right) \subset \left(Q_1(\boldsymbol{p}, \boldsymbol{p}^*), \dots, Q_M(\boldsymbol{p}, \boldsymbol{p}^*)\right)$$

Proof: In this case, the algebraic set of zeros of the exhaustive summary C contains the algebraic set of zeros of the canonical set B. Thus, the ideal generated by the exhaustive summary is a subset of the ideal generated by the canonical set. ■

Thus the reparameterizability of the coefficients of the input-output equations (equivalently, the exhaustive summary) by the canonical set implies the ideal generated by the canonical set must be at least as big as the ideal generated by the exhaustive summary.

In summary, we have found sufficient conditions for the local or global identifiability of new parameter combinations (Theorem 1). In addition, we have found necessary conditions for the reparameterizability of the coefficients of the input-output equations in terms of the (simplified) canonical set (Theorem 2).

Since our algorithm is simply extending the functionality of DAISY, it can be used on the same class of problems. Thus it can be used for linear or nonlinear models. We have provided examples of the method for both compartmental models and other types as well.

6. Case Study: Classic Unidentifiable 2-Compartment Model

$$\dot{x}_1 = -(k_{01} + k_{21})x_1 + k_{12}x_2 + u$$

$$\dot{x}_2 = k_{21}x_1 - (k_{02} + k_{12})x_2$$

$$y = \frac{x_1}{v}$$

Definition:

x_1, x_2 state variables

u input

y output

k_{01} , k_{02} , k_{12} , k_{21} , v unknown parameters

As in DAISY, we perform Ritt's pseudodivision algorithm to get an equation purely in terms of input/output and parameters:

$$v\ddot{y} + (k_{01} + k_{21} + k_{12} + k_{02})v\dot{y} - (k_{12}k_{21} - (k_{12} + k_{02})(k_{01} + k_{21}))vy - (k_{12} + k_{02})u - \dot{u} = 0$$

Thus our coefficients are:

$$\begin{gathered} v \\ (k_{01} + k_{02} + k_{12} + k_{21})v \\ (k_{01}k_{02} + k_{01}k_{12} + k_{02}k_{21})v \\ k_{12} + k_{02} \end{gathered} \tag{6.1}$$

When we run DAISY, these coefficients are set equal to those with $\{k_{01}, k_{02}, k_{12}, k_{21}, v\}$ replaced by $\{8, 7, 13, 12, 3\}$ to get the exhaustive summary:

$$\begin{gathered} v - 3 = 0 \\ k_{01}v + k_{02}v + k_{12}v + k_{21}v - 120 = 0 \\ k_{01}k_{02}v + k_{01}k_{12}v + k_{02}k_{21}v - 732 = 0 \\ k_{02} + k_{12} - 20 = 0 \end{gathered} \tag{6.2}$$

These pseudo-randomly generated numbers are used instead of symbolic values in order to save computation time [2]. The Buchberger Algorithm yields an infinite number of solutions for the parameters, provided by DAISY in the form:

$$k_{21} = \frac{-156}{k_{02} - 20}$$

$$k_{01} = \frac{4(5k_{02} - 61)}{k_{02} - 20}$$

$$k_{12} = -k_{02} + 20$$

$$v = 3$$

Here only v is identifiable, so our system is unidentifiable. This is where DAISY terminates. We take this result a step further and find combinations of parameters that yield a unique solution. We do this by replacing $\{8 , 7 , 13 , 12 , 3\}$ with symbolic values $\{\alpha , \beta , \gamma , \delta , \epsilon\}$ in (6.2):

$$v - \epsilon$$

$$k_{01}v + k_{02}v + k_{12}v + k_{21}v - \alpha\epsilon - \beta\epsilon - \gamma\epsilon - \delta\epsilon$$

$$k_{01}k_{02}v + k_{01}k_{12}v + k_{02}k_{21}v - \alpha\beta\epsilon - \beta\gamma\epsilon - \alpha\delta\epsilon$$

$$k_{02} + k_{12} - \beta - \delta$$

We choose to use symbolic values instead of numerical values so that we can form the simplified canonical set without confusion. Then we find Gröbner Bases for the system using different rankings. To form a complete set of Gröbner Basis elements, we try the following rankings of parameters, found by shifting the ordering: $\{k_{01} , k_{02} , k_{12} , k_{21} , v\}$, $\{k_{02} , k_{12} , k_{21} , v , k_{01}\}$, … , $\{v , k_{01} , k_{02} , k_{12} , k_{21}\}$. They are:

$$\{v - \epsilon , -k_{12}k_{21}\epsilon + \gamma\delta\epsilon , k_{02} + k_{12} - \beta - \delta , k_{01}\epsilon + k_{21}\epsilon - \alpha\epsilon - \gamma\epsilon\}$$

$$\{v - \epsilon , k_{01}k_{12}\epsilon - k_{12}\alpha\epsilon - k_{12}\gamma\epsilon + \gamma\delta\epsilon , k_{01}\epsilon + k_{21}\epsilon - \alpha\epsilon - \gamma\epsilon , k_{02} + k_{12} - \beta - \delta\}$$

$$\{-k_{01}k_{02}\epsilon + k_{02}\alpha\epsilon + k_{01}\beta\epsilon - \alpha\beta\epsilon + k_{02}\gamma\epsilon - \beta\gamma\epsilon + k_{01}\delta\epsilon - \alpha\delta\epsilon , v - \epsilon , k_{02} + k_{12} - \beta - \delta , k_{01}\epsilon + k_{21}\epsilon - \alpha\epsilon - \gamma\epsilon\}$$

$$\{k_{02}k_{21}\epsilon - k_{21}\beta\epsilon - k_{21}\delta\epsilon + \gamma\delta\epsilon , k_{01}\epsilon + k_{21}\epsilon - \alpha\epsilon - \gamma\epsilon , v - \epsilon , k_{02} + k_{12} - \beta - \delta\}$$

$$\{-k_{12}k_{21}\epsilon + \gamma\delta\epsilon , k_{02} + k_{12} - \beta - \delta , k_{01}\epsilon + k_{21}\epsilon - \alpha\epsilon - \gamma\epsilon , v - \epsilon\}$$

We pick the elements that can be decoupled:

$$\{v - \epsilon , -k_{12}k_{21}\epsilon + \gamma\delta\epsilon , k_{02} + k_{12} - \beta - \delta , k_{01}\epsilon + k_{21}\epsilon - \alpha\epsilon - \gamma\epsilon\}$$

We then form the canonical set, which in this case contains as many elements as the identifiable set:

$$\{v - \epsilon , -k_{12}k_{21} + \gamma\delta , k_{02} + k_{12} - \beta - \delta , k_{01} + k_{21} - \alpha - \gamma\}$$

In general, we choose the M 'simplest' elements, i.e., with the lowest degree and fewest number of terms. Here M = P-s = 5-1 = 4.

Then the simplified canonical set is:

$$q_1 = v$$

$$q_2 = -k_{12}k_{21}$$

$$q_3 = k_{02} + k_{12}$$

$$q_4 = k_{01} + k_{21}$$

Now we find the Grobner Basis $\widehat{\boldsymbol{G}}$ for each coefficient and find the following reparameterization of (6.1):

$$q_1$$

$$q_1q_3 + q_1q_4$$

$$q_1q_2 + q_1q_3q_4$$

$$q_3$$

This confirms that our original input-output coefficients are spanned by the elements we chose.

We now test the injectivity condition, i.e. set these new coefficients equal to those with $\{q_1, q_2, q_3, q_4\}$ replaced with symbolic values $\{\theta, \mu, \pi, \rho\}$.

$$q_1 - \theta$$

$$q_1q_3 + q_1q_4 - \theta\pi - \theta\rho$$

$$q_1q_2 + q_1q_3q_4 - \theta\mu - \theta\pi\rho$$

$$q_3 - \pi$$

We solve the system via the Buchberger Algorithm and get a unique solution for $\{q_1, q_2, q_3, q_4\}$.

Thus, the uniquely identifiable combinations found are:

$$v$$

$$k_{12}k_{21}$$

$$k_{02} + k_{12}$$

$$k_{01} + k_{21}$$

Notice that these combinations could be obtained from a single Gröbner Basis alone. This is not true in general, as we see in the next example.

Now, we reparameterize our original system as:

$$\dot{x}_1 = -q_4x_1 + k_{12}x_2 + u$$

$$\dot{x}_2 = \frac{q_2}{k_{12}} x_1 - q_3 x_2$$

$$y = \frac{x_1}{q_1}$$

We see that k_{12} still appears in our system. One way to fix this is to introduce a new variable, $x_2^{'} = k_{12} x_2$, and our system has only uniquely identifiable parameters:

$$\dot{x}_1 = -q_4 x_1 + x_2^{'} + u$$

$$\dot{x}_2^{'} = q_2 x_1 - q_3 x_2^{'}$$

$$y = \frac{x_1}{q_1}$$

7. Examples

Example: 2-Compartment Nonlinear Model

The following example is taken from Saccomani et al [17].

$$\dot{x}_1 = -\left(k_{21} + \frac{V_M}{K_M + x_1}\right) x_1 + k_{12} x_2 + b_1 u$$

$$\dot{x}_2 = k_{21} x_1 - (k_{02} + k_{12}) x_2$$

$$y = c_1 x_1$$

$$x_1(0) = 0$$

$$x_2(0) = 0$$

Definition:

x_1, x_2 state variables

u input

y output

k_{02} , k_{12} , k_{21} , V_M , K_M , b_1 , c_1 unknown parameters

For a pseudo-random point of $\{k_{21}, k_{12}, V_M, K_M, k_{02}, c_1, b_1\} = \{8, 7, 13, 12, 3, 6, 13\}$, DAISY returns the following solution:

$$V_M = \frac{78}{c_1}$$

$$k_{21} = 8$$

$$k_{12} = 7$$

$$b_1 = \frac{78}{c_1}$$

$$k_{02} = 3$$

$$K_M = \frac{72}{c_1}$$

Here, k_{21}, k_{12}, k_{02}, are identifiable, while V_M, b_1, K_M, c_1 are unidentifiable. It is easy to see that an identifiable set of solutions is formed by cross-multiplying, to obtain $\{c_1V_M\ , k_{21}\ , k_{12}\ , b_1c_1\ , k_{02}\ , c_1K_M\}$.

The input-output equation is of the form:

$$\begin{aligned}-b_1c_1^3K_M^2\dot{u} - 2b_1c_1^2K_M\dot{u}y - b_1c_1\dot{u}y^2 + c_1^2K_M^2\ddot{y} + 2c_1K_M\ddot{y}y + \ddot{y}y^2 \\ + (c_1^2k_{02}K_M^2 + c_1^2k_{12}K_M^2 + c_1^2k_{21}K_M^2 + c_1^2K_MV_M)\dot{y} \\ + (2c_1k_{02}K_M + 2c_1k_{12}K_M + 2c_1k_{21}K_M)\dot{y}y + (k_{02} + k_{12} + k_{21})\dot{y}y^2 \\ - (b_1c_1^3k_{02}K_M^2 + b_1c_1^3k_{12}K_M^2)u - (2b_1c_1^2k_{02}K_M + 2b_1c_1^2k_{12}K_M)uy \\ - (b_1c_1k_{02} + b_1c_1k_{12})uy^2 + (c_1^2k_{02}k_{21}K_M^2 + c_1^2k_{02}K_MV_M + c_1^2k_{12}K_MV_M)y \\ + (2c_1k_{02}k_{21}K_M + c_1k_{02}V_M + c_1k_{12}V_M)y^2 + k_{02}k_{21}y^3 = 0\end{aligned}$$

We now form the exhaustive summary, find the Gröbner Bases in the 7 shifted orderings of $\{k_{21}, k_{12}, V_M, K_M, k_{02}, c_1, b_1\}$. We pick the simplest combinations, which are $\{q_1 = c_1V_M\ , q_2 = k_{21}\ , q_3 = k_{12}\ , q_4 = b_1c_1\ , q_5 = k_{02}\ , q_6 = c_1K_M\}$, reparameterize the input-output coefficients in terms of these, form the exhaustive summary, and solve to get a unique solution for $\{c_1V_M\ , k_{21}\ , k_{12}\ , b_1c_1\ , k_{02}\ , c_1K_M\}$. Thus we have found our simplified canonical set.

We then tried to see if the canonical set can be obtained from a single Gröbner Basis. We wrote a computer program to try all 7!=5040 permutations of the parameters (using numerical values for p* to speed up computation time [2]) and we found no single Gröbner Basis contained the canonical set. At most, a single Gröbner Basis had 4 of the 6 elements.

Now we reparameterize our original system. Let $x_1^{'} = c_1x_1$ and $x_2^{'} = c_1x_2$. Then our system becomes:

$$\dot{x}_1^{'} = -\left(q_2 + \frac{q_1}{q_6 + x_1^{'}}\right)x_1^{'} + q_3x_2^{'} + q_4u$$

$$\dot{x}_2^{'} = q_2x_1^{'} - (q_3 + q_5)x_2^{'}$$

$$y = x_1^{'}$$

$$x_1^{'}(0) = 0$$

$$x_2'(0) = 0$$

Example: 3-Compartment Model

$$\dot{x}_1 = k_{13}x_3 + k_{12}x_2 - (k_{21} + k_{31})x_1 + u$$

$$\dot{x}_2 = k_{21}x_1 - (k_{12} + k_{02})x_2$$

$$\dot{x}_3 = k_{31}x_1 - (k_{13} + k_{03})x_3$$

$$y = \frac{x_1}{v}$$

Definition:

x_1, x_2, x_3 state variables

u input

y output

$k_{12}, k_{21}, k_{13}, k_{31}, k_{02}, k_{03}, v$ unknown parameters

For a pseudo-random point of $\{k_{12}, k_{21}, k_{13}, k_{31}, k_{02}, k_{03}, v\} = \{8, 7, 13, 12, 3, 6, 13\}$, DAISY returns the following solutions:

$$k_{21} = \frac{19k_{03} - 153}{k_{03} - 11}$$

$$k_{31} = -\frac{56}{k_{03} - 11}$$

$$k_{12} = \frac{156(k_{03} - 11)}{19k_{03} - 153}$$

$$k_{13} = -k_{03} + 11$$

$$k_{02} = \frac{205k_{03} - 1191}{19k_{03} - 153}$$

$$v = 13$$

or

$$k_{21} = \frac{19k_{03} - 205}{k_{03} - 19}$$

$$k_{31} = -\frac{156}{k_{03} - 19}$$

$$k_{12} = \frac{56(k_{03} - 19)}{19k_{03} - 205}$$

$$k_{13} = -k_{03} + 19$$

$$k_{02} = \frac{3(51k_{03} - 397)}{19k_{03} - 205}$$

$$v = 13$$

Here only v is identifiable. The input-output equation is of the form:

$$\begin{aligned}(k_{02}k_{03} + k_{03}k_{12} + k_{02}k_{13} + k_{12}k_{13})u + (k_{02} + k_{03} + k_{12} + k_{13})\dot{u} + \ddot{u} \\ - (k_{02}k_{03}k_{21} + k_{02}k_{13}k_{21} + k_{02}k_{03}k_{31} + k_{03}k_{12}k_{31})vy \\ - (k_{02}k_{03} + k_{03}k_{12} + k_{02}k_{13} + k_{12}k_{13} + k_{02}k_{21} + k_{03}k_{21} + k_{13}k_{21} + k_{02}k_{31} \\ + k_{03}k_{31} + k_{12}k_{31})v\dot{y} - (k_{02} + k_{03} + k_{12} + k_{13} + k_{21} + k_{31})v\ddot{y} - v\dddot{y} = 0\end{aligned}$$

We now form the exhaustive summary and find the Gröbner Bases in the 7 shifted orderings of $\{k_{12}, k_{21}, k_{13}, k_{31}, k_{02}, k_{03}, v\}$. Due to the large number of parameters, we choose numerical values in the exhaustive summary instead of symbolic values to save computation time. In this case, most of our Gröbner Basis elements are quadratic, but our coefficients of the input-output equation are multilinear, thus we take *factors* of the Gröbner Basis elements. We pick the simplest combinations, which are $\{q_1 = v,\ q_2 = k_{12}k_{21},\ q_3 = k_{13}k_{31},\ q_4 = k_{02} + k_{12},\ q_5 = k_{03} + k_{13},\ q_6 = k_{21} + k_{31}\}$, reparameterize the input-output coefficients in terms of these, form the exhaustive summary, and solve to get two distinct solutions. This is due to the symmetry of the problem. In particular, only v and $k_{21} + k_{31}$ are globally identifiable.

Note that the inspection method gives us the following globally identifiable parameter combinations:

$$v$$

$$k_{02}k_{03} + k_{03}k_{12} + k_{02}k_{13} + k_{12}k_{13}$$

$$k_{02} + k_{03} + k_{12} + k_{13}$$

$$k_{02}k_{03}k_{21} + k_{02}k_{13}k_{21} + k_{02}k_{03}k_{31} + k_{03}k_{12}k_{31}$$

$$k_{02}k_{21} + k_{03}k_{21} + k_{13}k_{21} + k_{02}k_{31} + k_{03}k_{31} + k_{12}k_{31}$$

$$k_{21} + k_{31}$$

Since these are complicated expressions, we cannot reparameterize our original equations over them. However, one could reparameterize using the companion matrix form with these parameter combinations. We will see this done in the following example.

Now we reparameterize our original system using our simplified canonical set. Let $x_2' = k_{12}x_2$ and $x_3' = k_{13}x_3$. Then our original system becomes:

$$\dot{x}_1 = x_3' + x_2' - q_6 x_1 + u$$

$$\dot{x}_2' = q_2 x_1 - q_4 x_2'$$

$$\dot{x}_3' = q_3 x_1 - q_5 x_3'$$

$$y = \frac{x_1}{q_1}$$

Example: 4-Compartment Model

The following example came from Evans and Chappell [9]. It describes the pharmacokinetics of bromosulphthalein. They find a reparameterization to make the model structurally locally identifiable [9]. We solve a related problem, with an input rather than an initial condition, and attempt to make the model structurally globally identifiable by a reparameterization.

$$\dot{x}_1 = -a_{31} x_1 + a_{13} x_3 + u$$

$$\dot{x}_2 = -a_{42} x_2 + a_{24} x_4$$

$$\dot{x}_3 = a_{31} x_1 - (a_{03} + a_{13} + a_{43}) x_3$$

$$\dot{x}_4 = a_{42} x_2 + a_{43} x_3 - (a_{04} + a_{24}) x_4$$

$$y_1 = x_1$$

$$y_2 = x_2$$

Definition:

x_1, x_2, x_3, x_4 state variables

u input

y_1, y_2 output

$a_{03}, a_{04}, a_{13}, a_{24}, a_{31}, a_{42}, a_{43}$ unknown parameters

For a pseudo-random point of $\{a_{03}, a_{04}, a_{13}, a_{24}, a_{31}, a_{42}, a_{43}\} = \{8, 7, 13, 12, 3, 6, 13\}$, DAISY returns the following solutions:

$$a_{04} = \frac{\pm\sqrt{457a_{43}^2 - 7800a_{43} + 24336} + 25a_{43} - 156}{2a_{43}}$$

$$a_{42} = \frac{\mp\sqrt{457a_{43}^2 - 7800a_{43} + 24336} + 25a_{43} - 156}{2a_{43}}$$

$$a_{03} = -a_{43} + 21$$

$$a_{24} = \frac{156}{a_{43}}$$

$$a_{31} = 3$$

$$a_{13} = 13$$

Here only a_{31} and a_{13} are identifiable. The input-output equations are of the form:

$$a_{13}\ddot{y}_2 + (a_{42}a_{13} + a_{13}a_{04} + a_{13}a_{24})\dot{y}_2 + a_{13}a_{42}a_{04}y_2 - a_{43}a_{24}\dot{y}_1 - a_{43}a_{31}a_{24}y_1 + a_{24}a_{43}u = 0$$

$$\ddot{y}_1 + (a_{31} + a_{03} + a_{13} + a_{43})\dot{y}_1 + (a_{31}a_{03} + a_{31}a_{43})y_1 - \dot{u} + (a_{03} + a_{13} + a_{43})u = 0$$

We now form the exhaustive summary and find the Gröbner Bases in the 7 shifted orderings of $\{a_{03}, a_{04}, a_{13}, a_{24}, a_{31}, a_{42}, a_{43}\}$. We pick the simplest combinations, which are $\{q_1 = a_{13}, q_2 = a_{31},\ q_3 = a_{04}a_{42},\ q_4 = a_{24}a_{43},\ q_5 = a_{03} + a_{43},\ q_6 = a_{04} + a_{24} + a_{42}\}$, reparameterize the input-output coefficients in terms of these, form the exhaustive summary, and solve to get a unique solution for $\{a_{13}, a_{31},\ a_{04}a_{42},\ a_{24}a_{43},\ a_{03} + a_{43},\ a_{04} + a_{24} + a_{42}\}$. Thus we have found our simplified canonical set, which agrees with the identifiable combinations found in [9]. However, our method guarentees the global identifiability of these parameter combinations, while Evans and Chappell can only show local identifiability using their Taylor Series approach [9].

Even though the input-output equations can be rationally reparameterized, the original equations cannot be rationally reparameterized. The reparameterization involves the square root function and can be found in Evans and Chappell [9]. Thus we see that input-output reparameterization is only a necessary condition for the original system to be reparameterized.

Alternatively, we can always reparameterize our model using the normal canonical form (companion matrix). Let $y_1 = v_1, \dot{y}_1 = \dot{v}_1 = v_2, y_2 = v_3, \dot{y}_2 = \dot{v}_3 = v_4, u = u_1, \dot{u}_1 = u_2$. Then the input-output equations become:

$$\dot{v}_1 = v_2$$

$$\dot{v}_2 = -(q_1 + q_2 + q_5)v_2 - q_2q_5v_1 + u_2 - (q_1 + q_5)u_1$$

$$\dot{v}_3 = v_4$$

$$q_1\dot{v}_4 = -q_1q_6v_4 - q_1q_3v_3 + q_4v_2 + q_2q_4v_1 - q_4u_1$$

$$y_1 = v_1$$

$$y_2 = v_3$$

Example: SIR (Susceptible Infected Recovered) model

$$\dot{S} = \mu N - S\left(\mu + \frac{\beta}{N} I\right)$$

$$\dot{I} = I\left(\frac{\beta}{N} S - (\mu + \nu)\right)$$

$$y = kI$$

Definition:

S, I state variables

y output

μ, ν, β, N, k unknown parameters

For a pseudo-random point of $\{\mu, \nu, \beta, N, k\} = \{8, 7, 13, 12, 3\}$, DAISY returns the following solution:

$$k = \frac{36}{N}$$

$$\nu = 7$$

$$\beta = 13$$

$$\mu = 8$$

Here, ν, β, μ are identifiable, while k and N are unidentifiable. It is easy to see that an identifiable set of solutions is formed by cross-multiplying, to obtain $\{kN, \nu, \beta, \mu\}$.

The input-output equation is of the form:

$$(-\beta k N \mu + k N \mu^2 + k N \mu \nu) y^2 + (\beta \mu + \beta \nu) y^3 + k N \mu y \dot{y} + \beta y^2 \dot{y} - k N \dot{y}^2 + k N y \ddot{y} = 0$$

We now form the exhaustive summary and find the Gröbner Bases in the 5 shifted orderings of $\{\mu, \nu, \beta, N, k\}$. We pick the simplest combinations, which are $\{q_1 = kN, q_2 = \nu, q_3 = \beta, q_4 = \mu\}$, reparameterize the input-output coefficients in terms of these, form the exhaustive summary, and solve to get a unique solution for $\{kN, \nu, \beta, \mu\}$. Thus we have found our simplified canonical set.

Now we reparameterize our original equations. Let $I' = kI$ and $S' = kS$. Then our system becomes:

$$\dot{S}' = q_1 q_4 - S'\left(q_4 + \frac{q_3}{q_1} I'\right)$$

$$\dot{I}' = I'\left(\frac{q_3}{q_1} S' - (q_2 + q_4)\right)$$

$$y = I'$$

Example: Unknown Initial Conditions

The following is taken from Saccomani et al [20].

$$\dot{x}_1 = p_1 u x_3$$

$$\dot{x}_2 = p_2 x_1$$

$$\dot{x}_3 = p_3 x_1 x_2$$

$$y_1 = x_1$$

$$y_2 = x_2$$

$$x_1(0) = x_{10}$$

$$x_2(0) = x_{20}$$

$$x_3(0) = x_{30}$$

Definition:

x_1, x_2, x_3 state variables

u input

y_1, y_2 output

p_1, p_2, p_3 unknown parameters

Here we have unknown initial conditions. This example illustrates how our algorithm can take advantage of the full extent of the DAISY program, e.g. incorporating initial conditions.

For a pseudo-random point of $\{p_1, p_2, p_3\} = \{8, 7, 13\}$, DAISY returns the following solution:

$$p_1 = \frac{104}{p_3}$$

$$p_2 = 7$$

Here only p_2 is identifiable. It is easy to see that an identifiable set of solutions is formed by cross-multiplying, to obtain $\{p_1 p_3, p_2\}$.

The input-output equations are:

$$\dot{y}_1 \dot{u}(-\dot{u}\dot{y}_1 + 2u\ddot{y}_1) + 2p_2 p_1 p_3 \dot{y}_1 u^3 y_1^2 + u^4 y_1^2 p_1^2 p_3(-2p_2 x_{30} + p_3 x_{20}^2) - \ddot{y}_1^2 u^2 = 0$$

$$\dot{u}\dot{y}_1 - \ddot{y}_1 u + u^2 y_1 y_2 p_1 p_3 = 0$$

We now form the exhaustive summary and find the Gröbner Bases in the 3 shifted orderings of $\{p_1, p_2, p_3\}$. Note in this case the only coefficients used are $2p_2p_1p_3$ and p_1p_3, since $p_1^2p_3(-2p_2x_{30} +$ $p3x202$ cannot be used to test identifiability since $\boldsymbol{x(0)}$ is not known [20]. We pick the simplest combinations, which are $\{q_1 = p_1p_3, q_2 = p_2\}$, reparameterize the input-output coefficients in terms of these (only the coefficients not involving $\boldsymbol{x}(0)$), form the exhaustive summary, and solve to get a unique solution for $\{p_1p_3, p_2\}$. Thus we have found our simplified canonical set.

Now we reparameterize our original equations. Let $x_3^{'} = p_1x_3$. Then our system becomes:

$$\dot{x}_1 = ux_3^{'}$$

$$\dot{x}_2 = q_2x_1$$

$$\dot{x}_3^{'} = q_1x_1x_2$$

$$y_1 = x_1$$

$$y_2 = x_2$$

$$x_1(0) = x_{10}$$

$$x_2(0) = x_{20}$$

$$x_3^{'}(0) = p_1x_{30}$$

Example: Rational coefficients

The following is taken from Margaria et al [13].

$$\dot{x}_1 = p_1x_1 - p_2x_1x_2$$

$$\dot{x}_2 = p_3x_2(1 - p_4x_2) + p_5x_1x_2$$

$$y_1 = x_1$$

Definition:

x_1, x_2 state variables

y_1 output

p_1, p_2, p_3, p_4, p_5 unknown parameters

For a pseudo-random point of $\{p_1, p_2, p_3, p_4, p_5\} = \{8, 7, 13, 12, 3\}$, DAISY returns the following solution:

$$p_1 = 8$$

$$p_3 = 13$$

$$p_4 = 12p_2/7$$

$$p_5 = 3$$

Here p_1, p_3, p_5 are identifiable while p_2, p_4 are unidentifiable. It is easy to see that an identifiable set of solutions is formed by cross-multiplying, to obtain $\{p_1, p_3, p_5, p_4/p_2\}$.

The input-output equation is of the form:

$$p_2 y\ddot{y} + (-p_2 + p_3 p_4)\dot{y}^2 + (-2p_1 p_3 p_4 - p_2 p_3)\dot{y}y - p_2 p_5 \dot{y}y^2 + (p_1^2 p_3 p_4 + p_1 p_2 p_3)y^2 + p_1 p_2 p_5 y^3 = 0$$

As discussed in [2], the input-output equation is made monic by dividing by p_2. Now the resulting coefficients are thus rational. We now form the exhaustive summary and find the Gröbner Bases in the 5 shifted orderings of $\{p_1, p_2, p_3, p_4, p_5\}$. We pick the simplest combinations, which are $\{q_1 = p_1, q_2 = p_3, q_3 = p_5, q_4 = p_4/p_2\}$, reparameterize the input-output coefficients in terms of these, form the exhaustive summary, and solve to get a unique solution for $\{p_1, p_3, p_5, p_4/p_2\}$. Thus we have found our simplified canonical set. Thus we have a case where we can decouple a Gröbner Basis term to get $q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*)$ where $q_i(\boldsymbol{p})$ is rational.

Now we reparameterize our original system. Let $x_2^{'} = p_2 x_2$. Then our original system becomes:

$$\dot{x}_1 = q_1 x_1 - x_1 x_2^{'}$$

$$\dot{x}_2^{'} = q_2 x_2^{'}(1 - q_4 x_2^{'}) + q_3 x_1 x_2^{'}$$

$$y_1 = x_1$$

Example: HIV/AIDS Model

The following four-dimensional HIV/AIDS model is taken from Saccomani and Bellu [18].

$$\dot{x}_1 = -\beta x_1 x_4 - d x_1 + s$$

$$\dot{x}_2 = \beta q_1 x_1 x_4 - k_1 x_2 - \mu_1 x_2$$

$$\dot{x}_3 = \beta q_2 x_1 x_4 + k_1 x_2 - \mu_2 x_3$$

$$\dot{x}_4 = -c x_4 + k_2 x_3$$

$$y_1 = x_1$$

$$y_2 = x_4$$

Definition:

x_1, x_2, x_3, x_4 state variables

y_1, y_2 output

$\beta, d, s, q_1, k_1, \mu_1, q_2, k_2, \mu_2, c$ unknown parameters

For a pseudo-random point of $\{\beta, d, s, q_1, k_1, \mu_1, q_2, k_2, \mu_2, c\} = \{8, 7, 13,12,3,6,13,16,9,10\}$, DAISY returns the following solutions:

$$q_1 = \frac{576}{k_1 k_2}, d = 7, s = 13, \mu_1 = -k_1 + 9, c = 9, \mu_2 = 10, q_2 = \frac{208}{k_2}, \beta = 8$$

$$q_1 = \frac{368}{k_1 k_2}, d = 7, s = 13, \mu_1 = -k_1 + 10, c = 9, \mu_2 = 9, q_2 = \frac{208}{k_2}, \beta = 8$$

$$q_1 = \frac{576}{k_1 k_2}, d = 7, s = 13, \mu_1 = -k_1 + 9, c = 10, \mu_2 = 9, q_2 = \frac{208}{k_2}, \beta = 8$$

The model is unidentifiable with only β, d, s globally identifiable and c, μ_2 locally identifiable. After adding initial conditions to the model, Saccomani and Bellu obtain that β, d, s are globally identifiable and all the other parameters are locally identifiable.

It is clear that by moving all the parameters to one side of the equations, we have that $q_1 k_1 k_2$ and $\mu_1 + k_1$ are locally identifiable parameter combinations and $q_2 k_2$ is a globally identifiable parameter combination. However, there is no mention of this in [18].

The input-output equations are of the form:

$$\dot{y}_1 + \beta y_1 y_2 + d y_1 - s = 0$$

$$\dddot{y}_2 + (c + k_1 + \mu_1 + \mu_2)\ddot{y}_2 - \beta q_2 k_2 \dot{y}_2 y_1 + (ck_1 + c\mu_1 + c\mu_2 + k_1\mu_2 + \mu_1\mu_2)\dot{y}_2 + \beta^2 q_2 k_2 y_1 y_2^2 \\ + \beta k_2 (d q_2 - k_1 q_1 - k_1 q_2 - \mu_1 q_2) y_1 y_2 + (-\beta q_2 k_2 s + c k_1 \mu_2 + c\mu_1\mu_2) y_2 = 0$$

We now form the exhaustive summary and find the Gröbner Bases in the 10 shifted orderings of $\{\beta, d, s, q_1, k_1, \mu_1, q_2, k_2, \mu_2, c\}$. We pick the simplest combinations, which are $\{q_1 k_1 k_2,\ d,\ s,\ \mu_1 + k_1,\ c,\ \mu_2,\ q_2 k_2,\ \beta\}$, reparameterize the input-output coefficients in terms of these, form the exhaustive summary, and solve to get a finite number of solutions for $\{q_1 k_1 k_2,\ d,\ s,\ \mu_1 + k_1,\ c,\ \mu_2,\ q_2 k_2,\ \beta\}$. Thus we have found our simplified canonical set. Note that we found local identifiability *without* using initial conditions. In particular, we find that $\{d,\ s,\ q_2 k_2,\ \beta\}$ are globally identifiable.

Note that the inspection method gives us the following globally identifiable parameter combinations:

$$\beta, d, s$$

$$q_2 k_2$$

$$c + k_1 + \mu_1 + \mu_2$$

$$ck_1 + c\mu_1 + c\mu_2 + k_1\mu_2 + \mu_1\mu_2$$

$$-k_1 k_2 q_1 - k_1 - \mu_1$$

$$ck_1\mu_2 + c\mu_1\mu_2$$

Since these are complicated expressions, we cannot reparameterize our original equations over them. However, one could reparameterize using the normal canonical form with these parameter combinations.

Now we reparameterize our original system using our simplified canonical set. Let $x_2^{'} = k_1k_2x_2$ and $x_3^{'} = x_3/q_2$. Then our original system becomes:

$$\dot{x}_1 = -\beta x_1x_4 - dx_1 + s$$

$$\dot{x}_2^{'} = \beta(q_1k_1k_2)x_1x_4 - (k_1 + \mu_1)x_2^{'}$$

$$\dot{x}_3^{'} = \beta x_1x_4 + (1/q_2k_2)x_2^{'} - \mu_2x_3^{'}$$

$$\dot{x}_4 = -cx_4 + k_2q_2x_3^{'}$$

$$y_1 = x_1$$

$$y_2 = x_4$$

We leave the reparameterization in the original parameter $\boldsymbol{p}$ to avoid confusion, since q_1 and q_2 are elements of $\boldsymbol{p}$.

8. Discussion

In all of the examples, we were able to go from infinitely many solutions (unidentifiability), as returned by DAISY, to one or finitely many solutions (identifiability). The original equations can be reparameterized in terms of the simplified canonical set, although sometimes this is not an easy task (as shown in Evans and Chappell [9]).

The reparameterizations of the original equations can sometimes be done by inspection, as demonstrated in most of our examples, but there are examples that cannot be rationally reparameterized, as demonstrated in the 4-compartment model. Although we can use the normal canonical form (i.e. companion matrix for linear problems), we prefer to use the original state variables, thus the inspection method is preferred over the normal canonical form.

The reparameterizations we found give a compact form where the new parameters in the equations are identifiable. However, since we make a change of variable with an unidentifiable parameter, the new state variable may not always be a useful quantity to examine. There are two cases: either the rescaled state variable is in the output (as in cases where both the input and output variables are multiplied by parameters in the original equations) or the output does not contain the rescaled variable. The latter condition may be preferable. However, the former condition can be remedied by a reparameterization by the companion matrix method.

The reparameterization of the input-output equations (Step 4) is not necessary to prove identifiability of the simplified canonical set. As stated above, if our canonical set contains decoupled elements, then the simplified canonical set is by definition identifiable. The reason the reparameterization of the coefficients is done is really to test if the original equations have a chance of being reparameterized in terms of the simplified canonical set. If the coefficients cannot be reparameterized in terms of the simplified canonical set, then the original system cannot be either. Also, we saw that the ability to reparameterize the coefficients of the input-output coefficients led to the interesting mathematical fact that the ideal generated by the canonical set is the same as the ideal generated by the exhaustive summary, when the canonical set contains only entire elements from Gröbner Bases. Additionally, the reparameterization step helps give us a useful *set* of identifiable combinations. We conjecture that there always exists a simplest set of identifiable combinations that lead to reparameterization of the input-output equations.

The "inspection" method, as discussed in the Differential Algebra Approach section, is the method of extracting simpler identifiable combinations from the coefficients of the input-output equations by subtracting/dividing the terms. The advantages of this method are it requires less computationally expensive machinery (no Gröbner Basis) and the correct solution can sometimes be found much faster. The disadvantages of this method are that, like many rule-based approaches, the procedure is ad hoc and may be forced to run through a large number of cases. In a sense, this method is doing exactly what a Gröbner Basis does, however the user/programmer must decide how to simplify the equations.

The inspection method can be used to easily find the simplest globally identifiable parameter combinations of the 2-compartment model, 4-compartment model, the SIR model, the unknown initial conditions example, and the rational coefficients example. The number of coefficients and their degrees of the 2-compartment nonlinear model make inspection a bit difficult. For the HIV/AIDS model, inspection can only give us the global identifiability of $\{\beta, d, s, q_2 k_2\}$, but we do not get the local identifiability of $\{q_1 k_1 k_2,\ \mu_1 + k_1, c, \mu_2\}$ so easily. Likewise, in the 3-compartment model, we do not easily get the local identifiability of $\{k_{12}k_{21}\ ,\ k_{13}k_{31}\ ,\ k_{02} + k_{12}\ ,\ k_{03} + k_{13}\}$ by inspection. However, the inspection method gives us globally identifiable combinations for both the HIV/AIDS model and the 3-compartment model, but due to their complicated nature, are not as useful since we must use the normal canonical form to reparameterize. In contrast, the combinations found from our algorithm are concise and may be useful quantities since we can reparameterize the original system. Additionally, our procedure can be automated whereas the inspection procedure is harder to automate.

In some cases, we have seen that the canonical set is actually apparent from the individual parameter solutions provided by DAISY. In the 2-compartment nonlinear model, the SIR model, the unknown initial conditions example, the rational coefficients example, and the HIV model, one can easily "cross-multiply" and find identifiable combinations. However, in the 2-compartment, 3-compartment, and 4-compartment models, this cannot be done. If we cross-multiply the equations in the solution and bring all the variables to one side, we obtain terms from Gröbner Bases that cannot be decoupled to form identifiable combinations. Thus we do not always obtain all the simplest identifiable parameter combinations from DAISY. Also, this means that not every term in a Gröbner Basis can be decoupled to obtain an identifiable combination.

Theorem 2 can be verified numerically by using a random numerical $\boldsymbol{p}^*$. If symbolic $\boldsymbol{p}^*$ is used, the Gröbner Bases will not be exactly the same because of how symbolic packages (like Mathematica) treat symbolic variables versus numbers, hence the bases will not be reduced. Note that Theorem 2 still applies in the case of rational coefficients of the input-output equations. In our example, we treat the term $1/p_2$ as a parameter and thus can still find a Gröbner Basis.

9. Conclusion

We have proposed an algorithm to find identifiable combinations of parameters in nonlinear ODE models and have found necessary and sufficient conditions for steps of the algorithm to work. We hope to examine our algorithm further to find what class of problems always give a canonical set that leads to reparameterization of the input-output equations. We are currently preparing a software distribution of our extended algorithm.

Acknowledgement

We would like to thank Professor Chris Anderson of the UCLA Mathematics Department for his constructive comments concerning this work.

References

[1] R. Bellman and K. J. Astrom, On structural identifiability, Math. Biosci. 7:329-339, 1970.

[2] G. Bellu, M. P. Saccomani, S. Audoly, and L D'Angio, DAISY: A new software tool to test global identifiability of biological and physiological systems, Computer Methods and Programs in Biomedicine 88:52-61, 2007.

[3] M. J. Chappell and R. N. Gunn, A procedure for generating locally identifiable reparameterisations of unidentifiable non-linear systems by the similarity transformation approach, Math. Biosci. 148: 21-41, 1998.

[4] C. Cobelli and J. J. DiStefano, III, Parameter and structural identifiability concepts and ambiguities: A critical review and analysis, Amer. J. Physiology: Regulatory, Integrative and Comparative Physiology 239: R7-R24, 1980.

[5] L. Denis-Vidal, G. Joly-Blanchard, C. Noiret, and M. Petitot, An Algorithm to test Identifiability of Non-linear Systems, Proc. 5th IFAC Conf. Nonlinear Control Systems, NOLCOS, St. Petersburg, 2001.

[6] L. Denis-Vidal and G. Joly-Blanchard, Equivalence and identifiability analysis of uncontrolled nonlinear dynamical systems, Automatica 40: 287-292, 2004.

[7] 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 57: 35-44, 2001.

[8] J. J. DiStefano, III and C. Cobelli, On parameter and structural identifiability: Nonunique observability/reconstructibility for identifiable systems, other ambiguities, and new definitions, IEEE Transactions on Automatic Control 25 (4): 830-833, 1980.

[9] N. Evans and M. Chappell, Extensions to a procedure for generating locally identifiable reparameterisations of unidentifiable systems, Math. Biosci. 168:137-159, 2000.

[10] K. R. Godfrey and J. J. DiStefano, III, Identifiability of model parameters, in: E. Walter (Ed.), Identifiability of Parametric Models, Pergamon, Oxford, Ch. 1, 1987, pp. 1-20.

[11] R. N. Gunn, M. J. Chappell, V. J. Cunningham, Reparameterisation of unidentifiable systems using the Taylor series approach, in: D.A. Linkens, E. Carson (Eds.), Proceedings of the Third IFAC Symposium on Modelling and Control in Biomedical Systems, Pergamon, Oxford, 1997, p. 247.

[12] L. Ljung and T. Glad, On global identifiability for arbitrary model parameterization, Automatica 30 (2):265-276, 1994.

[13] G. Margaria, E. Riccomagno, M. Chappell, and H. Wynn, Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences, Math. Biosci. 174:1-26, 2001.

[14] F. Ollivier, Le problem de l'identifiabilite structurelle globale: approche theoritique, methods effectives et bornes de complexite, PhD thesis, GAGE, Centre de Mathematiques, Ecole Polytechnique, 1990.

[15] F. Ollivier, Identifiabilite des systems, Technical Report, GAGE, Centre de Mathematiques, Ecole Polytechnique, pp. 43-54, 1998.

[16] H. Pohjanpalo, System Identifiability Based on the Power Series Expansion of the Solution, Math. Biosci. 41: 21-33, 1978.

[17] M. P. Saccomani, S. Audoly, G. Bellu, and L. D'Angio, A new differential algebra algorithm to test identifiability of nonlinear systems with given initial conditions, Proceedings of the 40th IEEE Conference on Decision and Control, Orlando, Florida, USA, December 2001.

[18] M.P. Saccomani and G. Bellu, DAISY: an Efficient Tool to Test Global Identifiability. Some Case Studies, 16th Mediterranean Conference on Control and Automation, Congress Centre, Ajaccio, France, June 25-27, 2008.

[19] M. P. Saccomani, S. Audoly, G. Bellu, and L. D'Angio, Parameter Identifiability of Nonlinear Biological Systems, in: L. Benvenuti, A. De Santis, and L. Farina (Eds.), Positive Systems, LNCIS 294, Springer, Berlin, 2003, pp. 87-93.

[20] M. P. Saccomani, S. Audoly, and L. D'angio, Parameter identifiability of nonlinear systems: the role of initial conditions, Automatica 39: 619-632, 2003.

[21] D. Shannon and M. Sweedler, Using Groebner bases to determine algebra membership, split surjective algebra homomorphisms and determine birational equivalence, Journal of Symbolic Computation 6 (2-3): 267-273, 1988.

[22] N. Verdiere, L. Denis-Vidal, G. Joly-Blanchard, D. Domurado, Identifiability and estimation of pharmacokinetic parameters for the ligands of the macrophage mannose receptor, Int. J. Applied Math Comput. Sci. 15 (4):517-526, 2005.