A new method for a priori practical identifiability

Peter Thompson, Benjamin Jan Andersson, Nicolas Sundqvist, and Gunnar Cedersund

Abstract

Background and objective: Practical identifiability analysis – to determine whether a model property can be determined from given data – is central to model-based data analysis in biomedicine. The main approaches used today all require that the coverage of the parameter space be exhaustive, which is usually not possible. An attractive alternative could be to use structural identifiability methods, since they do not need such a parameter coverage. However, current structural methods are unsuited for practical identifiability analysis, since they assume that all higher-order derivatives of the measured variables are available, and we do not know the implications of this assumption. Herein, we provide new definitions and methods that allow for this assumption to be relaxed.

Methods: The new methods and definitions are valid for ordinary differential equations, and use a combination of differential algebra and modulus calculus, implemented in Maple.

Results: We introduce the concept of $(\nu_1, ..., \nu_m)$-identifiability, which differs from previous definitions in that it assumes that only the first ν_i derivatives of the measurement signal y_i are available. This new type of identifiability can be determined using our new algorithm, as is demonstrated by applications to various published biomedical models. Our methods allow for identifiability of not only parameters, but of any model property, i.e. observability. These new results provide further strengthening of conclusions made in previous analysis of these models. Importantly, our analysis can for the first time quantify the impact of the assumption that all derivatives are available in specific examples. If one e.g. assumes that only up to third order derivatives, instead of all derivatives, are available, the number of identifiable parameters drops from 17 to 1 for a Drosophila model, and from 21 to 6 for an NF-κB model. In both these models, the previously obtained identifiability is present only if at least 20 derivatives of all measurement signals are available.

Conclusions: Our results demonstrate that the assumption regarding availability of derivatives done in traditional structural identifiability analysis causes a big overestimation regarding the number of parameters that can be estimated. Our new methods and algorithms allow for this assumption to be relaxed, which moves structural identifiability methodology one step closer to practical identifiability analysis.

1 Background

Biology and medicine involve the study of a myriad of time-dependent variables, which cross-talk with each other in complex networks. To deal with this complexity, one often makes use of mechanistic mathematical models, often formulated using ordinary differential equations (ODEs). These ODEs represent hypotheses regarding the biological mechanisms that can explain given biological data [1]. The model-based data analysis of such real data revolves around a correct uncertainty analysis of parameters and predictions: to what degree can one determine the parameters and predictions of interest? This uncertainty analysis is formally called identifiability analysis, if it deals with parameter uncertainty, and observability analysis, if it deals with other model properties, such as states and composite variables [2]. Both observability and identifiability analysis can be done in many different ways, but they are generally subdivided into two types of approaches: structural and practical.

Practical identifiability analysis revolves around the specific data one has collected in a specific situation ([2]). In other words, practical identifiability analysis looks at the real situation, where the data has noise,

and where the experiment might not have been done in an optimal way to generate information about all parameters. Historically, one has often used sensitivity-based methods to study practical identifiability, such as analysis of the Fisher Information Matrix ([3]), and the covariance of the parameters ([4]). In the last 10 years, these methods have been increasingly replaced by Markov Chain Monte Carlo (MCMC) and profile-likelihood based methods ([5], [2], [4]). The benefit of these more modern methods is that they study global properties, and do not have to assume that the system is identifiable to assess identifiability ([3], [2]). These methods are highly useful and flexible, but they still have limitations. Most importantly, they rely on the exhaustive coverage of the entire parameter space, either using the MCMC sampling or using the optimization step in profile likelihood. For large models, this optimization step is difficult and is time-consuming even for an expert in the field ([6], [7]). This difficulty will always leave the analysis results with an inherent doubt: if the coverage of the parameter space fails, the degree of identifiability (the calculated confidence intervals) is not correct.

Structural identifiability analysis has historically been the mirror image of practical identifiability analysis. Structural methods normally do not take a specific data set or the noise level into account. Instead structural identifiability analysis only considers the form of the equations, including the measurement equations. Furthermore, structural identifiability analysis does not rely on an exhaustive coverage of the entire parameter space. Instead, structural identifiability analysis makes use of powerful theories from differential algebra and differential geometry to investigate whether there are structural limits for whether a parameter or a model property ever can be determined from a given measurement. Methods have been developed to determine which states and parameters are locally identifiable ([8]) and globally identifiable ([9], [10]), to discover parameter combinations that are identifiable ([11], [6], [12], [13], [14]), and to suggest modifications in the modeling process based on identifiability ([15], [16]). In the last two decades, powerful methods have been developed that can do such an analysis also for relatively large biological models, featuring 50 to 100 states and parameters ([16]).

Because structural identifiability now can be applied to real, large models, it would be highly interesting if they also could perform practical identifiability analysis, where the limitations of a specific dataset are taken into account. The probably most prominent such assumption is that current methods assume that all derivatives of all measurement signals can be estimated from the data. In practical situations, with real data, this assumption is never fulfilled. It is therefore a critical flaw that we are lacking definitions and methods to analyze the consequences of this unfulfilled assumption.

In summary, there is therefore a need to develop new methods that combine the strengths of the two approaches: A method that does not rely on too idealistic assumptions regarding the real situations, but that also does not rely on optimization and coverage of the entire parameter space. However, such methods have not yet been proposed. Herein, we introduce a new type of practical *and* structural identifiability, which puts an upper limit on how many times each measurement signal can be differentiated. We also introduce a new algorithm that can calculate this new type of identifiability, and apply this to a series of published models. For all studied examples, traditional structural identifiability methods have widely overestimated the number of identifiable parameters: with a maximum of 3 derivatives available, more than 90% of the previously identifiable parameters lose their identifiability.

2 Results and Discussion

2.1 Problem: there is an upper limit for how many derivatives can be estimated from biological data

There is usually an upper limit for the number of derivatives of an output signal that can be estimated from data. This is illustrated in the following simulation. Consider the system

$$x_1' = \theta x_2, \quad x_2' = -\theta x_1 + 1, \quad y = x_1$$

where x_1 and x_2 are state variables, θ is a constant parameter, and y is an observable output variable; that is y is measurable by experiments but x_1, x_2 and θ are unknown. Suppose that, unbeknownst to the observer, $\theta = 1$, $x_1(0) = 10$, and $x_2(0) = 0$. Consequently the true signal of y is given by $y(t) = 1 + 9\cos(t)$. However, in practice, the measured signal will not show a smooth curve but rather a discrete time series that also has added noise. An illustration of this measurement signal can be found in Figure 1A, where 10 realizations of the function $y(t) + N(0, \sigma = 2)$, where the second term represents normally distributed noise with mean 0 and standard deviation 2, are depicted alongside the true function $y(t)$ (Figure 1A). The number of derivatives that can accurately be estimated can be evaluated by deriving the function of a polynomial interpolation of these data points and comparing these estimated derivatives to the derivatives of the true function (Figure 1B). More specifically, this example shows a 6th degree polynomial interpolation for t-values ranging from 0 to 6 in increments of 0.5. Subsequently, the 0th to the 5th derivatives of these interpolations are then compared to the derivatives of the true function $y(t)$. The number of derivatives that can accurately be estimated is dependent on the magnitude of noise in the system. More specifically, the magnitude of the added noise increases the number of accurately estimated derivatives declines (Figure 1C). To illustrate this dependency, the example above was repeated with the added noise sigma ranging from 0.1 to 10 in increments of 0.1 ($\sigma = [0.1, 0.2, \ldots, 10]$) and with 1000 data realizations generated at each noise level. Figure 1C shows the average number of derivatives that could accurately be estimated at each noise level. A derivative is considered accurate if the square root of the sum of squared residuals does not exceed the amplitude of y. From this figure, it is clear that the number of derivatives that can accurately be estimated diminishes rapidly as the signal-to-noise ratio approaches 1.

The following example illustrates how this lack of availability of all derivatives of y is associated with an unrealistic assessment of whether parameters are identifiable.

Example 2.1. Consider the system

$$\begin{cases} x_1' = \theta_1 x_2 \\ x_2' = -\theta_1 x_1 \\ x_3' = \theta_1 \theta_2 \\ y = x_1 + x_3, \end{cases}$$

where x_1, x_2, and x_3 are (unobserved) state variables, θ_1 and θ_2 are constant parameters, and y is an (observed) output variable.

Running the algorithm from [8] shows that x_1, x_2, x_3, θ_1, and θ_2 are locally identifiable. Let us attempt to estimate θ_1 from observation of y. A relation among θ_1 and y over $\mathbb{C}$ is $\theta_1^2 y'' + y^{(4)} = 0$. Thus we can estimate θ_1 by using our observed values of y'' and $y^{(4)}$ at some time t_0:

$$\theta_1 = \pm\sqrt{\frac{-y^{(4)}(t_0)}{y''(t_0)}}.$$

In practice, the measured signal y may be too noisy to give an accurate estimation of $y^{(4)}$. In such a situation, θ_1 is not identifiable from the given data, even though a structural analysis says that it is. Suppose now that

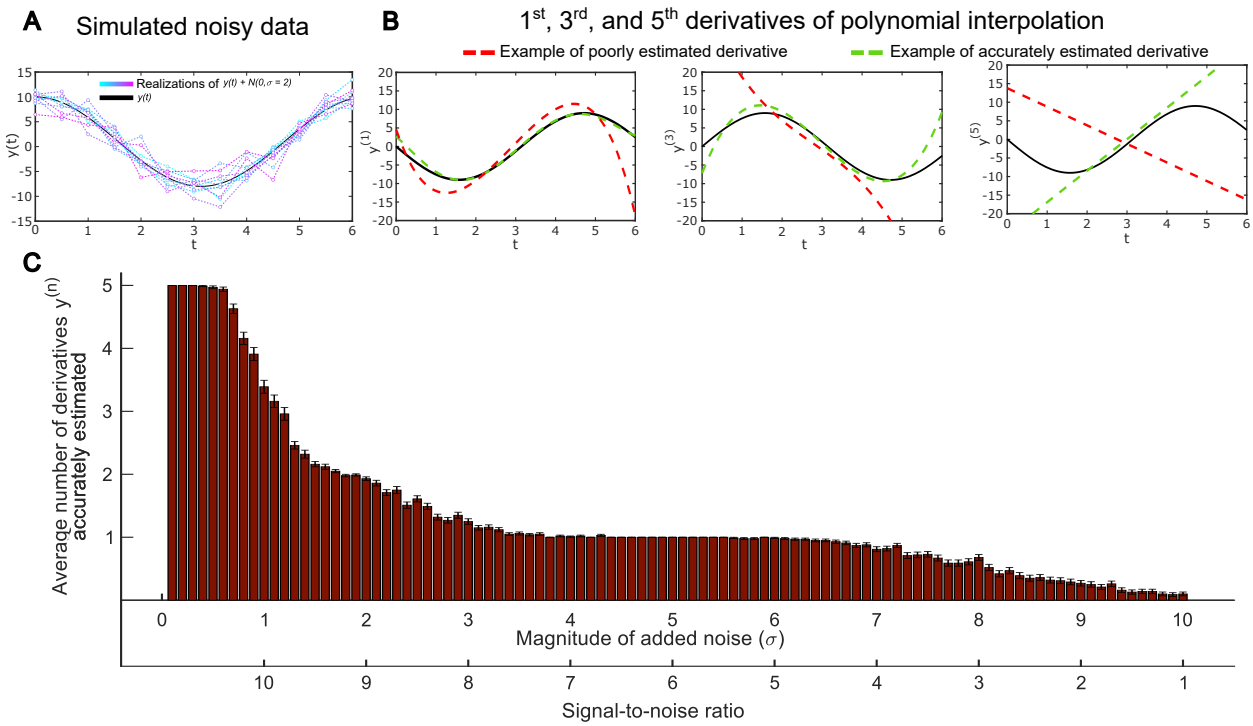

Figure 1: Illustration of the upper limit for how many derivatives can be estimated from noisy data. A: shows examples of 10 different realizations of noisy data (colored lines) and the true function $y(t) = 1 + 9\cos(t)$ (back line) **B:** Illustrations of the 1st, 3rd, and 5th derivatives of the function $y(t)$ with examples of a poorly estimated derivative (red line) and a accurately estimated derivative (green line). **C:** The relationship between the average number of accurately estimated derivatives and the magnitude of the added noise σ/the signal-to-noise ratio. The bars show the mean number of accurately estimated derivatives at each noise level and the error bars show the standard error of the mean (SEM).

we know that our measurements are good enough to give reliable estimates of y'' but of no higher derivative of y. To use structural methods to determine the degree of identifiability in such a situation, we need new definitions and methods.

Remark 2.2. Our approach assumes measurements of several derivatives of the output signal are made at a single time. In practice it is more common to measure the 0-th derivative of the output signal at several times. Having measurements with some uncertainty of $y(t_0),\dots,y(t_N)$ is equivalent to having measurements with some uncertainty of $y(t_0),\dots,y^{(N)}(t_0)$. Hence this aspect of our approach does not limit its generality.

For Example 2.1, one can estimate θ_1 using the relation $y''+\theta_1^2 y-\theta_1^3\theta_2 t-\theta_1^2 x_3(0)=0$ together with measurements of y and y'' at three different times. However knowledge of y,y',y'' at three times can be used to estimate $y,y',y'',y^{(3)},y^{(4)}$ at a single time.

2.2 A new definition for practical and structural identifiability: $(\nu_1,\dots,\nu_m)$-*identifiable*

We now provide an intuitively straightforward definition that is useful for understanding the new methods. In Section 5, we provide a precise definition, together with analytical proofs of all stated method properties.

Let ℓ, n, m, and r be positive integers. Let $\boldsymbol{\theta}=(\theta_1,\dots,\theta_\ell)$, $\mathbf{x}=(x_1,\dots,x_n)$, $\mathbf{y}=(y_1,\dots,y_m)$, and $\mathbf{u}=(u_1,\dots,u_r)$. Let $\mathbf{f}=(f_1,\dots,f_n)$ and $\mathbf{g}=(g_1,\dots,g_m)$ be tuples of rational functions in $\mathbf{x}$, $\mathbf{u}$, and $\boldsymbol{\theta}$ over $\mathbb{C}$. This setup determines a class of systems of ODEs with initial conditions, or a model class:

$$\Sigma(\mathbf{x}^*,\boldsymbol{\theta},\mathbf{u})=\begin{cases}\dfrac{d\mathbf{x}}{dt}=\mathbf{f}(\mathbf{x},\boldsymbol{\theta},\mathbf{u}),\\ \mathbf{y}=\mathbf{g}(\mathbf{x},\boldsymbol{\theta},\mathbf{u}),\\ \mathbf{x}(0)=\mathbf{x}^*,\end{cases}\tag{1}$$

where $\mathbf{x}$ are state variables, which usually correspond to time-varying concentrations or physical properties; $\mathbf{x}^*$ are the initial values of the state variables; $\boldsymbol{\theta}$ are constant parameters, e.g. corresponding to rate constants or volumes; $\mathbf{u}$ are input variables, which are observed and typically controlled by the experimentalist; and $\mathbf{y}$ are output variables, which are observed experimentally. With these notations in place, the new definition is roughly given by:

Definition 2.3 (roughly stated)**.** Let $\nu_1,\dots,\nu_m$ be non-negative integers. Rational expression $h\in\mathbb{C}(\mathbf{x},\boldsymbol{\theta})$ is said to be $(\nu_1,\dots,\nu_m)$-*identifiable* if perfect knowledge of the first ν_i derivatives of y_i at a particular time allows us to determine the value of h at that point in time, up to a finite set.

This definition, which is more technically stated in Definition 5.4, provides the essence of the new definition: given only a limited set of derivatives of the different measurement signals, can we determine the parameter θ_i to be locally identifiable, i.e. as belonging to a finite set? This definition is analogous to traditional definitions of structural identifiability, with the only addition being that we now only have access to a finite set of derivatives. In other words, in traditional structural identifiability analysis, no such constraints on the number of derivatives available is present. This new definition is closer to practical identifiability since, with real biological data, one can only estimate a few derivatives from the measurement data.

2.3 Solution: our new algorithms for calculating practical structural identifiability

For most models, it would be difficult to prove the $(\nu_1,\dots,\nu_m)$-identifiability of a given parameter directly from the definition. We provide Algorithm 1, which can do this using only straightforward algebraic computations. The algorithm applies to any rational combination of states and parameters, of which an individual parameter is a special case.

Algorithm 1: Determine whether a given element of $\mathbb{C}(\mathbf{x},\boldsymbol{\theta})$ is $\boldsymbol{\nu}$-identifiable

Input : Equations $\mathbf{x}' = \mathbf{f}(\mathbf{x},\boldsymbol{\theta},\mathbf{u}),\ \mathbf{y} = \mathbf{g}(\mathbf{x},\boldsymbol{\theta},\mathbf{u})$
non-negative integers $\nu_1,\ldots,\nu_m$
rational expression $h \in \mathbb{C}(\mathbf{x},\boldsymbol{\theta})$

Output: "Yes" if h is $(\nu_1,\ldots,\nu_m)$-identifiable
"No" if h is not $(\nu_1,\ldots,\nu_m)$-identifiable

Step 1: Write the matrix

$$J := \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} & \frac{\partial y_1}{\partial \theta_1} & \cdots & \frac{\partial y_1}{\partial \theta_\ell} \\ \frac{\partial y_1'}{\partial x_1} & \cdots & \frac{\partial y_1'}{\partial x_n} & \frac{\partial y_1'}{\partial \theta_1} & \cdots & \frac{\partial y_1'}{\partial \theta_\ell} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial y_1^{(\nu_1)}}{\partial x_1} & \cdots & \frac{\partial y_1^{(\nu_1)}}{\partial x_n} & \frac{\partial y_1^{(\nu_1)}}{\partial \theta_1} & \cdots & \frac{\partial y_1^{(\nu_1)}}{\partial \theta_\ell} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} & \frac{\partial y_m}{\partial \theta_1} & \cdots & \frac{\partial y_m}{\partial \theta_\ell} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial y_m^{(\nu_m)}}{\partial x_1} & \cdots & \frac{\partial y_m^{(\nu_m)}}{\partial x_n} & \frac{\partial y_m^{(\nu_m)}}{\partial \theta_1} & \cdots & \frac{\partial y_m^{(\nu_m)}}{\partial \theta_\ell} \end{pmatrix}.$$

Step 2: Compute the rank of J.

Step 3: Write the matrix J_h, obtained by adding the row

$$\begin{pmatrix} \frac{\partial h}{\partial x_1} & \cdots & \frac{\partial h}{\partial x_n} & \frac{\partial h}{\partial \theta_1} & \cdots & \frac{\partial h}{\partial \theta_\ell} \end{pmatrix}$$

to J. Then calculate the rank of J_h.

If $\operatorname{rank}(J_h) > \operatorname{rank}(J)$, conclude that h is not $(\nu_1,\ldots,\nu_m)$-identifiable. (Output "No").

If $\operatorname{rank}(J_h) = \operatorname{rank}(J)$, conclude that h is $(\nu_1,\ldots,\nu_m)$-identifiable. (Output "Yes").

Algorithm 2: Determine whether a given element of $\mathbf{x} \cup \boldsymbol{\theta}$ is $\boldsymbol{\nu}$-identifiable

Input : Equations $\mathbf{x}' = \mathbf{f}(\mathbf{x},\boldsymbol{\theta},\mathbf{u}),\ \mathbf{y} = \mathbf{g}(\mathbf{x},\boldsymbol{\theta},\mathbf{u})$
non-negative integers $\nu_1,\ldots,\nu_m$
choice of $\theta_i \in \boldsymbol{\theta}$ (resp. $x_i \in \mathbf{x}$)

Output: "Yes" if θ_i (resp. x_i) is $(\nu_1,\ldots,\nu_m)$-identifiable
"No" if θ_i (resp. x_i) is not $(\nu_1,\ldots,\nu_m)$-identifiable

Step 1: Write the matrix J as in Algorithm 1.

Step 2: Compute the rank of J.

Step 3: Remove the column corresponding to θ_i (resp. x_i) and compute the rank of the resulting matrix.

If the rank has decreased, conclude that θ_i (resp. x_i) is $(\nu_1,\ldots,\nu_m)$-identifiable. (Output "Yes")

If the rank is the same, conclude that θ_i (resp. x_i) is not $(\nu_1,\ldots,\nu_m)$-identifiable. (Output "No")

In the case where one is interested in just an individual parameter or state, one can perform Algorithm 2, where one column is removed instead of a row added, resulting in a smaller matrix.

The precise definition of $\boldsymbol{\nu}$-identifiability is given in Definition 5.4, and we prove that Algorithms 1 and 2 are correct in section 5.2.

The symbolic calculations can be impractically slow for all but the smallest models. Therefore we have also provided probabilistic analogs of these algorithms that are much faster in practice. These are presented and their correctness is proven in section 5.3.

2.4 Intuitive understanding of the new algorithm

Each row of J represents how an output, or one of its derivatives, varies with respect to small changes in each state and parameter. If the last row of J_h is a linear combination of the other rows, then the variation of h can be accounted for as a combination of the variations of the output derivatives. This, in turn, means that it seems like h can be expressed in terms of the known signals, i.e. that h is identifiable. Conversely, if the last row of J_h is linearly independent of the other rows, then the variation of h cannot be explained in terms of that of the available derivatives of the outputs, and one should not expect to be able to determine h from the available outputs and their derivatives.

2.5 Examples

2.5.1 Algorithm 1 determines identifiability of rational quantities, which is important in hypothesis testing

Algorithm 1 (and its faster probabilistic version Algorithm 3) can be used to analyze so-called core predictions (see [2]). A core prediction is a well-determined property, i.e. a model prediction with a small uncertainty. Such core predictions are therefore often tested in future experiments, and are a central part of model-based hypothesis testing. Algorithm 1 can test whether or not such a model property really can be well-determined, given the available data. We illustrate this new possibility in the following two examples.

Example 2.4. Figure 2 shows a model of insulin (ins) binding and activation of insulin receptor (IR). This model was originally presented in [17], and is explained more in detail therein and in the supplementary materials. This model was one of several hypotheses tested in the article. To draw one of the conclusions in that article, the authors looked at the proportion of insulin receptor bound to an internal membrane:

$$prop_i := \frac{IR_i}{IR + IR_{ins} + IR_P + IR_{iP} + IR_i}.$$

It was predicted that $prop_i$ should be between .55 and .80. Despite the fact that there are four outputs, none of the seven states (insulin receptor or substrate quantities at different locations) are identifiable, so it is not immediately obvious whether $prop_i$ can be estimated from the measurements. Algorithm 3 with $\nu_i = n + \ell - 1$ for all i shows that this quantity is indeed single-experiment locally identifiable (see Definition 5.8 and Corollary 5.10). The authors were able to experimentally determine the proportion to be 0.023 ± 0.008. Thus they were able to reject the model, which was the conclusion in that step in the paper.

In summary, our analysis shows that the proportion of receptors that are bound to an internal membrane indeed is an identifiable model property. This new result holds from a structural point of view, without assuming that all derivatives are known. These results also show this without relying on optimization or exhaustive parameter sampling, as is done using the traditional practical identifiability analysis done in the original papers [17] and [2]. This new result thus strengthens and confirms the conclusion drawn in the original paper.

$$
\begin{aligned}
IR' &= -k_1 \cdot IR \cdot u - k_{1,basal} \cdot IR + k_R \cdot IR_i + IR_{ins} \cdot k_{m,1} + k_{m,2} \cdot IR_P, \\
IR'_{ins} &= k_1 \cdot IR \cdot u - k_2 \cdot IR_{ins} - k_{m,1} \cdot IR_{ins}, \\
IR'_P &= k_2 \cdot IR_{ins} + k_{1,basal} \cdot IR - k_3 \cdot IR_P + k_{m,3} \cdot IR_{iP} - k_{m,2} \cdot IR_P, \\
IR'_{iP} &= k_3 \cdot IR_P - k_{m,3} \cdot IR_{iP} - k_D \cdot IR_{iP}, \\
IR'_i &= k_D \cdot IR_{iP} - k_R \cdot IR_i, \\
IRS' &= -k_4 \cdot (IR_P + IR_{iP}) \cdot IRS + k_{m,4} \cdot IRS_P, \\
IRS'_P &= k_4 \cdot (IR_P + IR_{iP}) \cdot IRS - k_{m,4} \cdot IRS_P, \\
y_A &= k_{Y,A} \cdot (IR_P + IR_{iP}), \\
y_B &= k_{Y,B} \cdot IRS_P, \\
y_C &= k_{Y,C} \cdot IRS_P, \\
y_D &= k_{Y,D} \cdot IRS_P
\end{aligned}
$$

States:	IR, IR_{ins}, IR_P, IR_{iP}, IR_i, IRS, IRS_P
Parameters:	k_1, $k_{1,basal}$, $k_{m,1}$, k_2, $k_{m,2}$, k_3, k_D, k_R, k_4, $k_{m,4}$, $k_{Y,A}$, $k_{Y,B}$, $k_{Y,C}$, $k_{Y,D}$
Outputs:	y_A, y_B, y_C, y_D
Inputs:	u

Figure 2: *Model for insulin uptake. (For more details, see "Supplementary Materials".)*

$$
\begin{aligned}
x'_1 &= \frac{247x_2}{250} - \frac{247x_1}{250} + \frac{17\theta_1}{5000000} - \frac{17\theta_2 x_1}{1500} - \frac{340\theta_3 x_1 x_4}{9}\left(1 - \frac{\theta_6 x_5}{\theta_7 + x_5}\right), \\
x'_2 &= \frac{247x_1}{250} - \frac{247x_2}{250}, \\
x'_3 &= \frac{247x_4}{250} - \frac{247x_3}{250} + \frac{100000000 x_7 \theta_4 \theta_8 x_2^2}{\frac{3537}{50} + 100000000 x_2^2}, \\
x'_4 &= \frac{247x_3}{250} - \frac{247x_4}{250} - \frac{17}{1500}\theta_5 x_4, \\
x'_5 &= \frac{10000x_1}{3} - \frac{11}{2}, \\
x'_6 &= \frac{20x_2}{3} - \frac{x_6}{500}, \\
x'_7 &= \left(-\frac{81}{2500000}x_6^2 + \frac{63}{100000}x_6 - \frac{1}{400}\right)\theta_9 x_7, \\
y_1 &= \frac{5000}{3}(x_1 + x_2), \\
y_2 &= \frac{5000}{3}(x_4 + x_3)
\end{aligned}
$$

Figure 3: *Model for interaction between liver and pancreatic cells. (For more details, see "Supplementary Materials".)*

Example 2.5. Figure 3 shows the model of the interaction between liver and pancreatic cells from [18]. There are seven states, $x_1, \ldots, x_7$, two outputs, y_1, y_2, and no inputs. Numerical values were inserted for parameters whose values were known from operating specifications or estimated from the literature. The remaining parameters, here denoted by $\theta_1, \ldots, \theta_9$, were to be determined from experimental data. One quantity of interest was

$$\theta_3\left(1 - \frac{\theta_6 x_5}{\theta_7 + x_5}\right),$$

representing hepatic insulin sensitivity (see [18, Figure 4 Plot E]). Algorithm 3 tells us that this quantity is

indeed identifiable. Moreover, by using different values of (ν_1, ν_2) we find that it is $(6,5)$-identifiable.

2.5.2 Identifiability tends to increase with increased order of observability

It is typically the case that as we increase the number of derivatives that can be reliably estimated, the number of identifiable parameters increases, often dramatically. Therefore assuming that all derivatives are available can give a large overestimate of the number of identifiable parameters. How large this overestimate is can now be determined by our new algorithms. We illustrate this capability in the following two examples.

Example 2.6 (Drosophila period protein)**.** Figure 4 gives the model that was used in [19] to model Drosophila period protein. A local structural identifiability analysis done in [8, p. 737] shows that all

$$
\begin{aligned}
M' &= \frac{v_s K_I^4}{K_I^4 + P_N^4} - \frac{v_m M}{K_m + M}, \\
P_0' &= k_s M - \frac{V_1 P_0}{K_1 + P_0} + \frac{V_2 P_1}{K_2 + P_1}, \\
P_1' &= \frac{V_1 P_0}{K_1 + P_0} + \frac{V_4 P_2}{K_4 + P_2} - P_1 \left(\frac{V_2}{K_2 + P_1} + \frac{V_3}{K_3 + P_1} \right), \\
P_2' &= \frac{V_3 P_1}{K_3 + P_1} - P_2 \left(\frac{V_4}{K_4 + P_2} + k_1 + \frac{v_d}{K_d + P_2} \right) + k_2 P_N, \\
P_N' &= k_1 P_2 - k_2 P_N, \\
y_1 &= P_N,
\end{aligned}
$$

Figure 4: *Model for Drosophilia period protein. (For more details, see "Supplementary Materials".)*

states and parameters other than M, v_s, v_m, K_m, and k_s are identifiable. However, Algorithm 4 shows that only P_N is identifiable using 19 or fewer derivatives of y_1. If in addition we observe M, it was shown in [8] that all states and parameters are identifiable. However Algorithm 4 shows this is the case only if at least 16 derivatives of the outputs are available. Figure 5 shows the number of states and parameters that are identifiable as a function of the maximum number of derivatives of each output that are available; for example $\nu_{max} = 3$ means that exactly $y_i, y_i', y_i'', y_i^{(3)}$ are available for all i. Blue circles represent output set $\{y_1 = P_N\}$ and red squares represent output set $\{y_1 = P_N, y_2 = M\}$.

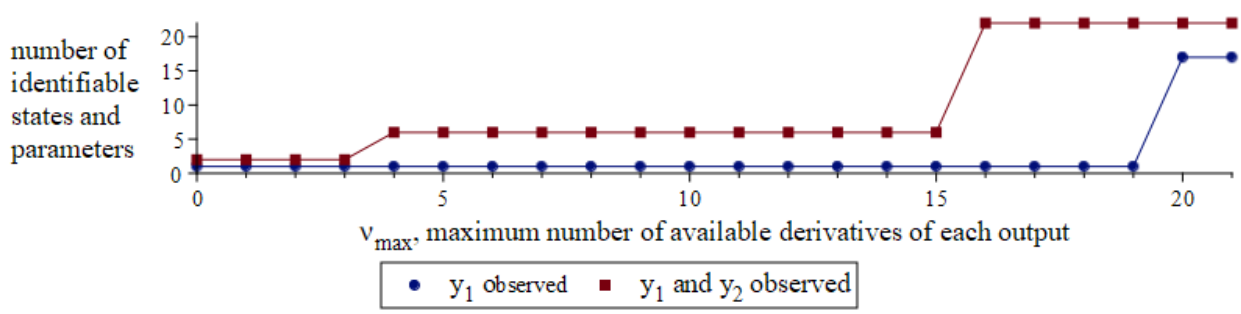

Figure 5: Variation of the number of identifiable parameters with maximum number of available output derivatives in a model of Drosophila period protein. The horizontal axis shows k: $\forall i\ (j \leq \nu_{max} \leftrightarrow y_i^{(j)}$ can be reliably estimated). The vertical axis shows the number of identifiable states and parameters in the model. Blue circles represent the case with only one output: $y_1 = P_N$. Red squares represent the case with two outputs: $y_1 = P_N$, $y_2 = M$.

Example 2.7 (NF-κB)**.** We consider a model of NF-κB regulatory module, as presented in Figure 6 below. This model was first introduced in [20], and an identifiability analysis was done in [21]. The model has 15 states, 28 parameters, 4 outputs, and 0 inputs.

$$
\begin{aligned}
x_1' &= -k_1x_1x_2 + \tfrac{1}{333}(-k_{14}x_1 + k_{15}x_9),\\
x_2' &= -k_1x_1x_2 + \tfrac{1}{333}k_{13}x_8,\\
x_3' &= k_1x_1x_2 - \tfrac{1}{333}k_{11}x_3,\\
x_4' &= k_3 + k_2x_2 - k_4x_4,\\
x_5' &= k_6 + k_5x_2 - k_7x_5,\\
x_6' &= k_9 + k_8x_2 - k_{10}x_6,\\
x_7' &= \tfrac{10}{16667}k_{11}x_3 - k_{21}x_7 + k_1x_8x_9 - k_{28}x_7x_{11},\\
x_8' &= k_{21}x_7 - \tfrac{10}{16667}k_{13}x_8 - k_1x_8x_9 + k_{26}x_{15},\\
x_9' &= k_{18}x_5 - k_{23}x_9 - k_1x_8x_9 + \tfrac{10}{16667}(k_{14}x_1 - k_{15}x_9) - k_{25}x_9x_{11},\\
x_{10}' &= k_{27}x_4 - k_{24}x_{10},\\
x_{11}' &= -k_{12}x_{11} - k_{16}x_{11} - k_{28}x_7x_{11} - k_{25}x_9x_{11} - k_{19}x_{10}x_{11} + k_{17}x_{13} + k_{22}x_{14} + k_{26}x_{15}
\end{aligned}
$$

$$
\begin{aligned}
x_{12}' &= k_{16}x_{11} + k_{19}x_{10}x_{11} - k_{12}x_{12},\\
x_{13}' &= k_{25}x_9x_{11} - k_{17}x_{13},\\
x_{14}' &= k_{20} - k_{12}x_{14} - k_{22}x_{14},\\
x_{15}' &= k_{28}x_7x_{11} - k_{26}x_{15},\\
y_1 &= x_4,\\
y_2 &= x_5,\\
y_3 &= x_6,\\
y_4 &= x_{10}
\end{aligned}
$$

Figure 6: *Model for NF-κB regulatory module. (For more details, see "Supplementary Materials".)*

The analysis in [21] shows that 21 states and parameters are identifiable, assuming that all derivatives of the four outputs can be estimated. However, using Algorithm 4, we found that if no more than 26 derivatives are available, then only 9 states and parameters are identifiable. Further analysis is shown in Figure 7.

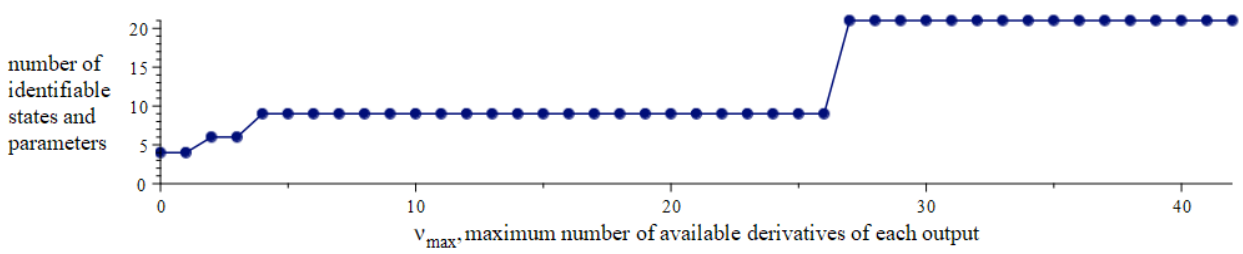

Figure 7: Variation of the number of identifiable parameters with maximum number of available output derivatives in a model of NF-κB regulatory module. The horizontal axis shows ν_{max}: $\forall i\ (j \leq \nu_{max} \leftrightarrow y_i^{(j)}$ can be reliably estimated$)$. The vertical axis shows the number of identifiable states and parameters in the model.

2.6 Relation to other available structural approaches to determine practical identifiability

We present a new method for determining structural identifiability in a more realistic setting, where only a given finite number of derivatives are available for each measurement signal. While this is novel, there are a few related methods that should be mentioned.

Most structural identifiability methods assume that all derivatives of outputs are measurable and all derivatives of all outputs are sufficiently varying. The latter condition is not always satisfied in practical modeling as, for example, a constant function may be used as an input, and this may lead to incorrect implementation of a model. In [22] a method for determining how many derivatives of the inputs must be non-zero was given, and an implementation was presented in [23].

In [16] an implementation was given for determining minimal sets of outputs that make a system structurally locally identifiable.

3 Theory

The definition of $\boldsymbol{\nu}$-identifiability and the justification of our results are made rigorous in Section 5. We give a brief summary of this theory here.

The notion of identifiability involves many subtleties. Our precise definition is a type of generic local single-experiment identifiability. This is done in terms of uniqueness (up to a finite set) of parameter and state values given known values of the output derivatives $\mathbf{y}^{(\leq \boldsymbol{\nu})}$ at a certain point in time. We show that if all derivatives of all outputs are available, then our definition is equivalent to the notion of local identifiability given in [10, Def. 2.5].

Our definition of $\boldsymbol{\nu}$-identifiability is intended to capture the property that a modeler really wants to ascertain, but it is often difficult to verify directly. Therefore we show that this property is equivalent to a property involving ranks of matrices, which is straightforward to check using existing software. This is in fact the property alluded to in Algorithm 1.

Although this rank property is straightforward to check, it involves computing ranks of symbolic matrices, which can cost the modeler time. Therefore, to alleviate this problem, we present probabilistic algorithms that substitute random integers for the variables and compute modulo a random prime. These are modifications of the method described in [8]. We present detailed probabilistic statements that a reader can use to create an implementation without investigating the proofs. We present detailed proofs that the reader can check without excessive pencil and paper computations.

Moreover, our work addresses an oversight in [8] surrounding the probability that a rational expression vanishes when evaluated at a random tuple of integers modulo a random prime. The author assumes that the denominator does not vanish, but then gives a bound for the probability that the numerator does not vanish regardless of this condition. We give a bound on the probability that the numerator does not vanish given that the denominator does not vanish.

4 Conclusions

Practical identifiability analysis lies at the heart of model-based data analysis. State-of-the-art numerical methods, such as profile likelihood and MCMC, suffer from the limitation that they require an exhaustive coverage of the parameter space, which is not possible for large and realistic models. This coverage is not needed for structural methods, but these methods traditionally assume that one can determine all derivatives of the measurements y. Such derivatives are not available in practice, but the consequences of this unfulfilled assumption have not been examined, because corresponding methods and algorithms are missing. Herein, we present a new definition – ν-identifability – which restricts the analysis to those derivatives of y that are available. We also present a new algorithm, which can determine such identifiability in practice. Applications to previously published models demonstrate that we can determine not only identifiability of parameters, but observability of any model property (Examples 2.4 and 2.5). These results allow us to strengthen previous conclusions. For instance, the previous rejection of a specific model for insulin signaling (Figure 2), was based on the model-guided experiments measuring the amount of internalized and dephosphorylated insulin receptors (Example 2.4), and we now show that this model property indeed can be identifiable, using a methodology that does not require coverage of the entire parameter space. Our analysis also shows that the number of parameters identifiable according to traditional structural methods are widely overestimated. If one e.g. assumes that only up to third order derivatives, instead of all derivatives, are available, the number of identifiable parameters drops from 17 to 1 for the Drosophila model (Figure 5), and

from 21 to 4 for an NF-κB model (Figure 7). In both these models, the previously obtained identifiability is present only if at least 20 derivatives of all measurement signals are available. Our results bring us one step closer to a structural approach for practical identifiability analysis.

5 Technical Details

5.1 Notation and Definitions

The notion of identifiability involves many subtleties. This subsection provides the setup necessary for rigorous treatment.

For positive integers n, ℓ, r, and m, tuples $\mathbf{f} \in \mathbb{C}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u})^n$ and $\mathbf{g} \in \mathbb{C}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u})^m$, we have a model class

$$\Sigma(t_0, \mathbf{x}^*, \boldsymbol{\theta}, \mathbf{u}) = \begin{cases} \dfrac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u}), \\ \mathbf{y} = \mathbf{g}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u}), \\ \mathbf{x}(t_0) = \mathbf{x}^*. \end{cases} \tag{2}$$

Fix positive integers n, r, and m, fix non-negative integer ℓ, and fix n-tuple $\mathbf{f}$ and m-tuple $\mathbf{g}$ of elements in $\mathbb{C}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u})$ for the rest of this section. The symbol $\mathbb{N}$ denotes the set of non-negative integers.

Notation 5.1. (Differential Algebraic Setting)

(a) A *differential ring* $(S, \,'\,)$ is a ring together with a *derivation*, that is, a function $'\colon S \to S$ satisfying $\forall a, b \in S\ (a+b)' = a' + b'$ and $(ab)' = a'b + ab'$. All rings are assumed to be commutative and with multiplicative identity. For $a \in S$ and $k \in \mathbb{N}$, $a^{(k)}$ denotes the image of a after k applications of $'$. The set $\{a^{(k)} \mid k \geq 0\}$ will be called the set of *derivatives* of a.

(b) Let $(S, \,'\,)$ be a differential ring and let $Z = \{Z_1, \ldots, Z_k\}$ be a set of indeterminates. Then $S\{Z\} = S\{Z_1, \ldots, Z_k\}$ denotes the polynomial ring in infinitely many indeterminates $S[Z_1, Z_1', Z_1'', \ldots, Z_k, Z_k', Z_k'', \ldots]$ with derivation $'$ extended by $(Z_i^{(j)})' = Z_i^{(j+1)}$.

(c) Let $(S_1, \,'\,)$ be a differential subring of $(S_2, \,'\,)$ and let $Z = \{Z_1, \ldots, Z_k\} \subseteq S_2$. Then $S_1\{Z\} = S_1\{Z_1, \ldots, Z_k\}$ denotes the smallest subring of S_2 containing S_1 and the derivatives of the elements of Z.

(d) Let $(S_1, \,'\,)$ be a differential ring that is a subring of a differential integral domain $(S_2, \,'\,)$ and let $Z = \{Z_1, \ldots, Z_k\} \subseteq S_2$. Then $S_1\langle Z \rangle = S_1\langle Z_1, \ldots, Z_k \rangle$ denotes the fraction field of $S_1\{Z\}$. This is a differential ring under the extension of $'$ via the quotient rule $(a/b)' = (ba' - ab')/b^2$.

(e) For any differential ring containing any elements of $\mathbb{C}(\boldsymbol{\theta})$ we assume $'$ maps each such element to 0.

(f) Let $(S, \,'\,)$ be a differential ring. A *differential ideal* I is an ideal satisfying $\forall a \in I\ a' \in I$. For subset $A \subseteq S$, we denote by $[A]$ the smallest differential ideal containing A.

(g) For an ideal I and element a in a ring S, we denote $I\colon a^\infty = \{s \in S \mid \exists k\colon a^k s \in I\}$. This set is also an ideal in S.

(h) Write $\mathbf{f}$ and $\mathbf{g}$ as fractions of elements of $\mathbb{C}[\mathbf{x}, \boldsymbol{\theta}, \mathbf{u}]$, let $Q \in \mathbb{C}[\mathbf{x}, \boldsymbol{\theta}, \mathbf{u}]$ be the LCM of the denominators, and let $\mathbf{F}$ and $\mathbf{G}$ be such that $\mathbf{f} = \mathbf{F}/Q$ and $\mathbf{g} = \mathbf{G}/Q$. Note that Q is unique up to multiplication by an element of $\mathbb{C}$ and choice of Q will not affect our definitions and results. We define the differential ideal of Σ as $\mathcal{J} = [Q\mathbf{x}' - \mathbf{F}, Q\mathbf{y} - \mathbf{G}] : Q^\infty \subseteq \mathbb{C}[\boldsymbol{\theta}]\{\mathbf{x}, \mathbf{y}, \mathbf{u}\}$. This is a prime (differential) ideal (cf [10, Lemma 3.2]) and thus $R := \mathbb{C}[\boldsymbol{\theta}]\{\mathbf{u}, \mathbf{x}, \mathbf{y}\}/\mathcal{J}$ is a differential ring that is an integral domain.

(i) For $\boldsymbol{\nu} \in \mathbb{N}^m$, we denote by $\mathbf{y}^{(\leq \boldsymbol{\nu})}$ the subset $\{y_i^{(j_i)} \mid 1 \leq i \leq m, 0 \leq j_i \leq \nu_i\}$ of R. We define $\nu_{max} := \max(\nu_1, \ldots, \nu_m)$.

Notation 5.2. (Analytic Setting)

(a) Let $t_0 \in \mathbb{C}$. Then $\mathbb{C}^\infty(t_0)$ denotes the set of all functions that are complex analytic in some neighborhood of $t = t_0$.

(b) Let $t_0 \in \mathbb{C}$. Then $\Omega^{t_0} \subseteq \mathbb{C}^n \times \mathbb{C}^\ell \times (\mathbb{C}^\infty(t_0))^r$ denotes the complement to the set where at least one of the denominators of $\mathbf{f}$ or $\mathbf{g}$ vanishes at $t = t_0$.

(c) For $t_0 \in \mathbb{C}$ and $h \in \mathbb{C}(\mathbf{x}, \boldsymbol{\theta})$, let $\Omega_h^{t_0} = \{(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^{t_0} \mid h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}})$ is defined$\}$.

(d) For $t_0 \in \mathbb{C}$ and $(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^{t_0}$, let $X(t_0, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in (\mathbb{C}^\infty(t_0))^n$ and $Y(t_0, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in (\mathbb{C}^\infty(t_0))^m$ denote the unique solution of the initial value problem $\Sigma(t_0, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})$ (cf. [24, Theorem 2.2.2]). For appropriate integers j and k, let $Y_j^{(k)}(t_0, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})$ denote the k-th derivative of the j-th component of $Y(t_0, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})$.

(e) For $s \in \mathbb{N}$, a subset $U \subseteq \mathbb{C}^s$ is called *Zariski open* if there exists P in $\mathbb{C}[Z_1, \ldots, Z_s]$ such that U is the complement to the zero set of P.

(f) For $s \in \mathbb{N}$ and $t_0 \in \mathbb{C}$, a subset $U \subseteq (\mathbb{C}^\infty(t_0))^s$ is called *Zariski open* if there exists P in $\mathbb{C}\{Z_1, \ldots, Z_s\}$ such that

$$U = \{\hat{\mathbf{z}} \in (\mathbb{C}^\infty(t_0))^s \mid P(\hat{\mathbf{z}})|_{t=t_0} \neq 0\}.$$

(g) For $s \in \mathbb{N}$, $t_0 \in \mathbb{C}$, and $W = \mathbb{C}^s$ or $(\mathbb{C}^\infty(t_0))^s$, the set of all nonempty Zariski open subsets of W will be denoted by $\tau(W)$.

(h) A subset $V \subseteq \mathbb{C}$ is called *codiscrete* if for any distinct $c_1, c_2 \in \mathbb{C} \backslash V$ there exist disjoint open sets $W_1 \ni c_1$ and $W_2 \ni c_2$ such that $W_1 \cap (\mathbb{C} \backslash V) = \{c_1\}$ and $W_2 \cap (\mathbb{C} \backslash V) = \{c_2\}$.

Notation 5.3. Let $h \in \mathbb{C}(\mathbf{x}, \boldsymbol{\theta})\langle \mathbf{u} \rangle$, let $t_1 \in \mathbb{C}$, and let $\hat{p} \in \Omega_h^{t_1}$.

- Substituting $(X(t_1, \hat{p}), \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})$ for $(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u})$ in h gives a function of one variable. The image of h under this substitution will be denoted by $h(t_1, \hat{p})(t)$.
- The connected component containing t_1 of the intersection of the domains of $h(t_1, \hat{p})(t)$, $\hat{\mathbf{u}}(t)$, $X(t_1, \hat{p})(t)$ and $Y(t_1, \hat{p})(t)$ will be denoted by $D_h(t_1, \hat{p})$.

Definition 5.4. Let $\boldsymbol{\nu} \in \mathbb{N}^m$. Let $h \in \mathbb{C}(\mathbf{x}, \boldsymbol{\theta})$. The expression h is said to be $\boldsymbol{\nu}$-identifiable if

$$\exists \Psi \in \tau(\mathbb{C}^{n+\ell}) \quad \exists U \in \tau((\mathbb{C}^\infty(0))^r) \quad \forall \hat{p} := (\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega_h^0 \cap (\Psi \times U)$$
$$\exists \text{ codiscrete } V \subseteq \mathbb{C} \quad \forall t_0 \in V \cap D_h(0, \hat{p}) \quad |S_h(t_0, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})| \text{ is finite,}$$

where

$$S_h(t_0, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) = \{h(\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}) \mid \tilde{p} := (\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega_h^{t_0} \text{ and } Y^{(\leq \boldsymbol{\nu})}(t_0, \tilde{p})(t_0) = Y^{(\leq \boldsymbol{\nu})}(0, \hat{p})(t_0)\}.$$

One can interpret this definition as follows: If at some time t_0, not belonging to a certain discrete "singular" set, we know the exact values of $Y^{(\leq \boldsymbol{\nu})}(t_0)$, then if we consider all possible values of the state variables and parameters at t_0 that produce these values of $Y_i^{(j)}(t_0)$, and then we consider the set of values of h obtained by evaluating h at these values, this set is finite.

5.1.1 Examples to illustrate the definition

Example 5.5. Consider the system

$$\Sigma: \quad x_1' = \theta_1\theta_2 x_1, \quad y_1 = x_1, \quad y_2 = \theta_2, \quad x_1(t_0) = x_1^*.$$

We show that θ_1 is $(1,0)$-identifiable. Note that $n = 1$ and $m = 2$. Since two parameters occur we can take ℓ to be any integer at least 2. Although no inputs occur the definitions require that we take r to be at least 1. Set $\ell = 2$ and $r = 1$. Let $\Psi = \{(\hat{x}_1^*, \hat{\theta}_1, \hat{\theta}_2) \in \mathbb{C}^3 \mid \hat{\theta}_2\hat{x}_1^* \neq 0\}$ and let $U = \mathbb{C}^\infty(0)$. Note that since there are no denominators we have $\Omega_{\theta_1}^{t_0} = \mathbb{C}^3 \times \mathbb{C}^\infty(t_0)$ for all t_0. Fix $\hat{p} := (\hat{x}_1^*, \hat{\theta}_1, \hat{\theta}_2, \hat{u}) \in \Omega_{\theta_1}^0 \cap (\Psi \times U)$. Let $V = \mathbb{C}$ and fix $t_0 \in V \cap D_{\theta_1}(0, \hat{p})$. It is trivial to obtain a formula for the solution of this system, and we have

$$Y_1(0,\hat{p})(t) = \hat{x}_1^* e^{\hat{\theta}_1\hat{\theta}_2 t}, \quad Y_2(0,\hat{p})(t) = \hat{\theta}_2.$$

Let $\tilde{p} := (\tilde{x}_1^*, \tilde{\theta}_1, \tilde{\theta}_2, \hat{u}) \in \Omega_{\theta_1}^{t_0}$, and note that

$$Y_1(t_0,\tilde{p})(t) = \tilde{x}_1^* e^{\tilde{\theta}_1\tilde{\theta}_2(t-t_0)}, \quad Y_2(t_0,\tilde{p})(t) = \tilde{\theta}_2. \tag{3}$$

For $\mathbf{v} = (1,0)$, the condition $Y^{(\leq \mathbf{v})}(t_0,\tilde{p})(t_0) = Y^{(\leq \mathbf{v})}(0,\hat{p})(t_0)$ gives

$$\hat{x}_1^* e^{\hat{\theta}_1\hat{\theta}_2 t_0} = \tilde{x}_1^*, \quad \hat{x}_1^*\hat{\theta}_1\hat{\theta}_2 e^{\hat{\theta}_1\hat{\theta}_2 t_0} = \tilde{x}_1^*\tilde{\theta}_1\tilde{\theta}_2, \quad \hat{\theta}_2 = \tilde{\theta}_2. \tag{4}$$

By the definition of Ψ, it follows that no side of the first or third of these equations is 0. Dividing the second equation by the product of the first and third proves that $\hat{\theta}_1 = \tilde{\theta}_1$. Since $\theta_1(\tilde{p}) = \tilde{\theta}_1$, we see that $S_{\theta_1}(t_0,\hat{p}) = \{\hat{\theta}_1\}$ and thus has cardinality 1. We conclude that θ_1 is $(1,0)$-identifiable.

We now show that θ_1 is not $(0,0)$-identifiable. Fix $\Psi \in \tau(\mathbb{C}^3)$ and $U \in \tau(\mathbb{C}^\infty(0))$. Let $\hat{p} := (\hat{x}_1^*, \hat{\theta}_1, \hat{\theta}_2, \hat{u}) \in \Omega_{\theta_1}^0 \cap (\Psi \times U)$. Fix codiscrete V, and let $t_0 \in V \cap D_{\theta_1}(0,\hat{p})$. Using the first and third equations of (4), we see that for all $k \in \mathbb{C}$ the tuple $\tilde{p}_k := (\hat{x}_1^* e^{\hat{\theta}_1\hat{\theta}_2 t_0}, k, \hat{\theta}_2, \hat{u})$ satisfies the condition $Y^{(\leq \mathbf{v})}(t_0,\tilde{p}_k)(t_0) = Y^{(\leq \mathbf{v})}(0,\hat{p})(t_0)$ for $\mathbf{v} = (0,0)$. Therefore

$$S_{\theta_1}(t_0,\hat{p}) \supseteq \{k \mid k \in \mathbb{C}\},$$

and the right-hand side is infinite. We conclude that θ_1 is not $(0,0)$-identifiable.

The next example includes an input and demonstrates why sometimes a discrete set must be excluded from V.

Example 5.6. Consider the system

$$\Sigma: \quad x_1' = \frac{1}{x_1}, \quad x_2' = f_2(x_1,x_2,u_1) \quad y_1 = x_2(x_1 - u_1), \quad y_2 = x_1 - u_1, \quad x_1(t_0) = x_1^*, \quad x_2(t_0) = x_2^*,$$

where f_2 is some rational function.

We have $n = 2$ and $m = 2$. Take $\ell = 0$ and $r = 1$. We show that x_2 is $(0,0)$-identifiable. We have $\Omega_{x_2}^0 = \{(\hat{x}_1^*, \hat{x}_2^*, \hat{u}_1) \mid \hat{x}_1^* \cdot (\operatorname{denom} f_2)(\hat{x}_1^*, \hat{x}_2^*, \hat{u}_1) \neq 0\}$. Let $\Psi = \mathbb{C}^2$ and let $U = \{z \in \mathbb{C}^\infty(0) \mid z(0)z'(0) - 1 \neq 0\}$. Let $\hat{p} := (\hat{x}_1^*, \hat{x}_2^*, \hat{u}_1) \in \Omega_{x_2}^0 \cap (\Psi \times U)$. Note that because of the way U was defined, $X_1(0,\hat{p}) - \hat{u}_1$ is not the zero function. Let V be the complement of the vanishing of $X_1(0,\hat{p}) - \hat{u}_1$, which is necessarily codiscrete since $X_1(0,\hat{p})$ and $\hat{u}_1$ are analytic. Let $t_0 \in V \cap D_{x_2}(0,\hat{p})$. If $\tilde{p} = (\tilde{x}_1^*, \tilde{x}_2^*, \hat{u}_1)$ is such that $Y_i(t_0,\tilde{p})(t_0) = Y_i(0,\hat{p})(t_0)$ for $i = 1,2$ then it follows that

$$x_2(\tilde{p}) = \tilde{x}_2^* = X_2(t_0,\tilde{p})(t_0) = \frac{Y_1(t_0,\tilde{p})(t_0)}{Y_2(t_0,\tilde{p})(t_0)} = \frac{Y_1(0,\hat{p})(t_0)}{Y_2(0,\hat{p})(t_0)} = X_2(0,\hat{p})(t_0).$$

Since $t_0 \in D_{x_2}(0,\hat{p})$ we know that $Y_1(0,\hat{p})(t_0)$ and $Y_2(0,\hat{p})(t_0)$ are defined and since $t_0 \in V$ we know that $Y_2(0,\hat{p})(t_0) \neq 0$. Thus $|S_{x_2}(t_0,\hat{p})| = 1$ and we conclude that x_2 is $(0,0)$-identifiable.

5.1.2 ν-identifiable implies single-experiment locally identifiable and the converse is true for sufficiently large ν

Many notions of identifiability are used in the literature. Some of these are stated and compared in [25]. In general, Definition 5.4 is not equivalent to any published definition as far as we are aware.

Definition 5.4 can be viewed as a type of single-experiment generic local identifiability. It is generic because there is a (possibly empty) set of numerical values of parameters and input functions for which h cannot be recovered but if the true values of the parameters and inputs lie outside this set h can be recovered. It is local because we allow h to lie in a finite set. One could change "is finite" to "equals one" to create a definition of a globally $\boldsymbol{\nu}$-identifiable function, however we do not address this in the current work. It is "single-experiment" because it refers to only one instance of the model. This is in contrast to a multi-experiment approach where one observes multiple instances of the model, usually with the same equation parameters but different initial conditions.

Since the equations of Σ are rational in $\mathbf{x}$, $\boldsymbol{\theta}$, and $\mathbf{u}$, there exist relations among sufficiently high derivatives of $\mathbf{y}$ and hence it is not meaningful to consider ν_i beyond a certain order. This is made precise by the following proposition.

Proposition 5.7. *Let $h \in \mathbb{C}(\mathbf{x},\boldsymbol{\theta})$ and $\boldsymbol{\nu} \in \mathbb{N}^m$. For each $i = 1,\dots,m$ let μ_i be the greatest non-negative integer such that the set $y_i, y_i', \dots, y_i^{(\mu_i)}$ is not algebraic over $\mathbb{C}\langle \mathbf{u}\rangle$. Then h is $\boldsymbol{\nu}$-identifiable if and only if h is $(\min(\nu_1,\mu_1),\dots,\min(\nu_m,\mu_m))$-identifiable. Moreover, $\forall i\ \mu_i \le n+\ell-1$.*

Proof. Since the transcendence degree of $\mathbb{C}\langle \mathbf{u}\rangle(\mathbf{x},\boldsymbol{\theta})$ over $\mathbb{C}\langle \mathbf{u}\rangle$ is $n+\ell$, it must be that $y_i, y_i', \dots, y_i^{(n+\ell)}$ is algebraic over $\mathbb{C}\langle \mathbf{u}\rangle$. Hence such a μ_i exists and moreover $\mu_i \le n+\ell-1$.

For the $\Leftarrow$ direction, note that if $\mathbf{a},\mathbf{b} \in \mathbb{N}^m$ are such that $a_i \ge b_i$ for all i, it follows from the definition that h is $\mathbf{b}$-identifiable implies h is $\mathbf{a}$-identifiable.

We now address the $\Rightarrow$ direction. Suppose h is $\boldsymbol{\nu}$-identifiable. By Proposition 5.12, h is algebraic over $\mathbb{C}\langle \mathbf{u}\rangle(\mathbf{y}^{(\le \boldsymbol{\nu})})$. For any i, by writing an algebraic dependence of $y_i, y_i', \dots, y_i^{(\mu_i+1)}$ over $\mathbb{C}\langle \mathbf{u}\rangle$ and differentiating (noting that the field has characteristic 0), we see that for all $s > \mu_i$ the element $y_i^{(s)}$ is algebraic over $\mathbb{C}\langle \mathbf{u}\rangle(y_i,\dots,y_i^{(\mu_i)})$. It follows that h is algebraic over $\mathbb{C}\langle \mathbf{u}\rangle(\{y_i,\dots,y_i^{(\min(\nu_i,\mu_i))}\}_{i=1,\dots,m})$. By Proposition 5.12 h is $(\min(\nu_1,\mu_1),\dots,\min(\nu_m,\mu_m))$-identifiable. $\square$

Definition 5.8. Let $h \in \mathbb{C}(\mathbf{x},\boldsymbol{\theta})$. The expression h is said to be single-experiment locally identifiable (SELI) if

$$\exists \Psi \in \tau(\mathbb{C}^{n+\ell}) \quad \exists U \in \tau((\mathbb{C}^\infty(0))^r) \quad \forall \hat{p} := (\hat{\mathbf{x}}^*,\hat{\boldsymbol{\theta}},\hat{\mathbf{u}}) \in \Omega_h^0 \cap (\Psi \times U) \quad |S_h(\hat{\mathbf{x}}^*,\hat{\boldsymbol{\theta}},\hat{\mathbf{u}})| \text{ is finite,}$$

where

$$S_h(\hat{\mathbf{x}}^*,\hat{\boldsymbol{\theta}},\hat{\mathbf{u}}) = \{h(\tilde{\mathbf{x}}^*,\tilde{\boldsymbol{\theta}}) \mid \tilde{p} := (\tilde{\mathbf{x}}^*,\tilde{\boldsymbol{\theta}},\hat{\mathbf{u}}) \in \Omega_h^0 \text{ and } Y(0,\tilde{p}) = Y(0,\hat{p})\}.$$

For $h \in \mathbf{x} \cup \boldsymbol{\theta}$, Definition 5.8 is equivalent to the definition of local identifiability given in Definition 2.5 of [10] (generalization to multiple inputs is asserted in Remark 2.2). In [10, Prop. 3.4 (a) $\Leftrightarrow$ (c)], it was shown that for $h \in \mathbf{x} \cup \boldsymbol{\theta}$, h is SELI if and only if h is algebraic over $\mathbb{C}\langle \mathbf{y},\mathbf{u}\rangle$. We extend this result to arbitrary $h \in \mathbb{C}(\mathbf{x},\boldsymbol{\theta})$.

Proposition 5.9. *Let $h \in \mathbb{C}(\mathbf{x},\boldsymbol{\theta})$. Then h is SELI if and only if h is algebraic over $\mathbb{C}\langle \mathbf{y},\mathbf{u}\rangle$.*

Proof. Let $\Sigma_1(0,\mathbf{x}^*,x_{n+1}^*,\boldsymbol{\theta},\mathbf{u})$ be the system obtained by adding to $\Sigma(0,\mathbf{x}^*,\boldsymbol{\theta},\mathbf{u})$ the equations $x_{n+1}' = \sum_{i=1}^n f_i \frac{\partial h}{\partial x_i}$, $y_{m+1} = x_{n+1} - h$, and $x_{n+1}(0) = x_{n+1}^*$. Note that Ω_h^0 in Σ is the projection of $\Omega_{x_{n+1}}^0$ onto all

coordinates but x_{n+1}. We divide the proof into the following three steps: h is Σ-SELI $\overset{1}{\Leftrightarrow}$ x_{n+1} is Σ_1-SELI $\overset{2}{\Leftrightarrow}$ x_{n+1} is algebraic over $\mathbb{C}\langle \mathbf{y}, y_{m+1}, \mathbf{u}\rangle$ $\overset{3}{\Leftrightarrow}$ h is algebraic over $\mathbb{C}\langle \mathbf{y}, \mathbf{u}\rangle$.

We show h is Σ-SELI $\Longleftrightarrow$ x_{n+1} is Σ_1-SELI. Suppose h is Σ-SELI. Let Ψ and U be as required by the definition. Define $\Psi_1 = \Psi \times \mathbb{C}$. Let $(\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_{x_{n+1}} \cap (\Psi_1 \times U)$. We verify that $|S_{x_{n+1}}(\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})| = |S_h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})|$. Since $Y'_{m+1} = 0$, we have

$$\begin{aligned} S_{x_{n+1}}(\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) = \{\tilde{x}^*_{n+1} \mid (\tilde{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_{x_{n+1}} \text{ and } Y(0, (\tilde{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}})) = Y(0, (\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})) \\ \text{and } \tilde{x}^*_{n+1} - h(\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}) = \hat{x}^*_{n+1} - h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}})\}, \end{aligned}$$

where $Y = (Y_1, \ldots, Y_m)$. It follows that

$$\begin{aligned} S_{x_{n+1}}(\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) &= \{h(\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}) + \hat{x}^*_{n+1} - h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}) \mid (\tilde{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_{x_{n+1}} \text{ and} \\ &\qquad Y(0, (\tilde{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}})) = Y(0, (\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}))\} \\ &= \{h(\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}) + \hat{x}^*_{n+1} - h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}) \mid (\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_h \text{ and} \\ &\qquad Y(0, (\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}})) = Y(0, (\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}))\}, \end{aligned}$$

where Ω^0_h is considered a subset of $\mathbb{C}^{n+\ell} \times \mathbb{C}^\infty(0)$. Thus

$$\begin{aligned} |S_{x_{n+1}}(\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})| &= |\{h(\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}) \mid (\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_h \text{ and } Y(0, (\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}})) = Y(0, (\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}))\}| \\ &= |S_h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})|, \end{aligned}$$

and we know that $|S_h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})| < \infty$ because h is Σ-SELI. This completes the first direction. Now suppose x_{n+1} is Σ_1-SELI. Let Ψ_1 and U be as required by the definition. Define Ψ to be the projection of Ψ_1 onto all coordinates other than x_{n+1}. Let $(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_h \cap (\Psi \times U)$. Let $\hat{x}^*_{n+1}$ be such that $(\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_{x_{n+1}} \cap (\Psi_1 \times U)$. Such a $\hat{x}^*_{n+1}$ exists because of the way Ψ was defined. Now

$$\begin{aligned} S_h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) &= \{h(\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}) \mid \ (\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_h \text{ and } Y(0, (\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}})) = Y(0, (\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}))\} \\ &= \{h(\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}) \mid \ (\tilde{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_{x_{n+1}} \text{ and } Y(0, (\tilde{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}})) = Y(0, (\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})) \\ &\qquad \text{and } \tilde{x}^*_{n+1} - h(\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}) = \hat{x}^*_{n+1} - h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}})\} \\ &= \{\tilde{x}^*_{n+1} - \hat{x}^*_{n+1} + h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}) \mid (\tilde{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_{x_{n+1}} \text{ and} \\ &\qquad Y_i(0, (\tilde{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}})) = Y_i(0, (\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})),\ i = 1, \ldots, m+1\}. \end{aligned}$$

Hence

$$\begin{aligned} |S_h(\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})| &= |\{\tilde{x}^*_{n+1} \mid \ (\tilde{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega^0_{x_{n+1}} \text{ and} \\ &\qquad Y_i(0, (\tilde{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}})) = Y_i(0, (\hat{\mathbf{x}}^*, \tilde{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})),\ i = 1, \ldots, m+1\}| \\ &= |S_{x_{n+1}}(\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})|, \end{aligned}$$

and we know that $|S_{x_{n+1}}(\hat{\mathbf{x}}^*, \hat{x}^*_{n+1}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})| < \infty$ because x_{n+1} is Σ_1-SELI.

From [10, Prop. 3.4 (a) $\Leftrightarrow$ (c)], we have that x_{n+1} is Σ_1-SELI $\Longleftrightarrow$ x_{n+1} is algebraic over $\mathbb{C}\langle \mathbf{y}, y_{m+1}, \mathbf{u}\rangle$.

We now show that h is algebraic over $\mathbb{C}\langle \mathbf{y}, \mathbf{u}\rangle$ $\Longleftrightarrow$ x_{n+1} is algebraic over $\mathbb{C}\langle \mathbf{y}, y_{m+1}, \mathbf{u}\rangle$. Suppose h is algebraic over $\mathbb{C}\langle \mathbf{y}, \mathbf{u}\rangle$. Then $x_{n+1} = h + y_{m+1}$ is algebraic over $\mathbb{C}\langle \mathbf{y}, y_{m+1}, \mathbf{u}\rangle$. We now address the other direction. Let $\mathcal{F} = \mathbb{C}\langle \mathbf{y}, \mathbf{u}\rangle$ and let $\mathcal{F}_1 = \mathbb{C}\langle \mathbf{y}, y_{m+1}, \mathbf{u}\rangle$. Suppose x_{n+1} is algebraic over $\mathcal{F}_1$. Then $h = x_{n+1} - y_{m+1}$ is algebraic over $\mathcal{F}_1$. Let $\sum_{i=0}^k a_i Z^i$ be the minimal polynomial of h over $\mathcal{F}_1$. Suppose y_{m+1}

appears in some a_i. Let α be the differential field automorphism on $\mathbb{C}(\mathbf{x}, x_{n+1}, \boldsymbol{\theta})\langle \mathbf{u} \rangle$ such that $\mathbb{C}(\mathbf{x}, \boldsymbol{\theta})\langle \mathbf{u} \rangle$ is fixed pointwise, $\alpha(y_{m+1}) = y_{m+1} + 1$, and $\alpha(x_{n+1}) = x_{n+1} + 1$. Now $\sum_{i=0}^{k}(a_i - \alpha(a_i))Z^i$ is a non-zero polynomial of lower degree with coefficients in $\mathcal{F}_1$ that has h as a root. Thus we have a contradiction. Therefore all the a_i lie in $\mathcal{F}$ and h is algebraic over $\mathcal{F}$. □

This leads to the main result of this section.

Corollary 5.10. *Let $h \in \mathbb{C}(\mathbf{x}, \boldsymbol{\theta})$ and $\boldsymbol{\nu} \in \mathbb{N}^m$.*

1. *h is $\boldsymbol{\nu}$-identifiable $\Rightarrow$ h is SELI.*
2. *If $\forall i\ \nu_i \geq n + \ell - 1$, then h is SELI $\Rightarrow$ h is $\boldsymbol{\nu}$-identifiable.*

Proof. Suppose h is $\boldsymbol{\nu}$-identifiable. Then by Proposition 5.12 we know h is algebraic over $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \boldsymbol{\nu})})$. It follows that h is algebraic over $\mathbb{C}\langle \mathbf{y}, \mathbf{u} \rangle$ and then by Proposition 5.9 we have that h is SELI.

Suppose that $\forall i\ \nu_i \geq n + \ell - 1$. Then by the arguments presented in Proposition 5.7 we have $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \boldsymbol{\nu})}) = \mathbb{C}\langle \mathbf{u}, \mathbf{y} \rangle$. Suppose h is SELI. By Proposition 5.9 we have that h is algebraic over $\mathbb{C}\langle \mathbf{u}, \mathbf{y} \rangle$, which equals $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \boldsymbol{\nu})})$. By Proposition 5.12 we have that h is $\boldsymbol{\nu}$-identifiable. □

5.2 Proof of Algorithms

The definition of $\boldsymbol{\nu}$-identifiability is stated in terms of analytic functions. The Proposition 5.12 gives a correspondence between the analytic property and an algebraic property. Its proof will use the following lemma.

Lemma 5.11. *Let $t_1, t_2 \in \mathbb{C}$, and let $\hat{p} \in \Omega^{t_1}$. The map*

$$h \mapsto h(t_1, \hat{p})(t_2)$$

gives a $\mathbb{C}$-algebra homomorphism $S \to \mathbb{C}$, where S is the ring $\{h \in Frac(R) \mid t_2 \in D_h(t_1, \hat{p})\}$. Moreover, under this map $\mathbf{y} \mapsto Y(t_1, \hat{p})(t_2)$.

Proof. Denote the stated map by φ. Since $X(t_1, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}), \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}$ satisfy all the relations among $\mathbf{x}, \boldsymbol{\theta}, \mathbf{u}$ and no denominator of S is sent to 0, we see that φ is a homomorphism on S. Now $\mathbf{y} = \mathbf{g}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u})$, so $\varphi(\mathbf{y}) = \mathbf{g}(\varphi(\mathbf{x}), \varphi(\boldsymbol{\theta}), \varphi(\mathbf{u})) = \mathbf{g}(X(t_1, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}), \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})(t_2)$. The existence and uniqueness theorem guarantees the existence and uniqueness of $X(t_1, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})$, and $Y(t_1, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})$ is given by $\mathbf{g}(X(t_1, \hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}), \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})$. □

Proposition 5.12. *Let $\boldsymbol{\nu} \in \mathbb{N}^m$. Expression $h \in \mathbb{C}(\mathbf{x}, \boldsymbol{\theta})$ is $\boldsymbol{\nu}$-identifiable if and only if h is algebraic over the subfield $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \boldsymbol{\nu})})$ of $Frac(R)$.*

Proof. Part 1: $\Leftarrow$ Assume that h is algebraic over $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \boldsymbol{\nu})})$. Consider the minimal polynomial of h over $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \boldsymbol{\nu})})$. Clearing denominators, we obtain the polynomial $P(Z) := \sum_{i=0}^{k} a_i Z^i$, where each $a_i \in \mathbb{C}\{\mathbf{u}\}[\mathbf{y}^{(\leq \boldsymbol{\nu})}]$ and $P(h) = 0$. By [26, Corollary 6.6], there exist $\Psi \in \tau(\mathbb{C}^{n+\ell})$ and $U \in \tau((\mathbb{C}^{\infty}(0))^r)$ such that for all $\hat{p} \in (\Psi \times U) \cap \Omega_h^0$ it holds that $a_k(0, \hat{p})(t)$ (see Notation 5.3) is not the zero function. Fix such Ψ and U and choose a $\hat{p} \in \Omega_h^0 \cap (\Psi \times U)$. Let V be the complement of the vanishing of $a_k(0, \hat{p})(t)$. Since $a_k(0, \hat{p})(t)$ is analytic about 0, we know that V is codiscrete.

Fix $t_0 \in V \cap D_h(0, \hat{p})$. Let $\tilde{\mathbf{x}}^*$, $\tilde{\boldsymbol{\theta}}$ be such that $\tilde{p} := (\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}, \hat{\mathbf{u}}) \in \Omega_h^{t_0}$ and $Y^{(\leq \boldsymbol{\nu})}(t_0, \tilde{p})(t_0) = Y^{(\leq \boldsymbol{\nu})}(0, \hat{p})(t_0)$. Since each a_i belongs to $\mathbb{C}\{\mathbf{u}\}[\mathbf{y}^{(\leq \boldsymbol{\nu})}]$, it follows from Lemma 5.11 that $a_i(t_0, \tilde{p})(t_0) = a_i(0, \hat{p})(t_0)$ for all i. Applying Lemma 5.11 to the equation $P(h) = 0$, we find that $\sum_{i=1}^{k} a_i(0, \hat{p})(t_0) h(t_0, \tilde{p})(t_0)^i = 0$. Thus $h(t_0, \tilde{p})(t_0)$ is a root of the non-zero polynomial $\sum_{i=1}^{k} a_i(0, \hat{p})(t_0) Z^i$. Noting that $h(\tilde{\mathbf{x}}^*, \tilde{\boldsymbol{\theta}}) = h(t_0, \tilde{p})(t_0)$ we conclude that $S_h(t_0, \hat{p})$ is finite.

Part 2: $\Rightarrow$ Assume that h is not algebraic over $\mathbb{C}\langle\mathbf{u}\rangle(\mathbf{y}^{(\leq\mathbf{v})})$ and that h is $\mathbf{v}$-identifiable. We will give a proof by contradiction. The proof can be divided into four steps:

Step 0 Label the rings that will be used in the proof.

Step 1 Choose $\hat{p} = (\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})$. This will be done in terms of the non-vanishing of minimal polynomials over fraction fields of intermediate rings. Fix V and t_0 and note that Lemma 5.11 with $(t_1, t_2, \hat{p}) = (0, t_0, \hat{p})$ gives a ring homomorphism $\varphi\colon R[h] \to \mathbb{C}$.

Step 2 Show that the set $\{\delta(h) \mid \delta\colon R[h] \to \mathbb{C}$ is a $\mathbb{C}$-algebra homomorphism such that $\delta(\mathbf{u}^{(\mathbb{N})}) = \varphi(\mathbf{u}^{(\mathbb{N})})$ and $\delta(\mathbf{y}^{(\leq\mathbf{v})}) = \varphi(\mathbf{y}^{(\leq\mathbf{v})})\}$ is infinite. This will involve careful extension of ring homomorphisms.

Step 3 Verify that $S_h(t_0, \hat{p})$ is infinite by noting that each δ corresponds to a tuple $(\delta(\mathbf{x}), \delta(\boldsymbol{\theta}), \hat{\mathbf{u}}) \in \Omega_h^{t_0}$.

Step 0 Label $x_1, \ldots, x_n, \theta_1, \ldots, \theta_\ell$ as $b_1, \ldots, b_{n+\ell}$. Without loss of generality, let k be such that $j \leq k$ if and only if b_j is algebraic over $\mathbb{C}\langle\mathbf{u}\rangle(\mathbf{y}^{(\leq\mathbf{v})})$. Let R_2 equal $\mathbb{C}\{\mathbf{u}\}[\mathbf{y}^{(\leq\mathbf{v})}, b_1, \ldots, b_k]$ and let F_2 be its field of fractions. Let μ be such that $b_{k+1}, \ldots, b_\mu, h$ is a transcendence basis for $\mathrm{Frac}(R) = \mathbb{C}\langle\mathbf{u}\rangle(\mathbf{x}, \boldsymbol{\theta})$ over F_2, relabeling if necessary. Let $R_3 = R_2[b_{k+1}, \ldots, b_\mu]$ and let F_3 be its field of fractions. Now $b_{\mu+1}, \ldots, b_{n+\ell}$ are each algebraic over $F_4 := F_2(b_{k+1}, \ldots, b_\mu, h)$. For $j = \mu+1, \ldots, n+\ell$ let $P_j(Z)$ be the minimal polynomial of b_j over F_4 multiplied by the LCM of the denominators, where F_4 is viewed as the field of fractions of $R_4 := R_2[b_{k+1}, \ldots, b_\mu, h]$. For each $j = \mu+1, \ldots, n+\ell$, let $L_j(b_{k+1}, \ldots, b_\mu, h) \in R_4$ be the leading coefficient of P_j. Write $h = n_h/d_h$, where $n_h, d_h \in R$. Let $P_{Qd_h}(Z)$ be the minimal polynomial of $Q \cdot d_h$ over F_4 multiplied by the LCM of the denominators. (Note that if $Qd_h \in F_4$ then $P_{Qd_h}(Z) = Z - Qd_h$.) Let $L_{Qd_h}(b_{k+1}, \ldots, b_\mu, h)$ and $C_{Qd_h}(b_{k+1}, \ldots, b_\mu, h)$ be its leading and constant coefficients, respectively.

$$
\begin{array}{rlrl}
\mathrm{Frac}(R) = & \mathbb{C}\langle\mathbf{u}\rangle(\mathbf{x}, \boldsymbol{\theta}) & R[h] = & \mathbb{C}\{\mathbf{u}\}[\mathbf{x}, \boldsymbol{\theta}, h] \quad (= \mathbb{C}[\boldsymbol{\theta}, h]\{\mathbf{u}, \mathbf{x}, \mathbf{y}\}/\mathcal{J}) \\
 & \vdots & & \vdots \\
 & & & R_4[Qd_h, b_{\mu+1}, b_{\mu+2}] \\
 & & & R_4[Qd_h, b_{\mu+1}] \\
 & F_4(Qd_h) & & R_4[Qd_h] \\
F_4 = & F_3(h) & R_4 := & R_3[h] \\
F_3 = & F_2(b_{k+1}, \ldots, b_\mu) & R_3 := & R_2[b_{k+1}, \ldots, b_\mu] \\
F_2 = & \mathbb{C}\langle\mathbf{u}\rangle\left(\mathbf{y}^{(\leq\mathbf{v})}, b_1, \ldots, b_k\right) & R_2 := & \mathbb{C}\{\mathbf{u}\}\left[\mathbf{y}^{(\leq\mathbf{v})}, b_1, \ldots, b_k\right]
\end{array}
$$

Table 1: Intermediate rings and their fraction fields

Step 1 Let Ψ and U be as guaranteed by the definition of $\mathbf{v}$-identifiability. Consider the subset of Ω_h^0 consisting of all $\hat{p}$ such that for all $j = \mu+1, \ldots, n+\ell-1$ the expressions L_j, L_{Qd_h}, and C_{Qd_h} evaluated at $\hat{p}$ and then $t = 0$ are not zero. This subset is non-empty because it is the intersection of finitely many non-empty Zariski open sets. Hence its intersection with $\Psi \times U$ is non-empty. Fix such a $\hat{p} = (\hat{\mathbf{x}}^*, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}})$. Let V be as required by the definition of $\mathbf{v}$-identifiability, let $t_0 \in V \cap D_h(0, \hat{p})$ (recall Notation 5.3), and let φ be the $\mathbb{C}$-algebra homomorphism from $R[h]$ to $\mathbb{C}$ given by applying the ring homomorphism described in Lemma 5.11 with $(t_1, t_2, \hat{p}) = (0, t_0, \hat{p})$.

Step 2 We now show that there are infinitely many $c \in \mathbb{C}$ such that there exists a $\mathbb{C}$-algebra homomorphism $\delta\colon R[h] \to \mathbb{C}$ such that $\delta(\mathbf{u}^{(\mathbb{N})}) = \varphi(\mathbf{u}^{(\mathbb{N})})$, $\delta(\mathbf{y}^{(\leq\mathbf{v})}) = \varphi(\mathbf{y}^{(\leq\mathbf{v})})$, and $\delta(h) = c$.

Define δ on R_2 by $\delta := \varphi|_{R_2}$. For each j, define $\delta(L_j)(b_{k+1}, \ldots, b_\mu, h) \in \mathbb{C}[b_{k+1}, \ldots, b_\mu, h]$ to be the element obtained by evaluating the coefficients (in R_2) of L_j via δ; define the analogous expressions with L_j replaced by L_{Qd_h} and C_{Qd_h}. By the way $\hat{p}$ was chosen, we have that $\delta(L_{Qd_h})(b_{k+1}, \ldots, b_\mu, h)$, $\delta(C_{Qd_h})(b_{k+1}, \ldots, b_\mu, h)$, and all $\delta(L_j)(b_{k+1}, \ldots, b_\mu, h)$ are non-zero. Using Lemma 5.13 we can extend δ to $R_3 := R_2[b_{k+1}, \ldots, b_\mu]$ in a way that makes neither $\delta(L_{Qd_h})(\delta(b_{k+1}), \ldots, \delta(b_\mu), h)$ nor $\delta(C_{Qd_h})(\delta(b_{k+1}), \ldots, \delta(b_\mu), h)$ nor any $\delta(L_j)(\delta(b_{k+1}), \ldots, \delta(b_\mu), h) \in \mathbb{C}[h]$ equal to zero. Choose such an extension and call this, somewhat abusing notation, δ. Since h is not algebraic over F_3, by Lemma 5.13 δ can be extended by mapping h to any element of $\mathbb{C}$. Now there are infinitely many c such that

$\delta(L_{Qd_h})(\delta(b_{k+1}),\ldots,\delta(b_\mu),c)\cdot\delta(C_{Qd_h})(\delta(b_{k+1}),\ldots,\delta(b_\mu),c)\cdot\prod_j\delta(L_j)(\delta(b_{k+1}),\ldots,\delta(b_\mu),c)\neq 0$. Fix such a c for the remainder of the proof and extend δ to R_4 by letting $\delta(h)=c$.

We now extend δ to $R_4[Qd_h]$. By the preceding discussion, the leading and constant coefficients of $\delta(P_{Qd_h})(Z)$ are non-zero, so by Lemma 5.13 we can extend δ to $R_4[Qd_h]$ so that $\delta(Qd_h)\neq 0$.

Next we extend δ to $R_4[Qd_h,b_{\mu+1}]$. Observe that the minimal polynomial $Q_{\mu+1}(Z)$ of $b_{\mu+1}$ over $F_4(Qd_h)$ is a factor of $P_{\mu+1}(Z)$. Therefore $\delta(Q_{\mu+1})(Z)$ is a non-zero factor of $\delta(P_{\mu+1})(Z)$. By Lemma 5.13 we can extend δ to $R_4[Qd_h,b_{\mu+1}]$.

The argument from the preceding paragraph can be repeated to show that δ can be extended to $R_4[Qd_h,b_{\mu+1},b_{\mu+2}]$. Applying this several more times, we see that we can extend δ to $R[h]$.

Let $T=\{\delta\mid\delta\colon R[h]\to\mathbb{C}$ is a $\mathbb{C}$-algebra homomorphism such that $\delta(\mathbf{u}^{(\mathbb{N})})=\varphi(\mathbf{u}^{(\mathbb{N})})$ and $\delta(\mathbf{y}^{(\leq\mathbf{v})})=\varphi(\mathbf{y}^{(\leq\mathbf{v})})\}$. We have shown that $\{\delta(h)\mid\delta\in T\}$ is infinite.

Step 3 We will now show that $S_h(t_0,\hat{p})$ is infinite. We have

$$S_h(t_0,\hat{p})\supseteq\{h(\delta(\mathbf{x}),\delta(\boldsymbol{\theta}))\mid\exists\delta\in T\ (\delta(\mathbf{x}),\delta(\boldsymbol{\theta}),\hat{\mathbf{u}})\in\Omega_h^{t_0}\text{ and }\\ Y^{(\leq\mathbf{v})}(t_0,\delta(\mathbf{x}),\delta(\boldsymbol{\theta}),\hat{\mathbf{u}})(t_0)=Y^{(\leq\mathbf{v})}(0,\hat{\mathbf{x}}^*,\hat{\boldsymbol{\theta}},\hat{\mathbf{u}})(t_0)\}.$$

By Lemma 5.11 it holds that $Y^{(\leq\mathbf{v})}(t_0,\delta(\mathbf{x}),\delta(\boldsymbol{\theta}),\hat{\mathbf{u}})(t_0)=\delta(\mathbf{y}^{(\leq\mathbf{v})})$ and $Y^{(\leq\mathbf{v})}(0,\hat{\mathbf{x}}^*,\hat{\boldsymbol{\theta}},\hat{\mathbf{u}})(t_0)=\varphi(\mathbf{y}^{(\leq\mathbf{v})})$. Hence

$$S_h(t_0,\hat{p})\supseteq\{\delta(h)\mid\exists\delta\in T\ (\delta(\mathbf{x}),\delta(\boldsymbol{\theta}),\hat{\mathbf{u}})\in\Omega_h^{t_0}\text{ and }\delta(\mathbf{y}^{(\leq\mathbf{v})})=\varphi(\mathbf{y}^{(\leq\mathbf{v})})\}.$$

Since for each $\delta\in T$ we have $\delta(Qd_h)\neq 0$, the first conjunct is satisfied. By the definition of T, the second conjunct is satisfied for all $\delta\in T$. Hence

$$S_h(t_0,\hat{p})\supseteq\{\delta(h)\mid\delta\in T\}.$$

We showed in Step 2 that the right hand side is infinite, and thus $S_h(t_0,\hat{p})$ is infinite. Therefore h is not $\mathbf{v}$-identifiable, contradicting our assumption. □

Lemma 5.13. *Let W be a (possibly infinite) set of indeterminates and let S be a $\mathbb{C}$-subalgebra of the field of rational functions $\mathbb{C}(W)$. Let $\varphi\colon S\to\mathbb{C}$ be a $\mathbb{C}$-algebra homomorphism. Let $f\in\mathbb{C}(W)\backslash\{0\}$.*

Suppose f is not algebraic over $\mathrm{Frac}(S)$. Then φ can be extended to $S[f]$ by mapping f to any element of $\mathbb{C}$.

Suppose f is algebraic over $\mathrm{Frac}(S)$. Let $P(Z)$ be the minimal polynomial of f over $\mathrm{Frac}(S)$ multiplied by the LCM of the denominators. Write $P(Z)=a_kZ^k+\ldots+a_0$, where $a_k,\ldots,a_0\in S$. If $\varphi(a_k)\neq 0$, then φ can be extended to $S[f]$. If furthermore $\varphi(a_0)\neq 0$, then in this extension $\varphi(f)\neq 0$.

Proof. Suppose f is not algebraic over $\mathrm{Frac}(S)$. Our result follows from [27, p. 99].

Suppose f is algebraic over $\mathrm{Frac}(S)$ and $\varphi(a_k)\neq 0$. If $f\in\mathrm{Frac}(S)$ then the result is trivial. Assume f is not in $\mathrm{Frac}(S)$. By [27, Theorem 3.2 p. 347], φ can be extended to $S[f]$ or $S[f^{-1}]$. If the former is true the proof is complete. Suppose φ can be extended to $\varphi_1\colon S[f^{-1}]\to\mathbb{C}$. Writing $P(Z)=\sum_{i=0}^k a_iZ^i$, we have that $\sum_{i=0}^k a_i(f^{-1})^{k-i}=0$. If $\varphi_1(f^{-1})=0$, then it follows that $\varphi_1\left(\sum_{i=0}^k a_i(f^{-1})^{k-i}\right)=\varphi_1(a_k)=0$, which contradicts our hypotheses. Thus $\varphi_1(f^{-1})\neq 0$. Now we can extend φ_1 to the ring $S[f^{-1}]_{\ker\varphi_1}$, which is equal to the subring of $\mathbb{C}(W)$ consisting of fractions with numerators in $S[f^{-1}]$ and denominators in $S[f^{-1}]\backslash\ker\varphi_1$ (cf. [27, p. 346]). Since $f^{-1}\notin\ker\varphi_1$, the element f belongs to $S[f^{-1}]_{\ker\varphi_1}$. By restricting φ_1 to $S[f]$ we have an extension of φ to $S[f]$.

Still assuming f is algebraic over $\mathrm{Frac}(S)$, suppose $\varphi(a_0)\neq 0$. The image of f must satisfy $\sum_{i=0}^k\varphi(a_i)\varphi(f)^i=0$. Thus it is impossible that $\varphi(f)=0$.

□

It is not always obvious whether a given field element is algebraic over a given subfield. The following proposition gives an equivalence that, for our purposes, reduces this question to the problem of checking the rank of a matrix with easily computable entries. We will use the following notation:

Notation 5.14. Let $W = W_1, \ldots, W_a$ be elements of a $\mathbb{C}$-algebra and let $Z = Z_1, \ldots, Z_b$ be an algebraically independent set over $\mathbb{C}$. We denote by $\frac{\partial W}{\partial Z}$ the matrix whose (i, j)-th entry is $\frac{\partial W_i}{\partial Z_j}$. If M is such a matrix and $Z_0 \in Z$, then $M \backslash Z_0$ denotes the result when the column corresponding to Z_0 is removed. If this column does not appear in M then $M \backslash Z_0$ is equal to M.

We will use the following algebraic fact, which generalizes [28, Thm. 2.3].

Proposition 5.15. *Let $p_1, \ldots, p_s$ each be algebraic over $\mathbb{C}(w_1, \ldots, w_a, z_1, \ldots, z_b)$, where $s \leq b$ and all w_i and z_i are indeterminates. The elements $p_1, \ldots, p_s$ are algebraically independent over $\mathbb{C}(w_1, \ldots, w_a)$ if and only if the matrix $\partial(p_1, \ldots, p_s)/\partial(z_1, \ldots, z_b)$ has rank equal to s.*

Proof. Let $M := \partial(p_1, \ldots, p_s)/\partial(z_1, \ldots, z_b)$ and let $M_1 := \partial(w_1, \ldots, w_a, p_1, \ldots, p_s)/\partial(w_1, \ldots, w_a, z_1, \ldots, z_b)$. Note that

$$M_1 = \begin{pmatrix} I & 0 \\ A & M \end{pmatrix},$$

where I is the $a \times a$ identity matrix and A is an $s \times a$ matrix.

Now $p_1, \ldots, p_s$ is algebraically independent over $\mathbb{C}(w_1, \ldots, w_a)$ iff $w_1, \ldots, w_a, p_1, \ldots, p_s$ is algebraically independent over $\mathbb{C}$ iff (by [28, Thm. 2.3]) $\operatorname{rank} M_1 = a + s$ iff $\operatorname{rank} M = s$. □

Proposition 5.16. *Let $\mathbf{v} \in \mathbb{N}^m$ and let $h \in \mathbb{C}(\mathbf{x}, \boldsymbol{\theta})$. Then h is algebraic over $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \mathbf{v})})$ if and only if the matrix $\frac{\partial(\mathbf{y}^{(\leq \mathbf{v})})}{\partial(\mathbf{x}, \boldsymbol{\theta})}$ has the same rank as $\frac{\partial(\mathbf{y}^{(\leq \mathbf{v})}, h)}{\partial(\mathbf{x}, \boldsymbol{\theta})}$.*

Proof. Note that a subset of $\mathbf{y}^{(\leq \mathbf{v})} \cup \{h\}$ is algebraically independent over $\mathbb{C}\langle \mathbf{u} \rangle$ if and only if it is algebraically independent over $\mathbb{C}(\mathbf{u}^{(\leq \mathbf{v}_{max})})$. Let T be a maximal subset of $\mathbf{y}^{(\leq \mathbf{v})}$ that is algebraically independent over $\mathbb{C}\langle \mathbf{u} \rangle$. Then by Proposition 5.15 with $w_1, \ldots, w_a = \mathbf{u}^{(\leq \mathbf{v}_{max})}$ and $z_1, \ldots, z_b = \mathbf{x} \cup \boldsymbol{\theta}$ we have $\operatorname{rank} \frac{\partial(T)}{\partial(\mathbf{x}, \boldsymbol{\theta})} = \operatorname{rank} \frac{\partial(\mathbf{y}^{(\leq \mathbf{v})})}{\partial(\mathbf{x}, \boldsymbol{\theta})} = |T|$.

Suppose h is algebraic over $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \mathbf{v})})$. Then $T \cup \{h\}$ is algebraic over $\mathbb{C}(\mathbf{u}^{(\leq \mathbf{v}_{max})})$. Then $\operatorname{rank} \frac{\partial(\mathbf{y}^{(\leq \mathbf{v})}, h)}{\partial(\mathbf{x}, \boldsymbol{\theta})} = \operatorname{rank} \frac{\partial(T, h)}{\partial(\mathbf{x}, \boldsymbol{\theta})} < |T| + 1$ by Proposition 5.15. Therefore $\operatorname{rank} \frac{\partial(\mathbf{y}^{(\leq \mathbf{v})}, h)}{\partial(\mathbf{x}, \boldsymbol{\theta})} = |T|$.

Suppose h is not algebraic over $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \mathbf{v})})$. Then $T \cup \{h\}$ is not algebraic over $\mathbb{C}(\mathbf{u}^{(\leq \mathbf{v}_{max})})$. Then $\operatorname{rank} \frac{\partial(\mathbf{y}^{(\leq \mathbf{v})}, h)}{\partial(\mathbf{x}, \boldsymbol{\theta})} = \operatorname{rank} \frac{\partial(T, h)}{\partial(\mathbf{x}, \boldsymbol{\theta})} = |T| + 1$ by Proposition 5.15. □

Corollary 5.17. *Algorithm 1 always terminates. The output is "Yes" if and only if h is $\mathbf{v}$-identifiable.*

Proof. Termination is obvious. The other result follows from Proposition 5.12 and Proposition 5.16. □

Proposition 5.18. *Let $\mathbf{v} \in \mathbb{N}^m$ and let J be as in Algorithm 2. Let $z \in \mathbf{x} \cup \boldsymbol{\theta}$. Then z is algebraic over $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \mathbf{v})})$ if and only if* $\operatorname{rank}(J \backslash z) = \operatorname{rank}(J) - 1$.

Proof. Without loss of generality we assume $z = \theta_\ell$. By Proposition 5.16, θ_ℓ is algebraic over $\mathbb{C}\langle \mathbf{u} \rangle(\mathbf{y}^{(\leq \mathbf{v})})$ if and only if $\operatorname{rank} J = \operatorname{rank} \frac{\partial(\mathbf{y}^{(\leq \mathbf{v})}, \theta_\ell)}{\partial(\mathbf{x}, \boldsymbol{\theta})}$. Now

$$\operatorname{rank} \frac{\partial(\mathbf{y}^{(\leq \mathbf{v})}, \theta_\ell)}{\partial(\mathbf{x}, \boldsymbol{\theta})} = \operatorname{rank} \begin{pmatrix} J \\ (0, \ldots, 0, 1) \end{pmatrix} = \operatorname{rank} \begin{pmatrix} J \backslash \theta_\ell & 0 \\ 0 & 1 \end{pmatrix} = \operatorname{rank}(J \backslash \theta_\ell) + 1.$$

□

Corollary 5.19. *Algorithm 2 always terminates. The output is "Yes" if and only if θ_i is $\mathbf{v}$-identifiable.*

Proof. Termination is obvious. The other result follows from Proposition 5.12 and Proposition 5.18. □

5.3 Probabilistic method for improved speed

5.3.1 Presentation of algorithms and summary of results

Algorithms 1 and 2 involve computing the rank of a matrix of rational expressions. It is usually much faster to insert random numbers for the variables and compute the rank of the resulting matrix. The disadvantage of this is that the numerical matrix may have lower rank than the symbolic. In [29] a method for doing this in the case where the coefficients of Σ are integers with user-specified probability of success was given. We adapt that method to our algorithms.

For the rest of this section, assume the coefficients of $\mathbf{f}$ and $\mathbf{g}$ in (2) and h in Algorithm 1 are integers. Note that the entries of J belong to $\mathbb{Q}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u}, \mathbf{u}', \ldots, \mathbf{u}^{(\nu_{max})})$. Our strategy involves randomly choosing $n + \ell + (\nu_{max} + 1)r$ non-negative integers and a prime number, and then evaluating the determinant of our matrix at these integers modulo the prime. Algorithm 3 implements this strategy on Algorithm 1.

Algorithm 3: Determines whether parameter combination h is $\boldsymbol{\nu}$-identifiable

Input : Equations $\mathbf{x}' = \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u})$, $\mathbf{y} = \mathbf{g}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u})$, where $(\mathbf{f}, \mathbf{g}) \notin \mathbb{C}^{n+m}$
$\boldsymbol{\nu} \in \mathbb{N}^m$
$h \in \mathbb{Q}(\mathbf{x}, \boldsymbol{\theta})$
$\alpha \in (0, 1)$

Output: "Yes" if h is $\boldsymbol{\nu}$-identifiable with probability at least α
"No" if h is not $\boldsymbol{\nu}$-identifiable with probability at least α

Step 1a: Compute the least $\mu \in \mathbb{N}$ such that $2 - (1 - 1/\mu)^{-2} \geq \alpha$.

Step 1b: Compute D and the least N as in Proposition 5.25 with $m^* = m + 1$.

Step 1c: Choose $(T, p) \in \{0, 1, 2, \ldots, \mu D\}^{n+\ell+(\nu_{max}+1)r} \times \{p \text{ prime } \mid N + 1 \leq p \leq 2N\}$ uniformly at random.

Step 1d: If $Q(T) \equiv 0 \bmod p$, repeat Step 1c. Otherwise continue to Step 1e.

Step 1e: Compute $J_h(T) \bmod p$.

Step 2: Compute $\operatorname{rank}(J_h(T) \bmod p)$ and $\operatorname{rank}(J(T) \bmod p)$.

Step 3: If $\operatorname{rank}(J_h(T) \bmod p) = \operatorname{rank}(J(T) \bmod p)$, output "Yes". Otherwise output "No".

While Algorithm 2 determines whether an individual parameter is $\boldsymbol{\nu}$-identifiable, Algorithm 4 uses this

concept to determine, with user-specified probability, all elements of $\mathbf{x} \cup \boldsymbol{\theta}$ that are $\boldsymbol{\nu}$-identifiable.

Algorithm 4: Determines the $\boldsymbol{\nu}$-identifiable subset of $\mathbf{x} \cup \boldsymbol{\theta}$

Input : Equations $\mathbf{x}' = \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u})$, $\mathbf{y} = \mathbf{g}(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u})$, where $(\mathbf{f}, \mathbf{g}) \notin \mathbb{C}^{n+m}$
$\boldsymbol{\nu} \in \mathbb{N}^m$
$\alpha \in (0,1)$

Output: Subset of $\mathbf{x} \cup \boldsymbol{\theta}$ consisting of exactly the $\boldsymbol{\nu}$-identifiable elements, with probability at least α

Step 1a: Compute the least $\mu \in \mathbb{N}$ such that $2 - (1 - 1/\mu)^{-2} \geq \alpha$.

Step 1b: Compute D and the least N as in Proposition 5.25 with $m^* = m$.

Step 1c: Choose $(T, p) \in \{0, 1, 2, \ldots, \mu D\}^{n+\ell+(\nu_{max}+1)r} \times \{p \text{ prime } \mid N+1 \leq p \leq 2N\}$ uniformly at random.

Step 1d: If $Q(T) \equiv 0 \bmod p$, repeat Step 1c. Otherwise continue to Step 1e.

Step 1e: Compute $J(T) \bmod p$.

Step 2: Compute $\operatorname{rank}(J(T) \bmod p)$.

Step 3: Let $S_{out} = \emptyset$.

For $z \in \mathbf{x} \cup \boldsymbol{\theta}$

Compute $\operatorname{rank}(J(T) \backslash z \bmod p)$.

If $\operatorname{rank}(J(T) \backslash z \bmod p) < \operatorname{rank}(J(T) \bmod p)$, add z to S_{out}.

End For

Step 4: Return S_{out}.

The main results on Algorithm 3 and Algorithm 4 are the following:

1. The expected time to reach Step 1e is negligible. Once Step 1e is reached the algorithm is guaranteed to terminate. (Proposition 5.29)

2. The probability that the output is correct is at least α. (Propositions 5.30 and 5.31)

3. For Algorithm 3, if in Step 2 $\operatorname{rank}(J(T) \bmod p) = n + \ell$, then h is $\boldsymbol{\nu}$-identifiable. For Algorithm 4, if in Step 2 $\operatorname{rank}(J(T) \bmod p) = n + \ell$, then every element of $\mathbf{x} \cup \boldsymbol{\theta}$ is $\boldsymbol{\nu}$-identifiable. (Proposition 5.32)

Remark 5.20. At the beginning of this section, we assumed n, m, and r are positive integers. As noted in Example 5.5, ℓ and r must be at least equal to the number of parameters and outputs, respectively, that appear in Σ, but can be chosen to be greater without changing identifiability results. In Algorithms 3 and 4 it is sufficient to use ℓ and r equal to the number of parameters and inputs, respectively, appearing in the equations. In particular one can use $r = 0$ if no inputs appear and the main results on the algorithms are correct.

Remark 5.21. We have presented Algorithms 3 and 4 only for the case where $\mathbf{f}$ and $\mathbf{g}$ are not all elements of $\mathbb{C}$, since removing this restriction would require addressing special cases in several of the proofs. If $(\mathbf{f}, \mathbf{g}) \in \mathbb{C}^{n+m}$, then h is $\boldsymbol{\nu}$-identifiable if and only if $h \in \mathbb{C}$. One could easily add a step at the beginning of either algorithm to accommodate this case.

Remark 5.22. Our algorithms do not specify the methods used to compute the ranks of matrices. One can use the state-of-the-art method for this to achieve maximum speed.

5.3.2 Proof of algorithms

In Algorithms 3 and 4, a random tuple of integers T and a random prime p are chosen, we check that the denominator of a determinant evaluated at T does not vanish modulo p, and then calculate the rank of a matrix after evaluating at T modulo p. We show that this gives the correct results with user-specified probability. The main results on Algorithms 3 and 4 are Propositions 5.29, 5.30, 5.31, and 5.32. The theory

is based on bounds on integer roots of polynomials with integer coefficients, as well as the distribution of the prime numbers.

First, we use Proposition 5.23, Lemma 5.24, and Lemma 5.26 to prove Proposition 5.25, which gives conditions on the sets from which we choose T and p so that the numerator of the determinant evaluated at T does not vanish modulo p with user-specified probability. This proposition is essentially a more precisely stated version of [29, Proposition 6], and the proof we give is outlined in [29].

Next, Proposition 5.28 shows that the probability that the numerator vanishes given that the denominator does not vanish can be specified by the user. Note that this is the true probability associated with the algorithm, since we must first check that our choice of (T,p) does not make the denominator vanish before proceeding. This issue is not addressed in similar algorithms (cf. [8, p. 739], [21]) and is non-trivial, as shown by Example 5.27.

Finally, we prove statements about the algorithms that are directly relevant to helping the user interpret them. Although in principle, arbitrarily many instances of (T,p) may need to be chosen before finding one that does not make the denominator vanish, Proposition 5.29 shows that the expected time for a successful choice is negligible. It also asserts the algorithm's termination after such a successful choice. Propositions 5.30 and 5.31 show that the algorithms produce the correct result with user-specified probability. Proposition 5.32 states that when the rank of the specialized matrix is full, the algorithms output the correct result with certainty.

Proposition 5.23 ([30] Prop. 98 p. 192)**.** *Let $P(Z_1,\dots,Z_k)$ be a polynomial of total degree D over an integral domain A. Let $S\subseteq A$. If an element $(z_1,\dots,z_k)$ is chosen from S^k uniformly at random, then*

$$\text{prob}(P(z_1,\dots,z_k)=0)\leq\frac{D}{|S|}.$$

Lemma 5.24 ([31] Lemma 18.9 p. 525)**.** *Let $S\subseteq\mathbb{N}$ be a nonempty finite set of prime numbers, let $M\in\mathbb{Z}\backslash\{0\}$, and let $C\geq|M|$. If p is chosen from S uniformly at random, then*

$$\text{prob}(p \textit{ divides } M)\leq\frac{\ln C}{|S|\ln\min S}.$$

Proposition 5.25. *Let J (resp. J_h) be as in Algorithm 2 (resp. Algorithm 1) and suppose J (resp. J_h) has at least one non-zero entry. Let*

1. $\mu\in\mathbb{N}^+$
2. $d_1=$ *maximum degree of the numerators and denominators of* $\mathbf{f}$ *and* $\mathbf{g}$ *(resp.* $\mathbf{f}$, $\mathbf{g}$ *and* h*)*
3. $d_2=\max\{\ln(|a|+1):a$ *is a coefficient in* $\mathbf{f}$ *or* $\mathbf{g}$ *(resp.* $\mathbf{f}$, $\mathbf{g}$, *or* h*)*$\}$
4. $D=(n+\ell)(2\nu_{max}+3)(n+m^*)d_1,\quad$ *where* $m^*=m$ *(resp.* $m^*=m+1$*)*
5. C *be such that* $\ln C=(2\ln(n+\ell+r+1)+\ln(\mu D+1))D+(n+\ell)(2\nu_{max}+3)((n+m^*)d_2+\ln(2nD))$, *where* $m^*=m$ *(resp.* $m^*=m+1$*)*
6. $N\in\mathbb{N}$ *such that* $N\geq 2\mu\ln C$
7. $S=\{p$ *prime* $\mid N+1\leq p\leq 2N\}$

Let J_0 be a square submatrix of J (resp. J_h) with rank equal to that of J (resp. J_h). Note that the numerator of $\det J_0$ *lies in* $\mathbb{Z}[\mathbf{x},\boldsymbol{\theta},\mathbf{u},\dots,\mathbf{u}^{(\nu_{max})}]$. *If a tuple of values T of the variables is chosen uniformly at random from* $\{0,1,2,\dots,\mu D\}^{n+\ell+(\nu_{max}+1)r}$ *and p is chosen uniformly at random from S, then*

$$\text{prob}(p \textit{ does not divide } \operatorname{num}\det J_0(T))\geq\left(1-\frac{1}{\mu}\right)^2,$$

where $\operatorname{num}\det J_0(T)$ *represents the specialization of the numerator of* $\det J_0$ *at the chosen values.*

Proof. We prove the main version first. The version for J_h will follow quickly from this.

Note that since J has at least one non-zero entry, d_1 is well-defined and positive, and hence D is positive. Let W denote $\operatorname{num}\det J_0(T)$. We have that

$$\text{prob}(p \text{ does not divide } W) = \text{prob}(W \neq 0) \cdot \text{prob}(p \text{ does not divide } W \text{ given } W \neq 0). \tag{5}$$

By Proposition 5.23, we have that

$$\text{prob}(W = 0) \leq \frac{\deg \operatorname{num}\det J_0}{\mu D + 1}.$$

We show that the degree of the numerator of any entry of J is no greater than $(2\nu_{max}+3)(n+m)d_1$. Fix $j \in \{1,\ldots,m\}$. We first show by induction that for $k \in \mathbb{N}$ we can write $y_j^{(k)} = A_k/Q^{2k+1}$, where $\deg A_k \leq (2k+1)(n+m)d_1$. For the base case $k=0$, recall from Notation 5.1 we have that $y_j = G_j/Q$, and $\deg G_j \leq (n+m)d_1$. For the inductive hypothesis, fix $k \geq 0$ and assume $y_j^{(k)} = A_k/Q^{2k+1}$ with $\deg A_k \leq (2k+1)(n+m)d_1$. Applying the quotient rule we have that

$$y_j^{(k+1)} = \frac{1}{Q^{4k+2}} \left(Q^{2k+1} \sum_{i=1}^{n} \frac{\partial A_k}{\partial x_i} \frac{F_i}{Q} - A_k(2k+1)Q^{2k} \sum_{i=1}^{n} \frac{\partial Q}{\partial x_i} \frac{F_i}{Q} \right), \tag{6}$$

recalling from Notation 5.1 that each $f_i = F_i/Q$. Since $\deg F_i$ and $\deg Q$ do not exceed $(n+m)d_1$, we conclude the inductive step. Continuing the proof of the bound on the degrees of the numerators of the entries of J, note that for any $z \in \mathbf{x} \cup \boldsymbol{\theta}$ we can write $\frac{\partial}{\partial z}(y_j^{(k)}) = \frac{1}{Q^{4k+2}} \left(Q^{2k+1} \frac{\partial A_k}{\partial z} - A_k(2k+1)Q^{2k} \frac{\partial Q}{\partial z} \right)$ where $\deg A_k \leq (2k+1)(n+m)d_1$. So we can write $\frac{\partial}{\partial z}(y_j^{(k)}) = A/Q^{2k+2}$ where $\deg A \leq (2k+2)(n+m)d_1$. In J_0 the value of k does not exceed ν_{max} so we conclude the argument.

Since J_0 is square it has at most $n+\ell$ rows. Therefore $\deg \operatorname{num}\det J_0$ is no greater than D. Thus, we have

$$\text{prob}(W = 0) \leq \frac{D}{\mu D + 1} \leq \frac{1}{\mu}. \tag{7}$$

Assume that $W \neq 0$. Lemma 5.26 below shows that $C \geq |W|$. Our assumptions imply that $\ln C \geq 3$ and hence $N \geq 6$. From [31, Exercise 18.18] it follows that S is non-empty. Using Lemma 5.24 with $M = W$, we have

$$\text{prob}(p \text{ divides } W) \leq \frac{\ln C}{|S| \ln \min S} \leq \frac{\ln C}{|S| \ln N}.$$

By [31, Exercise 18.18], we have $|S| > N/(2\ln N)$. It follows that

$$\text{prob}(p \text{ divides } W) \leq \frac{2\ln C}{N} \leq \frac{2\ln C}{2\mu \ln C} = \frac{1}{\mu}. \tag{8}$$

Combining (5), (7), and (8), we have our result.

We now prove the version with J_h. Let Σ_1 be the system obtained by adding the equation $y_{m+1} = h$ to Σ and set $\nu_{m+1} = 0$. Now J_h for Σ is equal to J for Σ_1. Our result follows from the main version of the proposition. □

Lemma 5.26. *In the setup of Proposition 5.25,* $\ln C \geq \ln |W|$.

Proof. For a polynomial p with integer coefficients, we shall define the height of p as $\text{ht}(p) := \max\{\ln(|a|+1) \mid a \text{ is a coefficient of } p\}$. We will use the following properties ([29, Lemma 1]): For polynomials $p_1,\ldots,p_s$ in $n+\ell+r$ variables, tuple of integers T_0, and partial derivative ∂, the following hold:

1. $\mathrm{ht}(\partial p_1) \leq \mathrm{ht}(p_1) + \ln \deg p_1$,
2. $\mathrm{ht}(p_1 + \ldots + p_s) \leq \max\{\mathrm{ht}(p_1), \ldots, \mathrm{ht}(p_s)\} + \ln s$,
3. $\mathrm{ht}(p_1 p_2) \leq \min\{\deg p_1, \deg p_2\} \cdot \ln(n+\ell+r+1) + \mathrm{ht}(p_1) + \mathrm{ht}(p_2)$, and
4. $\mathrm{ht}(p_1(T_0)) \leq \max\{\mathrm{ht}(t_0) \mid t_0 \in T_0\} \cdot \deg(p_1) + \mathrm{ht}(p_1)$.

Fix $j \in \{0, \ldots, m\}$ and write $y_j^{(k)} = A_k / Q^{2k+1}$, we will prove by induction that for all $k \in \mathbb{N}$

$$\mathrm{ht}(A_k) \leq (2k+1)(n+m)\left(2d_1 \ln(n+\ell+r+1) + d_2\right) + (k+1)\ln(2n(n+m)d_1) + (2k+1)\ln(2k+1). \quad (9)$$

We noted earlier that $\deg Q$, all $\deg F_i$, and $\deg G_j$ are no greater than $(n+m)d_1$. It follows from this and the product height property that $\mathrm{ht}(Q)$, $\mathrm{ht}(F_i)$, $\mathrm{ht}(G_j) \leq (n+m)\left(d_1 \ln(n+\ell+r+1) + d_2\right)$. For conciseness, we will use d_1^* and d_2^* to denote these degree and height bounds, respectively.

For the base case $k = 0$, note that $A_0 = G_j$. For the inductive hypothesis, fix $k \geq 0$ and suppose A_k satisfies (9). As shown in the (6), we have

$$A_{k+1} = \sum_{i=1}^{n} F_i \left(Q \cdot \frac{\partial A_k}{\partial x_i} - (2k+1) \cdot A_k \cdot \frac{\partial Q}{\partial x_i} \right).$$

Using the derivation and product height properties, as well as the bound $\deg A_k \leq (2k+1)d_1^*$, we have

$$\mathrm{ht}\left(Q \cdot \frac{\partial A_k}{\partial x_i} \right) \leq d_1^* \ln(n+\ell+r+1) + d_2^* + \mathrm{ht}(A_k) + \ln\left((2k+1)d_1^*\right)$$
$$\text{and } \mathrm{ht}\left((2k+1) \cdot A_k \cdot \frac{\partial Q}{\partial x_i} \right) \leq d_1^* \ln(n+\ell+r+1) + \ln(2k+2) + \mathrm{ht}(A_k) + d_2^* + \ln d_1^*.$$

By the sum height property we have that

$$\mathrm{ht}\left(Q \cdot \frac{\partial A_k}{\partial x_i} - (2k+1) \cdot A_k \cdot \frac{\partial Q}{\partial x_i} \right) \leq d_1^* \ln(n+\ell+r+1) + d_2^* + \mathrm{ht}(A_k) + \ln\left((2k+2)d_1^*\right) + \ln 2. \quad (10)$$

By the product height property we have

$$\mathrm{ht}\left(F_i \cdot \left(Q \cdot \frac{\partial A_k}{\partial x_i} - (2k+1) \cdot A_k \cdot \frac{\partial Q}{\partial x_i} \right) \right) \leq d_1^* \ln(n+\ell+r+1) + (\text{RHS of } (10)) + d_2^*. \quad (11)$$

By the sum height property we have

$$\begin{aligned}
\mathrm{ht}(A_{k+1}) &\leq 2d_1^* \ln(n+\ell+r+1) + 2d_2^* + \mathrm{ht}(A_k) + \ln\left(2(2k+2)d_1^*\right) + \ln n \\
&= 2(n+m)\left(2d_1 \ln(n+\ell+r+1) + d_2\right) + \mathrm{ht}(A_k) + \ln\left(2n(n+m)d_1\right) + \ln(2k+2) \\
&\leq (2k+3)(n+m)\left(2d_1 \ln(n+\ell+r+1) + d_2\right) + (k+2)\ln\left(2n(n+m)d_1\right) \\
&\quad + (2k+1)\ln(2k+1) + \ln(2k+2).
\end{aligned} \quad (12)$$

Noting that $(2k+1)\ln(2k+1) + \ln(2k+2) \leq (2k+3)\ln(2k+3)$, we conclude the inductive step.

Now for $z \in \mathbf{x} \cup \boldsymbol{\theta}$, we have that $\frac{\partial}{\partial z}(y_j^{(k)})$ is equal to $\left(Q \cdot \frac{\partial A_k}{\partial z} - (2k+1) \cdot A_k \cdot \frac{\partial Q}{\partial z} \right) / Q^{2k+2}$. A bound on the height of the numerator is given by the right-hand side of (10). Since RHS of (10) $\leq$ RHS of (11) $\leq$ RHS of (12), we see that a bound on this height is also given by the RHS of (9) with k replaced by $k+1$.

Observing that the maximum value of k occurring for an entry in J_0 is ν_{max} and applying the evaluation height property, we have that the numerator of each entry of $J_0(T)$ is bounded by

$$\begin{aligned} B :=& (2\nu_{max}+3)(n+m)\left((2\ln(n+\ell+r+1)+\ln(\mu D+1))\,d_1+d_2\right) \\ &+(\nu_{max}+2)\ln\left(2n(n+m)d_1\right)+(2\nu_{max}+3)\ln(2\nu_{max}+3). \end{aligned}$$

Since our bounds assume that entries in the same row have the same denominator, we have that $\operatorname{num}\det J_0(T) = \det(\operatorname{num} J_0(T))$, where $\operatorname{num} J_0$ is the matrix whose elements are the numerators of J_0. Henceforth we assume J_0 has polynomial entries with heights bounded by B when evaluated at T. Using Hadamard's Theorem, we have that $|W| \leq \prod_{i=1}^{\operatorname{rank} J_0} ||v_i||$, where $||v_i||$ is the square root of the sum of the squares of the elements of the i-th column of $J_0(T)$. Hence $||v_i|| \leq \sqrt{n+\ell}\cdot e^B$. Thus

$$\begin{aligned} \ln|W| \leq& (n+\ell)\Bigg((2\nu_{max}+3)(n+m)\Big(\big(2\ln(n+\ell+r+1)+\ln(\mu D+1)\big)d_1+d_2\Big) \\ &+(\nu_{max}+2)\ln\left(2n(n+m)d_1\right)+(2\nu_{max}+3)\ln(2\nu_{max}+3)+\frac{1}{2}\ln(n+\ell)\Bigg). \end{aligned}$$

Recalling that

$$\begin{aligned} \ln C =& \Big(2\ln(n+\ell+r+1)+\ln(\mu D+1)\Big)(n+\ell)(2\nu_{max}+3)(n+m)d_1 \\ &+(n+\ell)(2\nu_{max}+3)\Bigg((n+m)d_2+\ln\Big(2n(n+\ell)(2\nu_{max}+3)(n+m)d_1\Big)\Bigg), \end{aligned}$$

we conclude the proof. □

After a set of random integers and a random prime number is chosen, we must first check that no denominator vanishes modulo the prime number before evaluating the rank of the matrix. Thus the probability that the algorithm gives the correct answer is not simply the probability that p does not divide $\operatorname{num}\det J_0(T)$, but rather the probability that p does not divide $\operatorname{num}\det J_0(T)$ given that p does not divide $\operatorname{denom}\det J_0(T)$. If a random tuple T is chosen and used to evaluate two polynomials A and B, it is not necessarily the case that $\operatorname{prob}(A(T)\neq 0) \leq \operatorname{prob}(A(T)\neq 0 \text{ given } B(T)\neq 0)$, as the following example shows:

Example 5.27. Let k be a positive integer and let $A(Z)=Z$ and $B(Z)=(Z-1)(Z-2)\cdot\ldots\cdot(Z-k)$. If an integer T is chosen uniformly at random from $\{0,1,2,\ldots,k\}$, then $\operatorname{prob}(A(T)\neq 0)=\frac{k}{k+1}$ and $\operatorname{prob}(A(T)\neq 0 \text{ given } B(T)\neq 0)=0$.

The following proposition shows that the conditions used for Proposition 5.25 also give a bound on the conditional probability that is relevant to our algorithms.

Proposition 5.28. *Let J (resp. J_h) be as in Algorithm 2 (resp. Algorithm 1) and suppose J (resp. J_h) has at least one non-zero entry. Let μ, D, S, and J_0 be as in Proposition 5.25. If a tuple of values T of the variables is chosen uniformly at random from $\{0,1,2,\ldots,\mu D\}^{n+\ell+(\nu_{max}+1)r}$ and p is chosen uniformly at random from S, then*

$$\operatorname{prob}(p \textit{ does not divide } \operatorname{num}\det J_0(T) \textit{ given } p \textit{ does not divide } Q(T)) \geq 2-\left(1-\frac{1}{\mu}\right)^{-2},$$

where $\operatorname{num}\det J_0(T)$ *represents the specialization of the numerator of* $\det J_0$ *at the chosen values.*

Proof. Let M denote the sample space $\{0,\dots,\mu D\}^{n+\ell+(\nu_{max}+1)r}\times S$. Let

$$A=\{(T,p)\in M \mid p \text{ divides } \operatorname{num}\det J_0(T)\},\quad B=\{(T,p)\in M \mid p \text{ divides } Q(T)\}.$$

Now

$$\operatorname{prob}((T,p)\notin A \text{ given } (T,p)\notin B)=\frac{|M|-|A|-|B|+|A\cap B|}{|M|-|B|}\geq 1-\frac{|A|}{|M|-|B|}.$$

By Proposition 5.25 we have $|A|/|M|\leq 1-\left(1-\frac{1}{\mu}\right)^2$. We now prove the same bound holds for $|B|/|M|$. Now $\deg Q\leq (n+m)d_1$, so using Proposition 5.23 we have $\operatorname{prob}(Q(T)=0)\leq D/|S|=D/(\mu D+1)\leq 1/\mu$. By the first statement of [29, Prop. 3] with $j=0$ we have that the height of $Q(T)$ is no greater than $(n+m)(d_2+d_1(\ln(\mu D+1)+2\ln(n+\ell+r+1)))$, which is no greater than $\ln C$. Thus $C\geq |Q(T)|$ and if $Q(T)\neq 0$ then by Lemma 5.24 we have $\operatorname{prob}(p \text{ divides } Q(T))\leq \ln C/(|S|N)$, which by [31, Exercise 18.18] is no greater than $1/\mu$. Thus $|B|/|M|\leq 1-\left(1-\frac{1}{\mu}\right)^2$. Now we have

$$1-\frac{|A|}{|M|-|B|}\geq 1-\frac{1-\left(1-\frac{1}{\mu}\right)^2}{\left(1-\frac{1}{\mu}\right)^2}.$$

□

Proposition 5.29. *Fix the input to Algorithm 3 (or Algorithm 4). (i) The probability that Step 1e will be reached in no more than k iterations of Step 1c is at least $1-\left(\frac{1-\alpha}{2-\alpha}\right)^k$. (ii) When Step 1e is reached the algorithm is guaranteed to terminate.*

Proof. (i) It was shown in the proof of Proposition 5.28 that $\operatorname{prob}(Q(T)\equiv 0 \bmod p)\leq 1-(1-\frac{1}{\mu})^2$, and based on the way μ was chosen in Step 1a it is trivial to verify that this is no greater than $\frac{1-\alpha}{2-\alpha}$. Therefore for k independent choices the probability that $Q(T)\not\equiv 0 \bmod p$ for at least one of them is at least $1-\left(\frac{1-\alpha}{2-\alpha}\right)^k$.

(ii) This is obvious. □

Proposition 5.30. *Fix the input to Algorithm 3. If h is $\boldsymbol{\nu}$-identifiable, the probability that the output is "Yes" is at least α. If h is not $\boldsymbol{\nu}$-identifiable, the probability that the output is "No" is at least α.*

Proof. Consider the following subsets of $\{(T,p)\mid Q(T)\not\equiv 0 \bmod p\}$: $Y=\{(T,p)\mid \text{output is "Yes"}\}$, $N=\{(T,p)\mid \text{output is "No"}\}$, $E=\{(T,p)\mid \operatorname{rank}(J(T) \bmod p)=\operatorname{rank}J\}$, and $E_h=\{(T,p)\mid \operatorname{rank}(J_h(T) \bmod p)=\operatorname{rank}J_h\}$. We have

$$\operatorname{prob}(Y)\geq \operatorname{prob}(Y \text{ given } E)\operatorname{prob}(E)\quad\text{and}\quad \operatorname{prob}(N)\geq \operatorname{prob}(N \text{ given } E_h)\operatorname{prob}(E_h).$$

Proposition 5.28 gives us that $\operatorname{prob}(E)\geq\alpha$ and $\operatorname{prob}(E_h)\geq\alpha$.

Suppose h is $\boldsymbol{\nu}$-identifiable. By Propositions 5.12 and 5.16, we have $\operatorname{rank}J_h=\operatorname{rank}J$. We show that $\operatorname{prob}(Y \text{ given } E)=1$. Suppose $(T,p)\in E$. Then $\operatorname{rank}(J_h(T) \bmod p)\leq \operatorname{rank}J_h=\operatorname{rank}J=\operatorname{rank}(J(T) \bmod p)\leq \operatorname{rank}(J_h(T) \bmod p)$, and we deduce that $\operatorname{rank}(J_h(T) \bmod p)=\operatorname{rank}(J(T) \bmod p)$, so the algorithm outputs "Yes".

Suppose h is not $\boldsymbol{\nu}$-identifiable. By Propositions 5.12 and 5.16, we have $\operatorname{rank}J_h>\operatorname{rank}J$. We show that $\operatorname{prob}(N \text{ given } E_h)=1$. Suppose $(T,p)\in E_h$. Then $\operatorname{rank}(J_h(T) \bmod p)=\operatorname{rank}J_h>\operatorname{rank}J\geq \operatorname{rank}(J(T) \bmod p)$, so the algorithm outputs "No". □

Proposition 5.31. *Fix the input to Algorithm 4. Let S_{id} denote the $\boldsymbol{\nu}$-identifiable subset of $\mathbf{x}\cup\boldsymbol{\theta}$ and let S_{out} denote the output of Algorithm 4. The probability that $S_{out}=S_{id}$ is at least α.*

Proof. Because S_{out} depends on (T,p) we shall use the notation $S_{out}(T,p)$. Consider the following subsets of $\{(T,p) \mid Q(T) \not\equiv 0 \bmod p\}$: $Y = \{(T,p) \mid S_{out}(T,p) = S_{id}\}$ and $E = \{(T,p) \mid \operatorname{rank}(J(T) \bmod p) = \operatorname{rank} J\}$. We have

$$\operatorname{prob}(Y) \geq \operatorname{prob}(Y \text{ given } E)\operatorname{prob}(E).$$

By Proposition 5.28 we have $\operatorname{prob}(E) \geq \alpha$. We show $\operatorname{prob}(Y \text{ given } E) = 1$. Suppose $(T,p) \in E$. Let $z \in \mathbf{x} \cup \boldsymbol{\theta}$. Suppose z is $\boldsymbol{\nu}$-identifiable. By Propositions 5.12 and 5.16, we have $\operatorname{rank} J_z = \operatorname{rank} J$. Now $\operatorname{rank}(J_z(T) \bmod p) \leq \operatorname{rank} J_z = \operatorname{rank} J = \operatorname{rank}(J(T) \bmod p) \leq \operatorname{rank}(J_z(T) \bmod p)$. Since the final row of $J_z(T)$ must be $(0\ldots1\ldots0)$, it follows that $\operatorname{rank}(J(T)\backslash z \bmod p) = \operatorname{rank}(J_z(T)\backslash z \bmod p) = \operatorname{rank}(J_z(T) \bmod p) - 1 = \operatorname{rank}(J(T) \bmod p) - 1$. Therefore $z \in S_{out}(T,p)$. Suppose z is not $\boldsymbol{\nu}$-identifiable. It follows from Propositions 5.12 and 5.16 that $\operatorname{rank} J_z - 1 = \operatorname{rank} J$. Because $(T,p) \in E$ and the final row of $J_z(T)$ is $(0\ldots1\ldots0)$ we have that $\operatorname{rank}(J_z(T) \bmod p) = \operatorname{rank} J_z$. Now $\operatorname{rank}(J(T)\backslash z \bmod p) = \operatorname{rank}(J_z(T)\backslash z \bmod p) = \operatorname{rank}(J_z(T) \bmod p) - 1 = \operatorname{rank} J_z - 1 = \operatorname{rank} J$. Therefore $z \notin S_{out}(T,p)$. □

Proposition 5.32. *Suppose that in Step 2 of Algorithm 3 (resp. Algorithm 4)* $\operatorname{rank}(J(T) \bmod p) = n + \ell$. *Then h is $\boldsymbol{\nu}$-identifiable and the output is "Yes" (resp. every element of $\mathbf{x} \cup \boldsymbol{\theta}$ is $\boldsymbol{\nu}$-identifiable and the output is $\mathbf{x} \cup \boldsymbol{\theta}$).*

Proof. We first prove the statement regarding Algorithm 3. We have $n + \ell \geq \operatorname{rank} J \geq \operatorname{rank}(J(T) \bmod p)$, so $\operatorname{rank} J = n + \ell$. It follows that $\operatorname{rank} J_h = n + \ell$ and by Propositions 5.12 and 5.16 we have that h is $\boldsymbol{\nu}$-identifiable. Since $\operatorname{rank}(J(T) \bmod p) \leq \operatorname{rank}(J_h(T) \bmod p) \leq n + \ell$, it follows that $\operatorname{rank}(J(T) \bmod p) = \operatorname{rank}(J_h(T) \bmod p)$ and the output will be "Yes".

We now address Algorithm 4. By the preceding paragraph $\operatorname{rank} J = n + \ell$. Since for any $z \in \mathbf{x} \cup \boldsymbol{\theta}$ the matrix $J\backslash z$ has only $n + \ell - 1$ columns, it must be that $\operatorname{rank} J\backslash z < \operatorname{rank} J$ and hence by Propositions 5.12 and 5.18 each element of $\mathbf{x} \cup \boldsymbol{\theta}$ is $\boldsymbol{\nu}$-identifiable. Similarly, $\operatorname{rank}(J(T)\backslash z \bmod p) < \operatorname{rank}(J(T) \bmod p)$ and hence the output is $\mathbf{x} \cup \boldsymbol{\theta}$. □

References

[1] G. Cedersund and J. Roll, "Systems biology: Model based evaluation and comparison of potential explanations for given biological data," *FEBS Journal*, vol. 276, no. 4, pp. 903–922, 2009. [Online]. Available: https://doi.org/10.1111/j.1742-4658.2008.06845.x

[2] G. Cedersund, "Conclusions via unique predictions obtained despite unidentifiability – new definitions and a general method," *The FEBS Journal*, vol. 279, no. 18, pp. 3513–3527, 2012. [Online]. Available: https://doi.org/10.1111/j.1742-4658.2012.08725.x

[3] F.-G. Wieland, A. L. Hauber, M. Rosenblatt, C. Tönsing, and J. Timmer, "On structural and practical identifiability," *Current Opinion in Systems Biology*, vol. 25, pp. 60–69, 2021. [Online]. Available: https://doi.org/10.1016/j.coisb.2021.03.005

[4] G. Cedersund, *Prediction Uncertainty Estimation Despite Unidentifiability: An Overview of Recent Developments*. Cham: Springer International Publishing, 2016, pp. 449–466. [Online]. Available: https://doi.org/10.1007/978-3-319-21296-8_17

[5] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer, "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood," *Bioinformatics*, vol. 25, no. 15, pp. 1923–1929, aug 2009. [Online]. Available: https://doi.org/10.1093/bioinformatics/btp358

[6] A. F. Villaverde, A. Barreiro, and A. Papachristodoulou, "Structural identifiability of dynamic systems biology models," *PLOS Computational Biology*, vol. 12, no. 10, pp. 1–22, 10 2016. [Online]. Available: https://doi.org/10.1371/journal.pcbi.1005153

[7] A. Raue, M. Schilling, J. Bachmann, A. Matteson, M. Schelker, D. Kaschek, S. Hug, C. Kreutz, B. Harms, F. Theis, U. Klingmüller, and J. Timmer, "Lessons learned from quantitative dynamical modeling in systems biology," *PLOS ONE*, vol. 8, p. e74335, 12 2013. [Online]. Available: https://doi.org/10.1371/journal.pone.0074335

[8] A. Sedoglavic, "A probabilistic algorithm to test local algebraic observability in polynomial time," *Journal of Symbolic Computation*, vol. 33, no. 5, pp. 735–755, 2002. [Online]. Available: https://doi.org/10.1006/jsco.2002.0532

[9] 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*, vol. 88, no. 1, pp. 52–61, 2007. [Online]. Available: https://doi.org/10.1016/j.cmpb.2007.07.002

[10] H. Hong, A. Ovchinnikov, G. Pogudin, and C. Yap, "Global identifiability of differential models," *Communications on Pure and Applied Mathematics*, vol. 73, no. 9, pp. 1831–1879, 2020. [Online]. Available: https://doi.org/10.1002/cpa.21921

[11] N. Meshkat, C. E. Kuo, and J. DiStefano, III, "On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and COMBOS: A novel web implementation," *PLOS ONE*, vol. 9, no. 10, pp. 1–14, 10 2014. [Online]. Available: https://doi.org/10.1371/journal.pone.0110261

[12] A. Ovchinnikov, G. Pogudin, and P. Thompson, "Parameter identifiability and input–output equations," *Applicable Algebra in Engineering, Communication and Computing*, 2 2021. [Online]. Available: https://doi.org/10.1007/s00200-021-00486-8

[13] A. Ovchinnikov, A. Pillay, G. Pogudin, and T. Scanlon, "Computing all identifiable functions of parameters for ODE models," *Systems & Control Letters*, vol. 157, p. 105030, 2021. [Online]. Available: https://doi.org/10.1016/j.sysconle.2021.105030

[14] N. Meshkat, S. Sullivant, and M. Eisenberg, "Identifiability results for several classes of linear compartment models," *Bulletin of Mathematical Biology*, vol. 77, no. 8, pp. 1620–1651, Aug 2015. [Online]. Available: https://doi.org/10.1007/s11538-015-0098-0

[15] E. Gross, H. Harrington, N. Meshkat, and A. Shiu, "Linear compartmental models: Input-output equations and operations that preserve identifiability," *SIAM Journal on Applied Mathematics*, vol. 79, no. 4, pp. 1423–1447, 2019. [Online]. Available: https://doi.org/10.1137/18M1204826

[16] M. Anguelova, J. Karlsson, and M. Jirstrand, "Minimal output sets for identifiability," *Mathematical Biosciences*, vol. 239, no. 1, pp. 139–153, 2012. [Online]. Available: https://doi.org/10.1016/j.mbs.2012.04.005

[17] C. Brännmark, R. Palmér, S. T. Glad, G. Cedersund, and P. Strålfors, "Mass and information feedbacks through receptor endocytosis govern insulin signaling as revealed using a parameter-free modeling framework," *Journal of Biological Chemistry*, vol. 285, no. 26, pp. 20 171–20 179, 2010. [Online]. Available: https://doi.org/10.1074/jbc.M110.106849

[18] B. Casas, L. Vilén, S. Bauer, K. Kanebratt, C. W. Huldt, L. Magnusson, U. Marx, T. B. Andersson, P. Gennemark, and G. Cedersund, "Integrated experimental-computational analysis of a liver-islet microphysiological system for human-centric diabetes research," *bioRxiv*, 2021. [Online]. Available: https://doi.org/10.1101/2021.08.18.456693

[19] A. Goldbeter, "A model for circadian oscillations in the drosophila period protein (PER)," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 261, no. 1362, pp. 319–324, 1995. [Online]. Available: http://doi.org/10.1098/rspb.1995.0153

[20] T. Lipniacki, P. Paszek, A. R. Brasier, B. Luxon, and M. Kimmel, "Mathematical model of NF-κB regulatory module," *Journal of Theoretical Biology*, vol. 228, no. 2, pp. 195–215, 2004. [Online]. Available: https://doi.org/10.1016/j.jtbi.2004.01.001

[21] J. Karlsson, M. Anguelova, and M. Jirstrand, "An efficient method for structural identifiability analysis of large dynamic systems," *IFAC Proceedings Volumes*, vol. 45, no. 16, pp. 941–946, 2012, 16th IFAC Symposium on System Identification. [Online]. Available: https://doi.org/10.3182/20120711-3-BE-2027.00381

[22] A. F. Villaverde, N. D. Evans, M. J. Chappell, and J. R. Banga, "Sufficiently exciting inputs for structurally identifiable systems biology models," *IFAC-PapersOnLine*, vol. 51, no. 19, pp. 16–19, 2018, 7th Conference on Foundation of Systems Biology in Engineering FOSBE 2018. [Online]. Available: https://doi.org/10.1016/j.ifacol.2018.09.015

[23] ——, "Input-dependent structural identifiability of nonlinear systems," *IEEE Control Systems Letters*, vol. 3, no. 2, pp. 272–277, 2019. [Online]. Available: https://doi.org/10.1109/LCSYS.2018.2868608

[24] E. Hille, *Ordinary Differential Equations in the Complex Domain*, 1997.

[25] F. Anstett-Collin, L. Denis-Vidal, and G. Millérioux, "A priori identifiability: An overview on definitions and approaches," *Annual Reviews in Control*, vol. 50, pp. 139–149, 2020. [Online]. Available: https://doi.org/10.1016/j.arcontrol.2020.10.006

[26] H. Hong, A. Ovchinnikov, G. Pogudin, and C. Yap, "Global identifiability of differential models." [Online]. Available: https://doi.org/10.48550/arXiv.1801.08112

[27] S. Lang, *Algebra*, 3rd ed. New York: Springer, 2002.

[28] R. Ehrenborg and G. Rota, "Apolarity and canonical forms for homogeneous polynomials," *European Journal of Combinatorics*, vol. 14, no. 3, pp. 157–181, 1993. [Online]. Available: https://doi.org/10.1006/eujc.1993.1022

[29] A. Sedoglavic, "A probabilistic algorithm to test local algebraic observability in polynomial time," in *International Symposium on Symbolic and Algebraic Computation*, London, Ontorio, Canada, Jul. 2001. [Online]. Available: https://doi.org/10.1145/384101.384143

[30] F. Zippel, *Effective Polynomial Computation*. Boston: Springer, 1993.

[31] J. von zur Gathen and J. Gerhard, *Modern Computer Algebra*, 3rd ed. Cambridge University Press, 2013.