Finding identifiable parameter combinations in nonlinear ODE models and the rational reparameterization of their input-output equations

Nicolette Meshkat[1]*, Chris Anderson[2], and Joseph J. DiStefano III[3]

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

Abstract

When examining the structural identifiability properties of dynamic system models, some parameters can take on an infinite number of values and yet yield identical input-output data. These parameters and the model are then said to be unidentifiable. Finding identifiable combinations of parameters with which to reparameterize the model provides a means for quantitatively analyzing the model and computing solutions in terms of the combinations. In this paper, we revisit and explore the properties of an algorithm for finding identifiable parameter combinations using Gröbner Bases and prove useful theoretical properties of these parameter combinations. We prove a set of M algebraically independent identifiable parameter combinations can be found using this algorithm and that there exists a unique rational reparameterization of the input-output equations over these parameter combinations. We also demonstrate application of the procedure to a nonlinear biomodel.

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

1. Introduction

Parameter identifiability analysis for dynamic system ODE models addresses the question of which unknown parameters can be quantified from given input-output data. *Unidentifiable* parameters can take on an uncountably infinite number of values and yet result in identical input-output data. In such cases, the model and its parameter vector $\boldsymbol{p}$ are underdetermined with respect to the input-output data. This indeterminacy can be removed by finding the 'simplest' combinations of parameters that take on a unique or finite number of values, which are then used as candidates to reparameterize the model, rendering it *identifiable*. Thus the question becomes, how can identifiable parameter combinations be found?

This question has been partially answered for several model classes, under limited conditions. Evans and Chappell [1] and Gunn et al [2] adapt the Taylor series approach of Pohjanpalo [3] to find locally identifiable combinations. Chappell and Gunn [4] use the similarity transformation approach to generate locally identifiable reparameterizations. Thus, with these methods identifiability can only be guaranteed (at least) locally. The problem of finding identifiable parameter combinations has also been addressed using differential algebra methods, as Denis-Vidal et al [5] and Boulier [6] find globally identifiable combinations of parameters using an "inspection" method as discussed later in this paper. However, as shown by Meshkat et al [7], this method is difficult to implement as a fully automated

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

computational procedure.

In [7], an algorithm was outlined for finding the 'simplest' set of globally identifiable parameter combinations for a practical class of nonlinear ODE models. This algorithm extended the method of Saccomani and coworkers [8] using a variation of the Gröbner Basis approach. In this paper, we address several questions that arose in [7] regarding properties of the identifiable parameter combinations found, including algebraic independence and the existence of a rational reparameterization of the input-output equations derived from the original nonlinear model. Although a rational reparameterization of the original nonlinear model cannot always be done (as shown in [1]), we prove here that a unique rational reparameterization of the input-output equations can always be found over algebraically independent parameter combinations. In addition to being useful in quantifying the model and exercising its solutions, we will show that the ability to rationally reparameterize the input-output equations leads to a rigorous proof of identifiability.

2. Nonlinear ODE Model

The general form of the models under consideration is:

$$\dot{\boldsymbol{x}}(t,\boldsymbol{p}) = \boldsymbol{f}(\boldsymbol{x}(t,\boldsymbol{p}),\boldsymbol{u}(t),t;\ \boldsymbol{p}), t \in [t_0, T]$$

$$\boldsymbol{y}(t,\boldsymbol{p}) = \boldsymbol{g}(\boldsymbol{x}(t,\boldsymbol{p});\boldsymbol{p}) \tag{2.1}$$

Here $\boldsymbol{x}$ is a n-dimensional state variable, $\boldsymbol{p}$ is a P -dimensional parameter vector, $\boldsymbol{u}$ is the r-dimensional input vector, and $\boldsymbol{y}$ is the m-dimensional output vector. 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 [9]. For example, $\boldsymbol{p} \geq \boldsymbol{0}$ is common.

3. Identifiability and the Differential Algebra Approach

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 parameter values in the real "noisy" data problem, often called the *numerical identifiability* problem.

Structural identifiability can be expressed as an injectivity condition, as in [8]. 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. If there exists only one solution $\boldsymbol{p} = \boldsymbol{p}^*$, then this corresponds to global identifiability. If there exists finitely many distinct solutions for $\boldsymbol{p}$, then this corresponds to local identifiability. Infinitely many solutions for $\boldsymbol{p}$ corresponds to unidentifiability.

The *a priori structural identifiability* problem can be solved using the differential algebra approach of Saccomani et al [8], which follows methods developed by Ljung and Glad [10] and Ollivier [11,12]. Their program, DAISY, can be used to automatically check global identifiability of nonlinear dynamic models [13]. We summarize their approach below. A detailed description can be found in [7,13].

Using Ritt's pseudodivision algorithm, an input-output map can be determined in implicit form. The result of the pseudodivision algorithm is called the *characteristic set* [11]. Since the ideal generated by (2.1) is a prime ideal [14], the characteristic set is a "minimal" set of differential polynomials which generate the same differential ideal as the ideal generated by (2.1) [13]. The first m equations of the characteristic set are those independent of the state variables, and form the *input-output relations* [13]:

$$\boldsymbol{\Psi}(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{p}) = \boldsymbol{0} \tag{3.1}$$

The characteristic set is in general non-unique, but the coefficients of the input-output equations can be fixed uniquely by normalizing the equations to make them monic [13].

The m equations of the input-output relations $\boldsymbol{\Psi}(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{p}) = \boldsymbol{0}$ are polynomial equations in the variables $\boldsymbol{u}, \dot{\boldsymbol{u}}, \ddot{\boldsymbol{u}}, \ldots, \boldsymbol{y}, \dot{\boldsymbol{y}}, \ddot{\boldsymbol{y}}, \ldots$ with rational coefficients in the parameter vector $\boldsymbol{p}$. Specifically, these equations involve polynomials from the differential ring $R(\boldsymbol{p})[\boldsymbol{u}, \boldsymbol{y}]$, where $R(\boldsymbol{p})$ is the field of rational functions over the real numbers in the parameter vector $\boldsymbol{p}$. For each equation, we can write $\Psi_j(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{p}) = \sum_i c_i(\boldsymbol{p})\psi_i(\boldsymbol{u}, \boldsymbol{y})$, where $c_i(\boldsymbol{p})$ is a rational function in the parameter vector $\boldsymbol{p}$ and $\psi_i(\boldsymbol{u}, \boldsymbol{y})$ is a monomial function in the variables $\boldsymbol{u}, \dot{\boldsymbol{u}}, \ddot{\boldsymbol{u}}, \ldots, \boldsymbol{y}, \dot{\boldsymbol{y}}, \ddot{\boldsymbol{y}}, \ldots$, etc. We call $c_i(\boldsymbol{p})$ the coefficients of the input-output equations.

To form an injectivity condition, we set $\boldsymbol{\Psi}(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{p}) = \boldsymbol{\Psi}(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{p}^*)$. Then global identifiability becomes injectivity of the map $\boldsymbol{c}(\boldsymbol{p})$ [13]. That is, identifiability is determined by the equations

$$\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*) \tag{3.2}$$

for arbitrary $\boldsymbol{p}^*$ [13]. Thus, 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 $\boldsymbol{p}^*$ [13]. The equations $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ are called the *exhaustive summary* [11].

If there are finitely many distinct solutions for $\boldsymbol{p}$, then the model (2.1) is locally identifiable. The model (2.1) is unidentifiable if there are infinitely many solutions for $\boldsymbol{p}$, that is, the solution for $\boldsymbol{p}$ is expressed in terms of one or more free variables. Thus, determining structural identifiability is reduced to the nature of the solutions to $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$, which is typically solved by finding a Gröbner Basis and using elimination [13].

4. Some methods for finding identifiable parameter combinations

We focus on the case when (3.2) has infinitely many solutions (unidentifiability) in this paper. Unidentifiable models cannot be quantified from input-output data. A useful alternative is to find identifiable parameter combinations which can always be determined from input-output data, and attempt to reparameterize our model (2.1) in terms of these new parameters. Before we revisit our method for finding identifiable parameter combinations [7], we briefly present two other methods for finding identifiable parameter combinations using the differential algebra approach. Both procedures rely on using the exhaustive summary $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ to find parameter combinations that are either uniquely or finitely determined by $\boldsymbol{p}^*$.

Definition: Let s be the number of free parameters, defined as the number of total parameters P minus the number of equations M in the solution of $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$.

That is, there are s free parameters and M "non-free" parameters, where $P = M + s$. Sometimes identifiable combinations can easily be found directly from the solutions to the equations $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$, by algebraically manipulating their solutions to form $M = P - s$ parameter combinations in terms of $\boldsymbol{p}^*$ only. In other words, find solutions of the form $g(\boldsymbol{p}) = g(\boldsymbol{p}^*)$. For example, in the Nonlinear 2-Compartment Model in [7], the solution to $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ is of the following form, where $\boldsymbol{p} = \{k_{21}, k_{12}, V_M, K_M, k_{02}, c_1, b_1\}$ and $\boldsymbol{p}^* = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta\}$:

$$V_M = \frac{\gamma\zeta}{c_1}$$

$$k_{21} = \alpha$$

$$k_{12} = \beta$$

$$b_1 = \frac{\zeta\eta}{c_1}$$

$$k_{02} = \epsilon$$

$$K_M = \frac{\delta\zeta}{c_1}$$

Then clearly $\{c_1V_M, k_{21}, k_{12}, b_1c_1, k_{02}, c_1K_M\}$ are uniquely determined by $\boldsymbol{p}^*$ because we can move the parameter vector $\boldsymbol{p}$ all to one side of the equation. To verify global identifiability, one would then reparameterize $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ over these parameter combinations $\{c_1V_M, k_{21}, k_{12}, b_1c_1, k_{02}, c_1K_M\}$ and check the injectivity condition.

However, this ability to "move all parameters to one side of the equation" and thus "decouple" our parameter solution cannot always easily be done, as demonstrated in the Linear 2-Compartment Model below [7], where $\boldsymbol{p} = \{k_{01}, k_{02}, k_{12}, k_{21}, v\}$ and $\boldsymbol{p}^* = \{\alpha, \beta, \gamma, \delta, \epsilon\}$:

$$k_{21} = \alpha + \delta - \frac{\alpha k_{02} + \alpha\beta - \delta k_{02} + \beta\delta + \alpha\gamma}{k_{02} - \beta - \gamma}$$

$$k_{01} = \frac{\alpha k_{02} - \alpha\beta + \delta k_{02} - \beta\delta - \alpha\gamma}{k_{02} - \beta - \gamma}$$

$$k_{12} = -k_{02} + \beta + \gamma$$

$$v = \epsilon$$

Here we see that we it takes more effort to find the uniquely determined parameter combinations $\{v, k_{12}k_{21}, k_{02} + k_{12}, k_{01} + k_{21}\}$.

Why is this the case? The solutions to $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ can be found by solving for a Gröbner Basis, which is basically a way of simplifying the equations to a "triangular form", followed by elimination. No attempt to keep the $\boldsymbol{p}$ and $\boldsymbol{p}^*$ parameters separated is made during the elimination process, and so it is simply fortuitous if this occurs! Other examples of models whose parameter solutions cannot easily be decoupled can be found in [7].

The other way to find identifiable combinations is through the process called "inspection" [6]. The coefficients $c_i(\boldsymbol{p})$ of the input-output equations are assumed to be identifiable [15]. The process of "inspection" involves adding/subtracting/multiplying/dividing the coefficients $c_i(\boldsymbol{p})$ amongst each other to form simpler identifiable combinations, which are always of the form $g(\boldsymbol{p}) = g(\boldsymbol{p}^*)$. For example, if the coefficients are:

$$c_1(\boldsymbol{p}) = b_1 c_1$$

$$c_2(\boldsymbol{p}) = c_1 K_M$$

$$c_3(\boldsymbol{p}) = k_{02} + k_{12} + k_{21}$$

$$c_4(\boldsymbol{p}) = b_1 c_1 k_{02} + b_1 c_1 k_{12}$$

$$c_5(\boldsymbol{p}) = 2c_1 k_{02} k_{21} K_M + c_1 k_{02} V_M + c_1 k_{12} V_M$$

$$c_6(\boldsymbol{p}) = k_{02} k_{21}$$

Then it is obvious that the combinations $\{c_1 V_M, k_{21}, k_{12}, b_1 c_1, k_{02}, c_1 K_M\}$ are also uniquely determined by $\boldsymbol{p}^*$ and by reparameterizing $\boldsymbol{c}(\boldsymbol{p})$ over these combinations, one could verify injectivity and thus global identifiability of the reparameterized model.

However, if we instead have the following as coefficients:

$$c_1(\boldsymbol{p}) = v$$

$$c_2(\boldsymbol{p}) = (k_{01} + k_{21} + k_{02} + k_{12})v$$

$$c_3(\boldsymbol{p}) = (k_{01}k_{02} + k_{01}k_{12} + k_{02}k_{21})v$$

$$c_4(\boldsymbol{p}) = k_{02} + k_{12}$$

It is not so easy to simplify these coefficients to form simpler identifiable combinations. In this case, $\{v, k_{02} + k_{12}, k_{01} + k_{21}\}$ are clearly uniquely determined by $\boldsymbol{p}^*$, but the fourth combination $k_{12}k_{21}$ only becomes apparent if we factor $k_{01}k_{02} + k_{01}k_{12} + k_{02}k_{21}$ as $(k_{12} + k_{02})(k_{01} + k_{21}) - k_{12}k_{21}$. There are some examples of models in [7] where inspection gets trickier, primarily when there are a finite number of distinct solutions for $\boldsymbol{p}$, expressed in terms of one or more free variables. This is where a Gröbner Basis comes to the rescue.

5. Algorithm for finding identifiable parameter combinations

Our algorithm for finding identifiable combinations is based on the principle that a Gröbner Basis is in a sense a "simpler form" of $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$. When testing for identifiability using the differential algebra approach of [13], we are solving the system $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ by finding a Gröbner Basis and then by elimination, finding a solution for $\boldsymbol{p}$ in terms of $\boldsymbol{p}^*$ and possibly free parameters. Since a Gröbner Basis helps solve the system $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$, by reducing it to a simpler (triangular) form, we speculated and consequently demonstrated in [7] that the Gröbner Basis generates identifiable combinations that are 'simpler' than $\boldsymbol{c}(\boldsymbol{p})$.

There are at least M coefficients of the input-output equations, by definition. Thus, the exhaustive summary $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ is composed of at least M equations. From the exhaustive summary, we construct a Gröbner Basis of the form $\boldsymbol{G} = \{G_1(\boldsymbol{p}, \boldsymbol{p}^*), \dots, G_k(\boldsymbol{p}, \boldsymbol{p}^*)\}$, where G_i is a polynomial function with $k \geq M$, depending on the ranking of parameters. Our goal is to find M terms of the form

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

appearing either as an element by itself or as a factor of an element in a Gröbner Basis, since when set to zero this means $q_i(\boldsymbol{p})$ has either a unique or finite number of solutions, respectively. In other words, $G_i(\boldsymbol{p}, \boldsymbol{p}^*)$ is "decoupled" into a polynomial in $\boldsymbol{p}$ minus the same polynomial in $\boldsymbol{p}^*$.

For example, if there is a Gröbner Basis element $p_1p_2 - p_1^*p_2^*$, this means p_1p_2 has a unique solution. Or we may have $(p_3p_4 - p_3^*p_4^*)(p_3p_4 - p_5^*p_6^*)$ and $(p_5p_6 - p_3^*p_4^*)(p_5p_6 - p_5^*p_6^*)$ as elements, which means that p_3p_4 and p_5p_6 have a finite number of distinct solutions.

Note that instead of (5.1), we may 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, $p_1^*p_2p_3 - p_1^*p_2^*p_3^*$ reduces to $p_2p_3 - p_2^*p_3^*$.

Additionally, sometimes the Gröbner Basis element or factor can be rewritten in decoupled form in order to get an identifiable expression. For example, an element $p_2^*p_1 - p_1^*p_2$ can be decoupled as $\frac{p_1}{p_2} - \frac{p_1^*}{p_2^*}$.

Determination of additional expressions of the type (5.1) depends upon the choice of ranking of parameters when constructing the Gröbner Basis. The combinations we seek may not all appear in a single Gröbner Basis, hence the need for several rankings of parameters. One technique described in [7] is to try all P shifts of the parameter vector $\boldsymbol{p}$, since this forces each parameter to have the highest ranking, and thus be eliminated in that order. However, this may not give all M decoupled terms, thus different permutations of the parameter vector $\boldsymbol{p}$ may also need to be tested [7]. In practice, an *a priori* guess for identifiable combinations can help determine an optimal rank ordering, as demonstrated in the example below.

More than M decoupled elements can appear in the Gröbner Bases, so as stated in [7], we look for a set of the M 'simplest' algebraically independent combinations that as a set span all P parameters. We note that if the model is reducible, i.e. if one or more parameters in the model equations do not appear in the input-output equations, then we rename P to the number of parameters appearing in the input-output equations. By 'simplest', we mean the lowest degree and fewest number of terms. Algebraic independence will be defined in Section 8. We called a set of the M simplest terms of the form (5.1) the *canonical set* [7].

To formally check identifiability, one attempts to reparameterize the coefficients $\boldsymbol{c}(\boldsymbol{p})$ of the input-output equations over the terms $q = q_i(\boldsymbol{p})$. If a reparameterization $\tilde{\boldsymbol{c}}(\boldsymbol{q})$ exists, then injectivity of $\tilde{\boldsymbol{c}}(\boldsymbol{q})$ is tested, i.e. if $\tilde{\boldsymbol{c}}(\boldsymbol{q}) = \tilde{\boldsymbol{c}}(\boldsymbol{q}^*)$, does $\boldsymbol{q} = \boldsymbol{q}^*$? We found in [7] that if $q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*)$ appeared as an element in a Gröbner Basis, then global identifiability results, whereas if $q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*)$ appeared as a factor in a Gröbner Basis, then local identifiability results.

A more detailed explanation of our algorithm can be found in [7]. We summarize it as three basic steps:

Step 1: Search through all relevant rankings and determine elements of the Gröbner Bases (or factors, as needed) that *can be simplified* to the decoupled form $q_i(\boldsymbol{p}) - q_i(\boldsymbol{p}^*)$.

Step 2: Select the M 'simplest' algebraically independent combinations. By 'simplest', we mean the lowest degree and fewest number of terms. The set of M combinations must span all P parameters.

Step 3: Verify the injectivity condition of the model, that is, reparameterize $\boldsymbol{c}(\boldsymbol{p})$ as $\tilde{\boldsymbol{c}}(\boldsymbol{q})$ and then test if $\tilde{\boldsymbol{c}}(\boldsymbol{q}) = \tilde{\boldsymbol{c}}(\boldsymbol{q}^*)$ implies that $\boldsymbol{q}$ has a unique or finite number of solutions.

6. Example of finding identifiable parameter combinations

We now demonstrate our algorithm on a classic 2-compartment model that has been made nonlinear.

$$\dot{x}_1 = -(k_{21} + k_{01})x_1 \frac{V_M}{K_M + x_1} + k_{12}x_2 + b_1 u$$

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

$$y = c_1 x_1$$

Definitions:

x_1, x_2 state variables

u input

y output

$k_{01}, k_{02}, k_{12}, k_{21}, V_M, K_M, b_1, c_1$ unknown parameters

Figure 1.

Let $\boldsymbol{p} = \{k_{01}, k_{02}, k_{12}, k_{21}, V_M, K_M, b_1, c_1\}$ and $\boldsymbol{p}^* = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta, \theta\}$.

The input-output equation determined by Ritt's pseudodivision algorithm is:

$$\begin{aligned}-b_1c_1^3K_M^2\dot{u} - 2b_1c_1^2K_M\dot{u}y - b_1c_1\dot{u}y^2 + c_1^2K_M^2\ddot{y} + 2c_1K_My\ddot{y} + \ddot{y}y^2 \\ + (c_1^2k_{01}K_MV_M + c_1^2k_{02}K_M^2 + c_1^2k_{12}K_M^2 + c_1^2k_{21}K_MV_M)\dot{y} + 2(c_1k_{02}K_M + c_1k_{12}K_M)y\dot{y} \\ + (k_{02} + k_{12})\dot{y}y^2 - (b_1c_1^3k_{02}K_M^2 + b_1c_1^3k_{12}K_M^2)u - 2(b_1c_1^2k_{02}K_M + b_1c_1^2k_{12}K_M)uy \\ - (b_1c_1k_{02} + b_1c_1k_{12})uy^2 + (c_1^2k_{01}k_{02}K_MV_M + c_1^2k_{01}k_{12}K_MV_M + c_1^2k_{02}k_{21}K_MV_M)y \\ + (c_1k_{01}k_{02}V_M + c_1k_{01}k_{12}V_M + c_1k_{02}k_{21}V_M)y^2 = 0\end{aligned}$$

Notice there are 13 coefficients, but 5 of them are algebraically independent. Thus, we only use the coefficients that cannot be described as a polynomial or rational functions of the other coefficients, as discussed in Section 9 of this paper.

Five coefficients $\boldsymbol{c}(\boldsymbol{p})$ that satisfy this condition are:

$$b_1c_1$$

$$c_1K_M$$

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

$$c_1k_{01}k_{02}V_M + c_1k_{01}k_{12}V_M + c_1k_{02}k_{21}V_M$$

$$c_1^2k_{01}K_MV_M + c_1^2k_{02}K_M^2 + c_1^2k_{12}K_M^2 + c_1^2k_{21}K_MV_M$$

We solve $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$:

$$b_1c_1 = \eta\theta$$

$$c_1K_M = \zeta\theta$$

$$k_{02} + k_{12} = \beta + \gamma$$

$$c_1k_{01}k_{02}V_M + c_1k_{01}k_{12}V_M + c_1k_{02}k_{21}V_M = \alpha\beta\epsilon\theta + \alpha\gamma\epsilon\theta + \beta\delta\epsilon\theta$$

$$c_1^2k_{01}K_MV_M + c_1^2k_{02}K_M^2 + c_1^2k_{12}K_M^2 + c_1^2k_{21}K_MV_M = \alpha\epsilon\zeta\theta^2 + \delta\epsilon\zeta\theta^2 + \beta\zeta^2\theta^2 + \gamma\zeta^2\theta^2$$

To get:

$$b_1 = \frac{\eta\theta}{c_1}$$

$$k_{21} = \frac{\gamma\delta\epsilon\theta}{c_1V_M(-k_{02} + \beta + \gamma)}$$

$$k_{12} = -k_{02} + \beta + \gamma$$

$$k_{01} = \frac{(-\alpha(\beta + \gamma) - \beta\delta + k_{02}(\alpha + \delta))\epsilon\theta}{c_1 V_M (k_{02} - \beta - \gamma)}$$

$$K_M = \frac{\zeta\theta}{c_1}$$

Thus not all of the identifiable combinations are obvious from this solution.

In order to find identifiable parameter combinations, we now search through Gröbner Bases for decoupled terms or factors. An *a priori* guess for identifiable parameter combinations can help determine an optimal rank ordering.

For example, multiplication of the second and third equations in the solution to $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ and multiplication by $c_1 V_M$ yields that

$$c_1 V_M k_{12} k_{21} = \gamma\delta\epsilon\theta$$

thus, we guess that $c_1 V_M k_{12} k_{21}$ is an identifiable parameter combination. Thus, we should test a Gröbner Basis ranking where these parameters are eliminated last and grouped in that order. For example, the ranking $\{K_M, k_{02}, b_1, k_{01}, c_1, V_M, k_{12}, k_{21}\}$ keeps c_1, V_M, k_{12}, k_{21} together and eliminates them last.

The Gröbner Basis with ranking $\{K_M, k_{02}, b_1, k_{01}, c_1, V_M, k_{12}, k_{21}\}$ can be found using Mathematica:

$$\begin{pmatrix}
-\text{c1 k12 k21 vm}\,\zeta\theta + \gamma\delta\epsilon\zeta\theta^2 \\
-\text{k12 k21}\,\alpha\epsilon\zeta\theta^2 - \text{k12 k21}\,\delta\epsilon\zeta\theta^2 + \text{k01}\,\gamma\delta\epsilon\zeta\theta^2 + \text{k21}\,\gamma\delta\epsilon\zeta\theta^2 \\
-\text{k12 k21 vm}\,\alpha\zeta^2\theta^2 - \text{k12 k21 vm}\,\delta\zeta^2\theta^2 + \text{k01 vm}\,\gamma\delta\zeta^2\theta^2 + \text{k21 vm}\,\gamma\delta\zeta^2\theta^2 \\
-\text{k12 k21 vm}\,\alpha\zeta\eta\theta^2 - \text{k12 k21 vm}\,\delta\zeta\eta\theta^2 + \text{k01 vm}\,\gamma\delta\zeta\eta\theta^2 + \text{k21 vm}\,\gamma\delta\zeta\eta\theta^2 \\
-\text{c1 k12 k21 vm} + \text{c1 k01 vm}\,\beta + \text{c1 k21 vm}\,\beta + \text{c1 k01 vm}\,\gamma + \text{c1 k21 vm}\,\gamma - \alpha\beta\epsilon\theta - \alpha\gamma\epsilon\theta - \beta\delta\epsilon\theta \\
\text{c1 k01 vm}\,\zeta\theta + \text{c1 k21 vm}\,\zeta\theta - \alpha\epsilon\zeta\theta^2 - \delta\epsilon\zeta\theta^2 \\
-\text{b1}\,\alpha\beta\epsilon\theta - \text{b1}\,\alpha\gamma\epsilon\theta - \text{b1}\,\beta\delta\epsilon\theta - \text{k12 k21 vm}\,\eta\theta + \text{k01 vm}\,\beta\eta\theta + \text{k21 vm}\,\beta\eta\theta + \text{k01 vm}\,\gamma\eta\theta + \text{k21 vm}\,\gamma\eta\theta \\
\text{b1}\,\gamma\delta\epsilon\zeta\theta^2 - \text{k12 k21 vm}\,\zeta\eta\theta^2 \\
-\text{b1}\,\alpha\epsilon\zeta\theta^2 - \text{b1}\,\delta\epsilon\zeta\theta^2 + \text{k01 vm}\,\zeta\eta\theta^2 + \text{k21 vm}\,\zeta\eta\theta^2 \\
\text{b1 c1} - \eta\theta \\
\text{k02} + \text{k12} - \beta - \gamma \\
-\text{km}\,\alpha\beta\epsilon\theta - \text{km}\,\alpha\gamma\epsilon\theta - \text{km}\,\beta\delta\epsilon\theta - \text{k12 k21 vm}\,\zeta\theta + \text{k01 vm}\,\beta\zeta\theta + \text{k21 vm}\,\beta\zeta\theta + \text{k01 vm}\,\gamma\zeta\theta + \text{k21 vm}\,\gamma\zeta\theta \\
\text{km}\,\gamma\delta\epsilon\zeta\theta^2 - \text{k12 k21 vm}\,\zeta^2\theta^2 \\
-\text{km}\,\alpha\epsilon\zeta\theta^2 - \text{km}\,\delta\epsilon\zeta\theta^2 + \text{k01 vm}\,\zeta^2\theta^2 + \text{k21 vm}\,\zeta^2\theta^2 \\
\text{b1}\,\zeta\theta - \text{km}\,\eta\theta \\
\text{c1 km} - \zeta\theta
\end{pmatrix}$$

Notice there are many decoupled terms to choose from, but we only need $M = P - s$ = 8 − 3 = 5 parameter combinations. In this case, the decoupled terms can be found from a single Gröbner basis, but this is not true in general, as shown in [7].

Thus we pick a set of the 'simplest' algebraically independent parameter combinations:

$$q_1 = b_1 c_1$$

$$q_2 = c_1 K_M$$

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

$$q_4 = c_1 V_M k_{12} k_{21}$$

$$q_5 = c_1 V_M (k_{01} + k_{21})$$

Notice that the set of 'simplest' algebraically independent parameter combinations is not unique. For instance, b_1/K_M also appears as a decoupled term, thus the term $b_1 c_1$ could be replaced by it.

We can reparameterize our coefficients as $\tilde{\boldsymbol{c}}(\boldsymbol{q})$ by finding a Gröbner Basis of $\{c_i(\boldsymbol{p}) - \hat{c}_i,\ b_1 c_1 - \hat{q}_1,\ c_1 K_M - \hat{q}_2,\ k_{02} + k_{12} - \hat{q}_3,\ c_1 V_M k_{12} k_{21} - \hat{q}_4,\ c_1 V_M (k_{01} + k_{21}) - \hat{q}_5\}$ in the rank ordering $\{k_{01}, k_{02}, k_{12}, k_{21}, V_M, K_M, b_1, c_1, \hat{q}_1, \ldots, \hat{q}_5, \hat{c}_i\}$ for each coefficient $c_i(\boldsymbol{p})$. We get the following reparameterized coefficients of the input-output equation:

$$-q_1 q_2^2 \dot{u} - 2q_1 q_2 \dot{u} y - q_1 \dot{u} y^2 + q_2^2 \ddot{y} + 2q_2 y \ddot{y} + \ddot{y} y^2 + (q_2^2 q_3 + q_2 q_5)\dot{y} + 2q_2 q_3 y \dot{y} + q_3 \dot{y} y^2 \\ - q_1 q_2^2 q_3 u - 2q_1 q_2 q_3 u y - q_1 q_3 u y^2 + (q_2 q_3 q_5 - q_2 q_4) y + (q_3 q_5 - q_4) y^2 = 0$$

Then, when we set $\tilde{\boldsymbol{c}}(\boldsymbol{q}) = \tilde{\boldsymbol{c}}(\boldsymbol{q}^*)$, we get $\boldsymbol{q} = \boldsymbol{q}^*$, which means that our parameter combinations $\boldsymbol{q}$ are globally identifiable.

Thus, we can reparameterize our nonlinear model using a "canonical form", i.e. reduction to a first order system. Let $y = v_1, \dot{y} = \dot{v}_1 = v_2, u = u_1, \dot{u}_1 = u_2$. Then the input-output equation becomes:

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

$$\dot{v}_2 = (v_1 - q_2)^{-2}(q_1 q_2^2 u_2 + 2q_1 q_2 u_2 v_1 + q_1 u_2 v_1^2 - (q_2^2 q_3 + q_2 q_5) v_2 - 2q_2 q_3 v_1 v_2 - q_3 v_2 v_1^2 \\ + q_1 q_2^2 q_3 u_1 + 2q_1 q_2 q_3 u_1 v_1 + q_1 q_3 u_1 v_1^2 - (q_2 q_3 q_5 - q_2 q_4) v_1 - (q_3 q_5 - q_4) v_1^2)$$

where q_1, q_2, q_3, q_4, q_5 are all globally identifiable.

Using these $\boldsymbol{q}$, we seek to reparameterize the original model. In this case, this can be done by using the scaling: $X_1 = c_1 x_1$ and $X_2 = c_1 k_{12} x_2$

$$\dot{X}_1 = -X_1 \frac{q_5}{q_2 + X_1} + X_2 + q_1 u$$

$$\dot{X}_2 = X_1 \frac{q_4}{q_2 + X_1} - q_3 X_2$$

$$y = X_1$$

Notice that the reparameterized model no longer has a "compartment" form. This can be remedied by a simple algebra trick:

$$\dot{X}_1 = -X_1 \frac{(q_5 - q_4)}{q_2 + X_1} - X_1 \frac{q_4}{q_2 + X_1} + X_2 + q_1 u$$

$$\dot{X}_2 = X_1 \frac{q_4}{q_2 + X_1} - (q_3 - 1)X_2 - X_2$$

$$y = X_1$$

Figure 2.

7. Theoretical Considerations

As discussed above, there are 3 steps in our approach. First, we find parameter combinations $\boldsymbol{q}(\boldsymbol{p})$ from the Gröbner Bases of the exhaustive summary. Second, we reparameterize $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$ to get $\tilde{\boldsymbol{c}}(\boldsymbol{q})$. Third, we show that $\tilde{\boldsymbol{c}}(\boldsymbol{q}) = \tilde{\boldsymbol{c}}(\boldsymbol{q}^*)$ implies that $\boldsymbol{q}$ has a unique or finite number of solutions, and thus we have rigorously proven identifiability.

The main objective of the theoretical work is to show that the Gröbner Bases formed from the exhaustive summary $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ always provide a set of M combinations $\boldsymbol{q}(\boldsymbol{p})$ such that there exists a unique rational reparameterization of $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$. To establish this, we show:

I. A set of M algebraically independent parameter combinations $\boldsymbol{q}(\boldsymbol{p})$ can always be obtained from the Gröbner Bases of the exhaustive summary or from the exhaustive summary itself (Theorem 1).

II. For such a set of M combinations, there exists a unique rational reparameterization of $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$, i.e. $\tilde{\boldsymbol{c}}(\boldsymbol{q})$ (Theorem 2).

III. By construction of the $\boldsymbol{q}(\boldsymbol{p})$, we have that $\tilde{\boldsymbol{c}}(\boldsymbol{q}) = \tilde{\boldsymbol{c}}(\boldsymbol{q}^*)$ implies the combinations $\boldsymbol{q}(\boldsymbol{p})$ are either globally or locally identifiable, depending on whether they have a unique or finite number of solutions in the Gröbner Bases of the exhaustive summary (Theorem 3).

We address (I) in Section 8 and prove there are at least M algebraically independent parameter combinations $\boldsymbol{q}(\boldsymbol{p})$ obtained from decoupled terms or factors in the Gröbner Bases of the exhaustive summary (adjoined with the exhaustive summary). We also prove that if a decoupled term is contained in the ideal of other decoupled terms (thus “redundant”), then it is a polynomial or rational combination of these terms.

We address (II) in Section 9 and show that any term $c_i(\boldsymbol{p}) - c_i(\boldsymbol{p}^*)$ is “redundant” with respect to the ideal generated by $(q_1(\boldsymbol{p}) - q_1(\boldsymbol{p}^*), q_2(\boldsymbol{p}) - q_2(\boldsymbol{p}^*), \ldots, q_M(\boldsymbol{p}) - q_M(\boldsymbol{p}^*))$. We also show there are at most M algebraically independent parameter combinations $\boldsymbol{q}(\boldsymbol{p})$, and thus there are exactly M. Combining these results implies there is a unique rational reparameterization of the input-output equations over our M algebraically independent parameter combinations $\boldsymbol{q}(\boldsymbol{p})$.

We address (III) in Section 10 and show that a unique rational reparameterization of the coefficients of the input-output equations over our M algebraically independent parameter combinations implies that these combinations are in fact identifiable.

Before we address the algebraic independence of $\boldsymbol{q}(\boldsymbol{p})$, we must first show that the Gröbner Bases generate enough (that is, M) decoupled terms or factors that as a set span all P parameters. A simple explanation is that, even in the pathological case where Gröbner Bases do not generate enough combinations, we always have at least M coefficients $\boldsymbol{c}(\boldsymbol{p})$ of the input-output equations, which are known to be identifiable [15] and must span all P parameters, by definition. Thus, there will always be enough identifiable combinations to reparameterize $\boldsymbol{c}(\boldsymbol{p})$, i.e. we can trivially reparameterize $\boldsymbol{c}(\boldsymbol{p})$ over itself. The point of using a Gröbner Basis is that we can form 'simpler' parameter combinations. However, if no new ones are generated, this suggests the $\boldsymbol{c}(\boldsymbol{p})$ were "simple enough"! Thus, to simplify notation from now on, when we refer to the "decoupled combinations" $\boldsymbol{q}(\boldsymbol{p})$, we are referring to the parameter combinations $\boldsymbol{q}(\boldsymbol{p})$ obtained from decoupled terms or factors in the Gröbner Bases of $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$, and in the pathological case where not enough are generated, then $\boldsymbol{q}(\boldsymbol{p})$ possibly includes terms in $\boldsymbol{c}(\boldsymbol{p})$.

In our previous work [7], we referred to a decoupled combination found in the Gröbner Basis of (3.2) as "identifiable". We refrain from preemptively calling the decoupled combinations "identifiable" here, since the goal is to find the conditions to rationally reparameterize $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$ and thus get identifiability in the rigorous sense, i.e. the injectivity definition. We will see that a sufficient condition for finding a unique rational reparameterization is that our M parameter combinations $\boldsymbol{q}(\boldsymbol{p})$ are algebraically independent.

8. Algebraic Independence

Definition: A subset $S = \{\alpha_1, \dots, \alpha_n\}$ of a field L is *algebraically dependent* over a subfield K if there exists a nonzero polynomial P of n variables with coefficients in K such that

$P(\alpha_1, \dots, \alpha_n) = 0$ (*)

Definition: If S is not algebraically dependent, i.e. if there exists no nonzero polynomial P such that (*) holds, then S is *algebraically independent* [16].

In this paper, $L = R(\boldsymbol{p})$ and $K = R$. Thus, S is a subset of polynomials in $R(\boldsymbol{p})$.

Algebraic independence can be tested in the following way. Let the polynomials be $r_1(\boldsymbol{p}), r_2(\boldsymbol{p}), \dots, r_n(\boldsymbol{p})$ and let $\hat{\boldsymbol{r}}$ be a tag variable, i.e. a variable introduced in order to eliminate other variables [17]. Then form the Gröbner Basis of the set $\{r_1(\boldsymbol{p}) - \hat{r}_1, r_2(\boldsymbol{p}) - \hat{r}_2, \dots, r_n(\boldsymbol{p}) - \hat{r}_n\}$ with the ranking $\{p_1, \dots, p_P, \hat{r}_1, \dots, \hat{r}_n\}$. A polynomial in only $\hat{r}_1, \dots, \hat{r}_n$ will result if and only if the set is algebraically dependent. If no such polynomial results, then the set is algebraically independent [17].

8.1 Redundancy

When we consider solving a system of equations $\{\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)\}$, an interesting question arises: do we need every equation $c_i(\boldsymbol{p}) = c_i(\boldsymbol{p}^*)$, or are some of these equations redundant? Redundant means that the solution space of some equation contains the intersection of the solution spaces of other equations in the set. More precisely, suppose $S = \{f_1, \dots, f_m\}$ is a collection of polynomials in n variables $x_1, \dots, x_n$ and suppose a solution to the system of equations $f_1 = 0, f_2 = 0, \dots, f_m = 0$ is

$(a_1, \dots, a_n)$. Let I be the ideal generated by S. Then any element f of the ideal I also satisfies $f(a_1, \dots, a_n) = 0$ [16]. In other words, solving $f = 0$ would be redundant in solving the system $f_1 = 0, f_2 = 0, \dots, f_m = 0$. We now show redundancy can be related to algebraic dependence:

Lemma 1: Suppose there are m algebraically independent terms $f_i(\boldsymbol{p}), g_1(\boldsymbol{p}), \dots, g_{m-1}(\boldsymbol{p})$ and that the terms $f_i(\boldsymbol{p}) - \hat{f}_i,\ g_1(\boldsymbol{p}) - \hat{g}_1,\ g_2(\boldsymbol{p}) - \hat{g}_2, \dots, g_{m-1}(\boldsymbol{p}) - \hat{g}_{m-1}$ are consistent, i.e. solving $\{f_i(\boldsymbol{p}) = \hat{f}_i, \boldsymbol{g}(\boldsymbol{p}) = \hat{\boldsymbol{g}}\}$ does not result in the empty set. Then the set $\{f_i(\boldsymbol{p}) - \hat{f}_i,\ g_1(\boldsymbol{p}) - \hat{g}_1,\ g_2(\boldsymbol{p}) - \hat{g}_2, \dots, g_{m-1}(\boldsymbol{p}) - \hat{g}_{m-1}\}$ contains no redundant elements.

Proof: We prove the contra-positive. Assume that the set $\{f_i(\boldsymbol{p}) - \hat{f}_i,\ g_1(\boldsymbol{p}) - \hat{g}_1,\ g_2(\boldsymbol{p}) - \hat{g}_2, \dots, g_{m-1}(\boldsymbol{p}) - \hat{g}_{m-1}\}$ contains a redundant element. Without loss of generality, assume $f_i(\boldsymbol{p}) - \hat{f}_i$ is contained in the ideal $(g_1(\boldsymbol{p}) - \hat{g}_1,\ g_2(\boldsymbol{p}) - \hat{g}_2, \dots, g_{m-1}(\boldsymbol{p}) - \hat{g}_{m-1})$.

Then if $\boldsymbol{p}' = (p_1', \dots, p_P')$ is any solution to the system $g_1(\boldsymbol{p}) = \hat{g}_1,\ g_2(\boldsymbol{p}) = \hat{g}_2, \dots, g_{m-1}(\boldsymbol{p}) = \hat{g}_{m-1}$, then $\boldsymbol{p}'$ is also a solution to $f_i(\boldsymbol{p}) = \hat{f}_i$, i.e. $f_i(\boldsymbol{p}') - \hat{f}_i = 0$.

We have that $\boldsymbol{p}'$ is a function $\hat{g}_1, \hat{g}_2, \dots, \hat{g}_{m-1}$ and possibly $\boldsymbol{p}$, thus since $\hat{f}_i$ is a tag variable and not a function of $\boldsymbol{p}$, then $\hat{f}_i$ must be only a function of $\hat{g}_1, \hat{g}_2, \dots, \hat{g}_{m-1}$. Since $f_i(\boldsymbol{p}) - \hat{f}_i$ is contained in the ideal $(g_1(\boldsymbol{p}) - \hat{g}_1,\ g_2(\boldsymbol{p}) - \hat{g}_2, \dots, g_{m-1}(\boldsymbol{p}) - \hat{g}_{m-1})$, then $\hat{f}_i$ must be a *polynomial* or *rational* function of $\hat{g}_1, \hat{g}_2, \dots, \hat{g}_{m-1}$. Thus, $f_i(\boldsymbol{p}), g_1(\boldsymbol{p}), g_2(\boldsymbol{p}), \dots, g_{m-1}(\boldsymbol{p})$ are algebraically dependent. ■

Following the proof to Lemma 1, we can show that if $c_i(\boldsymbol{p}) - c_i(\boldsymbol{p}^*)$ is contained in the ideal $\left(q_1(\boldsymbol{p}) - q_1(\boldsymbol{p}^*), \dots, q_M(\boldsymbol{p}) - q_M(\boldsymbol{p}^*)\right)$, then $c_i(\boldsymbol{p})$ must be a *polynomial* or *rational* function of $q_1(\boldsymbol{p}), \dots, q_M(\boldsymbol{p})$ (which means that $c_i(\boldsymbol{p})$ and $q_1(\boldsymbol{p}), \dots, q_M(\boldsymbol{p})$ are algebraically dependent). This statement will be later used to prove that there exists a unique rational reparameterization of $c_i(\boldsymbol{p})$ in terms of $q_1(\boldsymbol{p}), \dots, q_M(\boldsymbol{p})$. We thus state it as a corollary:

Corollary 1: Suppose the $M + 1$ terms $c_i(\boldsymbol{p}) - \hat{c}_i, q_1(\boldsymbol{p}) - \hat{q}_1,\ q_2(\boldsymbol{p}) - \hat{q}_2, \dots, q_M(\boldsymbol{p}) - \hat{q}_M$ are consistent, i.e. solving $\{c_i(\boldsymbol{p}) = \hat{c}_i, q_1(\boldsymbol{p}) = \hat{q}_1,\ q_2(\boldsymbol{p}) = \hat{q}_2, \dots, q_M(\boldsymbol{p}) = \hat{q}_M\}$ does not result in the empty set. If $c_i(\boldsymbol{p}) - \hat{c}_i$ is contained in the ideal $(q_1(\boldsymbol{p}) - \hat{q}_1,\ q_2(\boldsymbol{p}) - \hat{q}_2, \dots, q_M(\boldsymbol{p}) - \hat{q}_M)$, then $\hat{c}_i$ must be a *polynomial* or *rational* function of $\hat{q}_1, \hat{q}_2, \dots, \hat{q}_M$.

Proof: This follows from the proof of Lemma 1. ■

We seek to prove that among the coefficients $\boldsymbol{c}(\boldsymbol{p})$ of the input-output equations, there exists a subset of M algebraically independent coefficients $c_1(\boldsymbol{p}), \dots, c_M(\boldsymbol{p})$. This will follow from the converse to Lemma 1, with an additional assumption:

Assumption: The number of non-redundant equations is equal to the number of non-free parameters in the solution.

In other words, for P parameters and M non-redundant equations, we can eliminate up to $M - 1$ parameters and solve for M non-free parameters in terms of s free parameters. This is a reasonable assumption because we are examining the unidentifiable case, where our parameter vector $\boldsymbol{p}$ is underdetermined with respect to the input-output data. This assumption will be used in Section 9 to link the ideas of algebraic independence of our polynomials to the dimension of our variety.

Converse to Lemma 1: Suppose there are m non-redundant and consistent terms $f_1(\boldsymbol{p}) - \hat{f}_1, \ldots, f_m(\boldsymbol{p}) - \hat{f}_m$, where m is the number of non-free parameters in its solution. Then the m terms $f_1(\boldsymbol{p}), \ldots, f_m(\boldsymbol{p})$ are algebraically independent.

Proof: Assume there are m non-redundant and consistent terms $f_1(\boldsymbol{p}) - \hat{f}_1, \ldots, f_m(\boldsymbol{p}) - \hat{f}_m$, where m is the number of non-free parameters in its solution $\boldsymbol{p}'$, in terms of free parameters and $\hat{f}_1, \ldots, \hat{f}_m$. Let $I = (\, f_1(\boldsymbol{p}) - \hat{f}_1, \ldots, f_m(\boldsymbol{p}) - \hat{f}_m\,)$, taken as an ideal over $R(\boldsymbol{p})$.

For a contradiction, assume that $f_1(\boldsymbol{p}), \ldots, f_m(\boldsymbol{p})$ are algebraically dependent. Then this means there exists a non-zero polynomial g such that $g\big(\boldsymbol{f}(\boldsymbol{p})\big) = 0$.

Then over the quotient ring $R(\boldsymbol{p})/I$, we can substitute $\hat{f}_i$ for $f_i(\boldsymbol{p})$ and obtain a condition:

$$g\big(\hat{\boldsymbol{f}}\big) = g\big(\hat{f}_1, \ldots, \hat{f}_m\big) = 0.$$

There are two possibilities for $g\big(\hat{\boldsymbol{f}}\big)$:

Case 1: $g\big(\hat{\boldsymbol{f}}\big)$ only involves $\hat{f}_i$ that are not in the solution $\boldsymbol{p}'$.

Then the corresponding terms $f_i(\boldsymbol{p}) - \hat{f}_i$ would be redundant since they are not incorporated in the solution $\boldsymbol{p}'$, a contradiction.

Case 2: $g\big(\hat{\boldsymbol{f}}\big)$ involves $\hat{f}_i$ in the solution $\boldsymbol{p}'$.

Then this means that the solution $\boldsymbol{p}'$ has an additional constraint on it, which means there is another way to write the solution using the polynomial relationship $g\big(\hat{\boldsymbol{f}}\big) = 0$. Then this other solution is either inconsistent with $\boldsymbol{p}'$, which means the terms $f_1(\boldsymbol{p}) - \hat{f}_1, \ldots, f_m(\boldsymbol{p}) - \hat{f}_m$ are inconsistent, or one or more of the terms $f_1(\boldsymbol{p}) - \hat{f}_1, \ldots, f_m(\boldsymbol{p}) - \hat{f}_m$ are redundant since the polynomial constraint $g\big(\hat{\boldsymbol{f}}\big) = 0$ implies that some $\hat{f}_i$ can be solved for in terms of other $\hat{\boldsymbol{f}}$ and thus not all $\hat{\boldsymbol{f}}$ are needed in the solution $\boldsymbol{p}'$. This provides a contradiction.

Thus, these two cases prove that the m terms $f_1(\boldsymbol{p}), \ldots, f_m(\boldsymbol{p})$ are algebraically independent. ■

Lemma 1 and the Converse together show that finding an algebraically independent set of M coefficients $c_1(\boldsymbol{p}), \ldots, c_M(\boldsymbol{p})$ is equivalent to finding a set of M non-redundant $c_1(\boldsymbol{p}) - c_i(\boldsymbol{p}^*), \ldots, c_M(\boldsymbol{p}) - c_i(\boldsymbol{p}^*)$ in the exhaustive summary (which is consistent since $\boldsymbol{p} = \boldsymbol{p}^*$ always satisfies these equations), where M is the number of non-free parameters in the solution. Since this is true by assumption (p. 13), then we have shown there exists a set of M algebraically independent coefficients $\boldsymbol{c}(\boldsymbol{p})$.

We now return to the question of how to determine which of the terms $c_i(\boldsymbol{p}) = c_i(\boldsymbol{p}^*)$ in the exhaustive summary are redundant. The above Converse shows that if the set is algebraically dependent, then at least one of the terms in the set is redundant, and thus not all the coefficients are needed to solve the exhaustive summary. In other words, one or more of the equations can be removed, but how to we determine which ones can be removed while retaining the same solution space?

For example, consider the ideal generated by $(p_1p_2 - p_1^*p_2^*,\ p_2p_3 - p_2^*p_3^*,\ p_1p_3 - p_1^*p_3^*,\ p_1p_2p_3 - p_1^*p_2^*p_3^*)$. The set of zeroes of all four polynomials is $\{p_1 = p_1^*, p_2 = p_2^*, p_3 = p_3^*\}$. Using tag variables and finding the Groebner Basis of $(p_1p_2 - \hat{a},\ p_2p_3 - \hat{b},\ p_1p_3 - \hat{c},\ p_1p_2p_3 - \hat{d})$ for the ranking $\{p_1, p_2, p_3, \hat{a}, \hat{b}, \hat{c}, \hat{d}\}$, it can be shown that $\hat{d}^2 = \hat{a}\hat{b}\hat{c}$, thus the terms $\{p_1p_2,\ p_2p_3,\ p_1p_3,\ p_1p_2p_3\}$ are algebraically dependent. However, if we remove the term $p_1p_2p_3 - p_1^*p_2^*p_3^*$ from the set, the set of zeroes is now $\{p_1 = \pm p_1^*, p_2 = \pm p_2^*, p_3 = \pm p_3^*\}$. In other words, even though the new set $\{p_1p_2 - p_1^*p_2^*,\ p_2p_3 - p_2^*p_3^*,\ p_1p_3 - p_1^*p_3^*\}$ is algebraically independent and non-redundant, this does not mean we have the same solution space as before. Notice, however, that if any of the terms $p_1p_2 - p_1^*p_2^*,\ p_2p_3 - p_2^*p_3^*$, or $p_1p_3 - p_1^*p_3^*$ were removed from our original set, the solution set would remain the same. In other words, we conjecture that we can only remove terms that are a polynomial/rational combination of the others, not terms which have a higher power equal to a polynomial/rational combination of the others.

Thus, we want to form a minimal set of $\boldsymbol{c}(\boldsymbol{p})$ needed to generate the same solution space as the original exhaustive summary (3.2). This question has been addressed in [15] as forming a generating set for $\boldsymbol{c}(\boldsymbol{p})$ in the sense of sub-algebras. Any term $c_i(\boldsymbol{p})$ that could be rewritten (using ideal operations) as other elements in $\boldsymbol{c}(\boldsymbol{p})$ was excluded from the generating set [15]. Thus, if there are more than M coefficients of the input-output equations, we only need to choose M coefficients that are not polynomial or rational combinations of each other for the exhaustive summary $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$. Since we have already proven that there exists a set of M algebraically independent coefficients, which is a stronger condition, then forming a set of M coefficients that are not polynomial/rational combinations of each other can be done. As discussed earlier with algebraic dependence, this is easily checked by taking a Gröbner Basis of $\{\boldsymbol{c}(\boldsymbol{p}) - \hat{\boldsymbol{c}}\}$. To simplify notation, from now on we will refer to the exhaustive summary equations $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ as having M equations, i.e. we will restrict $\boldsymbol{c}(\boldsymbol{p})$ to the M elements in its generating set.

8.2 Algebraic independence of $\boldsymbol{q}(\boldsymbol{p})$

We now show that if there exist M algebraically independent $\boldsymbol{c}(\boldsymbol{p})$, then there exists M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$. To prove this, we show that if there were less than M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$, for instance, $M-1$, then these $M-1$ $q_j(\boldsymbol{p})$ elements cannot be algebraically dependent with each of the M original coefficients $c_i(\boldsymbol{p})$ because then the original M coefficients would also be dependent. This contradiction implies that we can adjoin a subset of the algebraically independent $\boldsymbol{c}(\boldsymbol{p})$ to our set of algebraically independent $q_j(\boldsymbol{p})$ to obtain M algebraically independent decoupled combinations.

Theorem 1: Suppose there are (at least) M $\boldsymbol{q}(\boldsymbol{p})$ terms over P parameters obtained from the Gröbner Bases of the exhaustive summary (or from the exhaustive summary itself). Suppose there are exactly M

algebraically independent coefficients $c_i(\boldsymbol{p})$. Then there exist (at least) M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$.

Proof: We want to show that there exists a set of M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$. Assume every set of M elements chosen is algebraically dependent. In other words, the largest set of algebraically independent elements is less than M, say $M-1$ (proof follows similarly if any number less than $M-1$ is chosen).

Assume for a contradiction that these $M-1$ elements $q_j(\boldsymbol{p})$ are algebraically dependent with each of the M algebraically independent coefficients $c_i(\boldsymbol{p})$ (taken individually). Then we have the following polynomials:

$$f_1(q_1(p), q_2(p), \dots, q_{M-1}(p), c_1(p)) = 0$$

$$f_2(q_1(p), q_2(p), \dots, q_{M-1}(p), c_2(p)) = 0$$

$$\dots$$

$$f_M(q_1(p), q_2(p), \dots, q_{M-1}(p), c_M(p)) = 0$$

Where each f_i includes c_i and one or more q_j.

Then since each $f_i, 1 \le i \le M$ must include some $q_j, 1 \le j \le M-1$, then this means the f_i cannot be disjoint, i.e. they overlap in some q_j. Since there are M such f_i, using elimination we can form a polynomial $g\big(c_1(p), c_2(p), \dots, c_M(p)\big) = 0$ [16]. This implies that the $c_i(p)$ are algebraically dependent, thus we have a contradiction.

Thus, these $M-1$ elements $q_j(\boldsymbol{p})$ must be algebraically independent with (at least) one of the M original coefficients, say $c_i(\boldsymbol{p})$. Thus, we adjoin $c_i(\boldsymbol{p})$ to the set of $M-1$ algebraically independent $q_j(\boldsymbol{p})$ to get a total of M algebraically independent decoupled combinations (which is, of course, a simpler set of decoupled combinations than the original $\boldsymbol{c}(\boldsymbol{p})$ we started with). Thus, in the pathological case where there are not M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$, we can adjoin a subset of the M algebraically independent coefficients $\boldsymbol{c}(\boldsymbol{p})$ to get the "simplest set" of algebraically independent decoupled combinations. ∎

As described above, we take the Gröbner Basis of the set $\{q_1(\boldsymbol{p}) - \hat{q}_1, q_2(\boldsymbol{p}) - \hat{q}_2, \dots, q_M(\boldsymbol{p}) - \hat{q}_M\}$ with the ranking $\{p_1, \dots, p_P, \hat{q}_1, \dots, \hat{q}_M\}$ to test if our set of $\boldsymbol{q}(\boldsymbol{p})$ is algebraically independent.

9. Rational Reparameterization

9.1 Solution space

We now seek to prove that $c_i(\boldsymbol{p}) - c_i(\boldsymbol{p}^*)$ is in fact redundant with respect to the ideal generated by $(q_1(\boldsymbol{p}) - q_1(\boldsymbol{p}^*), q_2(\boldsymbol{p}) - q_2(\boldsymbol{p}^*), \dots, q_M(\boldsymbol{p}) - q_M(\boldsymbol{p}^*))$ which will immediately prove $c_i(\boldsymbol{p})$ can be rationally reparameterized over $\boldsymbol{q}(\boldsymbol{p})$. To do this, we examine the solution space generated by the exhaustive summary $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ and the solution space generated by $\{q_1(\boldsymbol{p}) = q_1(\boldsymbol{p}^*), q_2(\boldsymbol{p}) = q_2(\boldsymbol{p}^*), \dots, q_M(\boldsymbol{p}) = q_M(\boldsymbol{p}^*)\}$.

By the $\boldsymbol{p}$-solution (or simply solution space) of a polynomial set of $\{f_1(\boldsymbol{p}), f_2(\boldsymbol{p}), \dots, f_M(\boldsymbol{p})\}$, we mean the set of values of p's where each polynomial vanishes. This is also known as an algebraic set or variety, which we call $V_p(\boldsymbol{f})$.

Definition: A variety $V_p(\boldsymbol{f})$ is *irreducible* if whenever $V_p(\boldsymbol{f})$ is written in the form $V_p(\boldsymbol{f}) = V_1 \cup V_2$ where V_1 and V_2 are varieties, then either $V_p(\boldsymbol{f}) = V_1$ or $V_p(\boldsymbol{f}) = V_2$ [16].

For example, $V_p(p_1 - p_1^*)$ is irreducible, but $V_p\left(p_1^2 - p_1^{*2}\right)$ is not irreducible.

We examine the variety of $\boldsymbol{c}(\boldsymbol{p}) - \boldsymbol{c}(\boldsymbol{p}^*)$, which we call $V_p(\boldsymbol{c})$. Note that $V_p(\boldsymbol{c})$ is an intersection of varieties formed by $c_1(\boldsymbol{p}) - c_1(\boldsymbol{p}^*), c_2(\boldsymbol{p}) - c_2(\boldsymbol{p}^*), \dots, c_M(\boldsymbol{p}) - c_M(\boldsymbol{p}^*)$, i.e. the variety of any $c_i(\boldsymbol{p}) - c_i(\boldsymbol{p}^*)$ contains the variety $V_p(\boldsymbol{c})$. For unidentifiable systems, $V_p(\boldsymbol{c})$ has two forms:

Case 1: the solution can be written as M non-free parameters in terms of s free parameters (described as $\widetilde{\boldsymbol{p}}$) with only 1 solution branch:

$$\{p_1 = f_1(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*), p_2 = f_2(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*), \dots, p_M = f_M(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*)\}$$

Case 2: the solution can be written as M non-free parameters in terms of s free parameters with multiple branches of solutions:

$$\{p_1 = g_1^1(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*), p_2 = g_2^1(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*), \dots, p_M = g_M^1(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*)\}$$

$$\{p_1 = g_1^2(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*), p_2 = g_2^2(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*), \dots, p_M = g_M^2(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*)\}$$

$$\dots$$

$$\{p_1 = g_1^\eta(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*), p_2 = g_2^\eta(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*), \dots, p_M = g_M^\eta(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*)\}$$

Where η equals the number of distinct solutions in $\boldsymbol{p}$. In this case, we describe the solution

$$\{p_1 = g_1^i(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*), p_2 = g_2^i(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*), \dots, p_M = g_M^i(\widetilde{\boldsymbol{p}}, \boldsymbol{p}^*)\}$$

as the sub-variety $V_p^i(\boldsymbol{c})$. So we have that $V_p(\boldsymbol{c}) = \bigcup_{i=1}^{\eta} V_p^i(\boldsymbol{c})$.

Let $\{q_1(\boldsymbol{p}), q_2(\boldsymbol{p}), \dots, q_M(\boldsymbol{p})\}$ be a set of algebraically independent parameter combinations found from decoupled terms/factors in the Gröbner Bases of (3.2). There are also two cases for $q_j(\boldsymbol{p})$:

Case 1: $q_j(\boldsymbol{p})$ appears as an element in a Gröbner Basis, i.e. in the form $q_j(\boldsymbol{p}) - q_j(\boldsymbol{p}^*)$.

Case 2: $q_j(\boldsymbol{p})$ appears as a factor of a Gröbner Basis element, i.e. in the form $\left(q_j(\boldsymbol{p}) - f_{j_1}(\boldsymbol{p}^*)\right)\left(q_j(\boldsymbol{p}) - f_{j_2}(\boldsymbol{p}^*)\right) \dots \left(q_j(\boldsymbol{p}) - f_{j_{N_j}}(\boldsymbol{p}^*)\right)$ where N_j is the multiplicity of $q_j(\boldsymbol{p})$. We know that for each $q_j(\boldsymbol{p})$, one of the $f_{j_\alpha}(\boldsymbol{p}^*)$ must be $q_j(\boldsymbol{p}^*)$ since this represents the trivial solution.

Let the elements $q_j(\boldsymbol{p})$, where $j \in S$ and S is a subset of the indicies $\{1,2,\dots,M\}$, be the elements of $\boldsymbol{q}(\boldsymbol{p})$ which belong to case 2, i.e. those that appear as a factor in a Gröbner Basis. Thus, the elements $q_j(\boldsymbol{p})$, where $j \notin S$, are the elements of $\boldsymbol{q}(\boldsymbol{p})$ which belong to case 1.

We are now going to relate the variety $V_p(\boldsymbol{q})$ to the variety $V_p(\boldsymbol{c})$. The variety $V_p(\boldsymbol{c})$ corresponding to equations $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ is an intersection of varieties formed by its Gröbner Basis elements [16]. If one of the Gröbner Basis elements factorizes non-trivially, then a solution is formed by taking one of the factors and again intersecting it with other elements or factors of other elements in the Gröbner Basis. Thus, the variety generated by each element in a Gröbner Basis contains the variety $V_p(\boldsymbol{c})$. Then this means the variety of a Gröbner Basis element $q_j(\boldsymbol{p}) - q_j(\boldsymbol{p}^*)$ contains the variety $V_p(\boldsymbol{c})$. (*)

For a Gröbner Basis element $\left(q_j(\boldsymbol{p}) - f_{j_1}(\boldsymbol{p}^*)\right)\left(q_j(\boldsymbol{p}) - f_{j_2}(\boldsymbol{p}^*)\right)\dots\left(q_j(\boldsymbol{p}) - f_{j_{N_j}}(\boldsymbol{p}^*)\right)$, the variety of each $q_j(\boldsymbol{p}) - f_{j_\alpha}(\boldsymbol{p}^*)$ factor contains one or more sub-varieties $V_p^i(\boldsymbol{c})$. Since each element of a Gröbner Basis contains the solution space generated by the whole Gröbner Basis, then the union of the varieties of $q_j(\boldsymbol{p}) - f_{j_1}(\boldsymbol{p}^*), \dots, q_j(\boldsymbol{p}) - f_{j_{N_j}}(\boldsymbol{p}^*)$ contains the union of $V_p^i(\boldsymbol{c})$, i.e. $V_p(\boldsymbol{c})$. Thus, for every $V_p^i(\boldsymbol{c})$, there exists some factor $q_j(\boldsymbol{p}) - f_{j_\alpha}(\boldsymbol{p}^*)$ whose variety contains $V_p^i(\boldsymbol{c})$ (for all $j \in S$, for some α). (**)

We call $V_p^k(\boldsymbol{q})$ the sub-variety formed by the variety of $\{q_j(\boldsymbol{p}) - f_{j_\alpha}(\boldsymbol{p}^*)$ for all $j \in S$, for some α, where $1 \le \alpha \le N_j\}$, together with $\{q_j(\boldsymbol{p}) - q_j(\boldsymbol{p}^*)$ for $j \notin S\}$. In other words, $V_p^k(\boldsymbol{q})$ is generated by choosing a factor from elements like $\left(q_j(\boldsymbol{p}) - f_{j_1}(\boldsymbol{p}^*)\right)\left(q_j(\boldsymbol{p}) - f_{j_2}(\boldsymbol{p}^*)\right)\dots\left(q_j(\boldsymbol{p}) - f_{j_{N_j}}(\boldsymbol{p}^*)\right)$ where $j \in S$ and combining it with elements $q_j(\boldsymbol{p}) - q_j(\boldsymbol{p}^*)$ for $j \notin S$, and then finding the algebraic set of zeroes of this set. Here $1 \le k \le \mu$, where μ is the product of the multiplicities of all $q_j(\boldsymbol{p})$, i.e. $\mu = \prod_{j \in S} N_j$. Combining (*) and (**), we get that some $V_p^k(\boldsymbol{q})$ contains $V_p^i(\boldsymbol{c})$.

Loosely speaking, the dimension of a variety is the number of parameters that can vary freely [18]. We employ the dimension of a variety to prove the next two lemmas, using the following fact: If $V_p^k(\boldsymbol{q})$ contains $V_p^i(\boldsymbol{c})$, then the dimension of $V_p^k(\boldsymbol{q})$ is greater than or equal to the dimension of $V_p^i(\boldsymbol{c})$ and thus $V_p^k(\boldsymbol{q})$ has at least as many free parameters as $V_p^i(\boldsymbol{c})$ [16]. Before we examine when this containment becomes equality, we first prove that there are exactly M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$.

Lemma 2: Suppose there are at least M $\boldsymbol{q}(\boldsymbol{p})$ terms over P parameters found from decoupled terms/factors in the Gröbner Bases of the exhaustive summary (3.2). Then there are exactly M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$.

Proof: Theorem 2 showed there are at least M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$. We now show there are at most M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$.

Assume there are more than M algebraically independent parameter combinations $\boldsymbol{q}(\boldsymbol{p})$ in the Gröbner Bases, i.e. there are more than M terms of the form $q_j(\boldsymbol{p}) - q_j(\boldsymbol{p}^*)$ or $q_j(\boldsymbol{p}) - f_{j_\alpha}(\boldsymbol{p}^*)$ where $q_j(\boldsymbol{p})$ are algebraically independent. As described above, there exists a variety $V_p^k(\boldsymbol{q})$ that contains $V_p^i(\boldsymbol{c})$.

Thus, there exists a variety $\tilde{V}_p^k(\boldsymbol{q})$ generated by more than M terms of the form $q_j(\boldsymbol{p}) - q_j(\boldsymbol{p}^*)$ or $q_j(\boldsymbol{p}) - f_{j_\alpha}(\boldsymbol{p}^*)$ that contains $V_p^i(\boldsymbol{c})$. This implies that the dimension of $\tilde{V}_p^k(\boldsymbol{q})$ is greater than or equal to the dimension of $V_p^i(\boldsymbol{c})$. Since there are more than M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$ (by assumption) but exactly M algebraically independent $\boldsymbol{c}(\boldsymbol{p})$, then Lemma 1 implies that there are more non-redundant constraints, thus more non-free parameters that form $\tilde{V}_p^k(\boldsymbol{q})$ than form $V_p^i(\boldsymbol{c})$. Thus $\tilde{V}_p^k(\boldsymbol{q})$ is expressed in terms of fewer free parameters than $V_p^i(\boldsymbol{c})$. Thus the dimension of $\tilde{V}_p^k(\boldsymbol{q})$ is strictly less than the dimension of $V_p^i(\boldsymbol{c})$, a contradiction. Thus, there are exactly M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$. ■

Lemma 3: When the solution to (3.2) can be written as M non-free parameters in terms of s free parameters with only 1 solution branch (case 1), then $V_p(\boldsymbol{q})$ equals $V_p(\boldsymbol{c})$.

When the solution to (3.2) can be written as M non-free parameters in terms of s free parameters with multiple branches of solutions (case 2), some union of η different sub-varieties $V_p^k(\boldsymbol{q})$ equals the union of sub-varieties $V_p^i(\boldsymbol{c})$ for $1 \le i \le \eta$.

Proof:

First we examine case 1, where there is a single solution branch where non-free parameters can be written in terms of the free parameters in $V_p(\boldsymbol{c})$. As mentioned above, the variety $V_p(\boldsymbol{q})$ contains the variety $V_p(\boldsymbol{c})$ since the variety of each element $q_j(\boldsymbol{p}) - q_j(\boldsymbol{p}^*)$ contains the solution to $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$. By Lemma 2, the ideal generated by $\boldsymbol{q}(\boldsymbol{p}) - \boldsymbol{q}(\boldsymbol{p}^*)$ is generated by M algebraically independent elements, thus by Lemma 1, none of the M elements that generate the ideals are redundant. Likewise, there are M non-redundant generators in $\boldsymbol{c}(\boldsymbol{p}) - \boldsymbol{c}(\boldsymbol{p}^*)$. This means the solutions to $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ and $\boldsymbol{q}(\boldsymbol{p}) = \boldsymbol{q}(\boldsymbol{p}^*)$ can be solved for (over the complex numbers) in terms of M non-free parameters in s free parameters [16]. This means both varieties $V_p(\boldsymbol{q})$ and $V_p(\boldsymbol{c})$ are spanned by s free parameters. Thus, since each of the varieties can be parameterized in terms of s free parameters, then the dimensions for each of these varieties is the same. Since the variety $V_p(\boldsymbol{q})$ is formed by taking irreducible elements in the Gröbner Bases, then the varieties $V_p(\boldsymbol{q})$ and $V_p(\boldsymbol{c})$ themselves are irreducible. Thus, the fact that $V_p(\boldsymbol{q})$ contains $V_p(\boldsymbol{c})$ but the dimensions are the same implies that $V_p(\boldsymbol{q})$ equals $V_p(\boldsymbol{c})$ [19].

Next, we examine case 2, where there are multiple branches of solutions where the non-free parameters are written in terms of the free parameters in $V_p(\boldsymbol{c})$. As mentioned above, some $V_p^k(\boldsymbol{q})$ contains $V_p^i(\boldsymbol{c})$ since for every $V_p^i(\boldsymbol{c})$, there exists some factor $q_j(\boldsymbol{p}) - f_{j_\alpha}(\boldsymbol{p}^*)$ whose variety contains $V_p^i(\boldsymbol{c})$ (for all $j \in S$, for some α) and the variety of each element $q_j(\boldsymbol{p}) - q_j(\boldsymbol{p}^*)$ contains $V_p^i(\boldsymbol{c})$ (for $j \notin S$). We want to show that some $V_p^k(\boldsymbol{q})$ will result in some $V_p^i(\boldsymbol{c})$. Again, these varieties are irreducible since they are formed by taking irreducible factors. Again, the dimensions of $V_p^k(\boldsymbol{q})$ and $V_p^i(\boldsymbol{c})$ must be the same due to the number of free parameters, thus some $V_p^k(\boldsymbol{q})$ containing $V_p^i(\boldsymbol{c})$ implies that some $V_p^k(\boldsymbol{q})$ equals $V_p^i(\boldsymbol{c})$ [19]. Thus, some union of η different $V_p^k(\boldsymbol{q})$ equals $V_p(\boldsymbol{c})$. ■

Thus we have shown that the $\boldsymbol{p}$-solution space generated by $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ is the same as the $\boldsymbol{p}$-solution space generated by a subset of the union of the solutions for M $\boldsymbol{q}(\boldsymbol{p})$ terms. Mathematically, this is interesting because it means that our space $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ can be represented by the solution spaces associated with the simpler combinations $\boldsymbol{q}(\boldsymbol{p})$ instead. Thus, even though $\boldsymbol{q}(\boldsymbol{p})$ may not all come from a single Gröbner Basis, the decoupled terms of the form $q_j(\boldsymbol{p}) - q_j(\boldsymbol{p}^*)$ or $q_j(\boldsymbol{p}) - f_{j_\alpha}(\boldsymbol{p}^*)$ still behave like a basis for the ideal generated by the exhaustive summary.

9.2 Reparameterization of $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$

In [7], we showed that when a rational reparameterization of $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$ exists, then the ideal generated by $\boldsymbol{c}(\boldsymbol{p}) - \boldsymbol{c}(\boldsymbol{p}^*)$ is congruent to the ideal generated by $\boldsymbol{q}(\boldsymbol{p}) - \boldsymbol{q}(\boldsymbol{p}^*)$, i.e. that $V_p(\boldsymbol{q})$ equals $V_p(\boldsymbol{c})$. Now we show the converse is also true.

We show that Lemma 3 implies that each $c_i(\boldsymbol{p}) - c_i(\boldsymbol{p}^*)$ is redundant with respect to the ideal generated by $(q_1(\boldsymbol{p}) - q_1(\boldsymbol{p}^*), q_2(\boldsymbol{p}) - q_2(\boldsymbol{p}^*), \ldots, q_M(\boldsymbol{p}) - q_M(\boldsymbol{p}^*))$. Then by Corollary 1 we have that each $c_i(\boldsymbol{p})$ is an rational combination of $\boldsymbol{q}(\boldsymbol{p})$. Thus we always have a rational reparameterization of $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$.

Theorem 2: Assume there exists a set of M algebraically independent decoupled combinations (called $\boldsymbol{q}(\boldsymbol{p})$), found by using the Gröbner Bases of (3.2) and the original $\boldsymbol{c}(\boldsymbol{p})$. Then there exists a unique rational reparameterization, $\tilde{\boldsymbol{c}}(\boldsymbol{q})$, of $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$.

Proof: Lemma 3 implies that:

$c_i(\boldsymbol{p}) - c_i(\boldsymbol{p}^*)$ is contained in the ideal generated by $(q_1(\boldsymbol{p}) - q_1(\boldsymbol{p}^*), q_2(\boldsymbol{p}) - q_2(\boldsymbol{p}^*), \ldots, q_M(\boldsymbol{p}) - q_M(\boldsymbol{p}^*))$ (***)

since $V_p(\boldsymbol{q})$ equals $V_p(\boldsymbol{c})$ in case 1 or some $V_p^k(\boldsymbol{q})$ equals some $V_p^i(\boldsymbol{c})$ in case 2.

From Corollary 1, if $c_i(\boldsymbol{p}) - \hat{c}_i$ is contained in the ideal generated by $(q_1(\boldsymbol{p}) - \hat{q}_1, q_2(\boldsymbol{p}) - \hat{q}_2, \ldots, q_M(\boldsymbol{p}) - \hat{q}_M)$, then $\hat{c}_i$ equals a polynomial or rational function of $\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_M$. Applying this to (***), we have that $\hat{c}_i$ is a polynomial or rational function of $\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_M$, or in other words, each coefficient $c_i(\boldsymbol{p})$ is equal to a rational combination of $\boldsymbol{q}(\boldsymbol{p})$. This reparameterization is unique, since if there were two distinct reparameterizations $\tilde{\boldsymbol{c}}(\boldsymbol{q})$ and $\hat{\boldsymbol{c}}(\boldsymbol{q})$, then since $\boldsymbol{c}(\boldsymbol{p}) = \tilde{\boldsymbol{c}}(\boldsymbol{q}(\boldsymbol{p})) = \hat{\boldsymbol{c}}(\boldsymbol{q}(\boldsymbol{p}))$, this implies dependence amongst the $\boldsymbol{q}(\boldsymbol{p})$, a contradiction. ■

To find the rational reparameterization of $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$, one finds the Gröbner Basis of $\{c_i(\boldsymbol{p}) - \hat{c}_i, q_1(\boldsymbol{p}) - \hat{q}_1, \ldots, q_M(\boldsymbol{p}) - \hat{q}_M\}$ over the ranking $\{p_1, \ldots, p_P, \hat{q}_1, \ldots, \hat{q}_M, \hat{c}_i\}$ for each coefficient $c_i(\boldsymbol{p})$. As discussed in [7], a linear polynomial $f(\hat{q}_1, \ldots, \hat{q}_M) - g(\hat{q}_1, \ldots, \hat{q}_M)\hat{c}_i$ will result.

10. Global or Local Identifiability

Theorem 3: Assume there exists a unique rational reparameterization, $\tilde{\boldsymbol{c}}(\boldsymbol{q})$, of $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$. Then $\boldsymbol{q}$ is either globally or locally identifiable.

Proof:

Since each $\boldsymbol{q}(\boldsymbol{p})$ has at most a finite number of solutions in the Gröbner Basis of (3.2), then solving $\tilde{\boldsymbol{c}}(\boldsymbol{q}) = \tilde{\boldsymbol{c}}(\boldsymbol{q}^*)$ gives that each $\boldsymbol{q} = \boldsymbol{q}(\boldsymbol{p})$ is either globally or locally identifiable, depending on whether each $q_j(\boldsymbol{p})$ has a unique (case 1) or finite number (case 2) of solutions. In other words, if the $\boldsymbol{q}(\boldsymbol{p})$ only appeared as elements $q_j(\boldsymbol{p}) - q_j(\boldsymbol{p}^*)$ in the Gröbner Bases, then global identifiability results, and if at least one $q_j(\boldsymbol{p})$ appears as a factor $q_j(\boldsymbol{p}) - f_{j_\alpha}(\boldsymbol{p}^*)$ in a Gröbner Basis, then local identifiability results. ■

This means we can take for granted that a set of algebraically independent $\boldsymbol{q}(\boldsymbol{p})$ are identifiable and thus the $\tilde{\boldsymbol{c}}(\boldsymbol{q})$ reparameterization step mentioned is only a mathematical formality, as predicted in [7].

A useful interpretation of this theorem is that the identifiable combinations $\boldsymbol{q}(\boldsymbol{p})$ thus form a "basis" of all identifiable combinations, in that they are the simplest polynomial or rational functions that are globally or locally identifiable.

We summarize the results of this paper in a final theorem:

Theorem 4: Suppose we have an unidentifiable model of the form (2.1) and let the exhaustive summary (3.2) be described in terms of M non-redundant equations, where M is the number of non-free parameters in the solution to (3.2). Then the algorithm described in Section 5 can be used to find a set of M algebraically independent identifiable combinations $\boldsymbol{q}(\boldsymbol{p})$.

Proof: The Gröbner Bases of the exhaustive summary, in conjunction with the exhaustive summary, generate at least M decoupled parameter combinations $\boldsymbol{q}(\boldsymbol{p})$. We assume that the exhaustive summary is described in terms of M non-redundant equations, where M is the number of non-free parameters in its solution. The Converse to Lemma 1 proves there are M algebraically independent combinations $\boldsymbol{c}(\boldsymbol{p})$ in the input-output equations. Then Theorem 1 implies there are at least M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$ and Lemma 2 implies there are at most M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$, thus there are exactly M algebraically independent $\boldsymbol{q}(\boldsymbol{p})$. Theorem 2 implies there is a unique rational reparameterization of $\boldsymbol{c}(\boldsymbol{p})$ over $\boldsymbol{q}(\boldsymbol{p})$, call it $\tilde{\boldsymbol{c}}(\boldsymbol{q})$. Finally, Theorem 3 implies the parameter combinations $\boldsymbol{q}(\boldsymbol{p})$ are either globally or locally identifiable.

11. Conclusion

In this paper, we have demonstrated our algorithm on a nonlinear model and have proven that this procedure has a firm theoretical foundation. We have shown that the Gröbner Bases formed from the exhaustive summary (in conjunction with the exhaustive summary) can be used to provide a set of M 'simplest' algebraically independent parameter combinations $\boldsymbol{q}(\boldsymbol{p})$ to uniquely reparameterize the coefficients of the input-output equations as rational terms. These parameter combinations are found by searching for "decoupled" terms or factors in the Gröbner Bases of the exhaustive summary. A unique rational reparameterization of the coefficients over these parameter combinations immediately implies global identifiability when decoupled terms are used and local identifiability when decoupled factors are used. We have thus provided a class of nonlinear models for which identifiable parameter combinations can be used to rationally reparameterize the model in a "canonical form". One practical consequence of this work is the result that when seeking $\boldsymbol{q}(\boldsymbol{p})$ one need not only consider those arising from a single Gröbner Basis, but can consider $\boldsymbol{q}(\boldsymbol{p})$ arising from any ordering. Future work consists of developing efficient procedures for finding these algebraically independent identifiable combinations among the large number, $P!$, of possible Gröbner Bases.

References

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

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

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

[4] 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 (1998) 21-41.

[5] 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 (2001) 35-44.

[6] F. Boulier, Differential Elimination and Biological Modelling, Radon Series Comp. Appl. Math 2 (2007) 111-139.

[7] N. Meshkat, M. Eisenberg, and J. J. DiStefano III, An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner Bases, Math. Biosci. 222 (2009) 61-72.

[8] M. P. Saccomani, S. Audoly, G. Bellu, and L. D'Angiò, 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 (2001) 3108-3113.

[9] 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 (1980) 830-833.

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

[11] F. Ollivier, Le problème de l'identifiabilité structurelle globale: étude théoritique, méthodes effectives et bornes de complexité, PhD thesis, Ecole Polytechnique, 1990.

[12] F. Ollivier, Identifiabilité des systèmes, Technical Report 97-04, GAGE, Ecole Polytechnique, 1997, pp. 43-54.

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

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

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

[16] D. Cox, J. Little, and D. O'Shea, Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3rd Edition, Springer, 2007.

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

[18] K. Ueno, An Introduction to Algebraic Geometry, American Mathematical Society, 1995.

[19] I. Shafarevich, Basic Algebraic Geometry 1, Second Edition, Springer-Verlag, 1988.

Figure Legends

Figure 1. Nonlinear 2-Compartment Model

Figure 2. Reparameterized Nonlinear 2-Compartment Model